Vol.11.Issue.3.2023 (July-Sept.)
©KY PUBLICATIONS

BULLETIN OF MATHEMATICS AND STATISTICS RESEARCH

A Peer Reviewed International Research Journal

http://www.bomsr.com

Email:editorbomsr@gmail.com

INTERNATIONAL STANDARD SERIAL NUMBER 2348-0580

RESEARCH ARTICLE

ON THE SELECTION OF PROBABILITY DISTRIBUTONS FOR MODELLING OF DAILY RAINFALL AMOUNT IN NORTHERN ODISHA, INDIA

Tapan Kumar Malik[1], A.K. Mangaraj[2], L.N. Sahoo[3]
[1]Department of Statistics, Bhadrak Autonomous College, Bhadrak 756100, India
E-mail: tapupriti@gmail.com
[2]Department of Statistics, Rajendra University, Balangir 767002, India
E-mail: akmangaraj@gmail.com
[3]Former Professor of Statistics, Utkal University, Bhubaneswar–751004, India
E-mail: lnsahoostatuu@rediffmail.com

DOI:10.33329/bomsr.11.3.71

ABSTRACT

The main theme of this paper is to search for well-suited probability distribution in order to model the distribution of daily rainfall measurements in the Northern Odisha region during monsoon season. The study is focused on evaluating fitting performance of 18 theoretical probability distributions using data on daily rainfall volume for 35 years. Three goodness-of-fit tests *viz*., Kolmogorov-Smirnov, Anderson-Darling and Chi-Squared were deployed to test and compare fitting strength of the selected distributions, and to recognize the most appropriate one.

Keywords: Empirical distribution, goodness-of-fit, Northern Odisha, rainfall modelling, theoretical distribution.

INTRODUCTION

Rainfall is the most essential environmental factor whose incidence pattern affects human activities in diversified ways. Stochastic nature of the rainfall occurrence events makes it difficult for accurate description of the distribution of rainfall system. Modelling of daily rainfall amount using a probability distribution is therefore very essential for proper understanding of the shape of its underlying distribution. Such an attempt on spatiotemporal basis is not only paramount in the fields of hydrology, meteorology, climatology, fisheries, ecology, environment and others but also for efficient planning and execution of water resource management program for agricultural and industrial developments. The technique of stochastic modelling refers to the frequency analysis that involves in selecting a suitable probability distribution and estimating the parameters, investigating various rainfall characteristics for the trend detection, and to forecast rainfall amounts at different probability levels.

The search for a probability distribution that best fits precipitation intensity was the objective of several studies [See, for example, Suhaila and Jemain (2007a, 2007b, 2008), Olofintoye *et al.* (2009), Suhaila *et al*. (2011), Sukla *et al*. (2014), Zeng *et al*. (2015), Amin *et al*. (2016), Ghosh *et al*. (2016), Lei Ye *et al*. (2018), Sreedhar (2019), Baghel *et al*. (2019), Ogarekpe *et al*. (2020), Pina Ximenes *et al*. (2021), Moccia *et al.* (2021) and the references cited therein]. Because by using the said identified distribution, it is possible to predict future events, such as the probability of rain occurring in a given region or time period. But, from the available literature, there is no guarantee that a specific distribution can be considered to have a good fit for all situations. Hence, it is necessary to evaluate many available distributions in order to find the most suitable one that could provide reliable rainfall estimates.

The present investigation considers Northern Odisha region as its study area. Although there is not any clear cut boundary line to define this region, from the administrative point of view five districts are included in North Odisha. But, for studying agro climatic characteristics, the districts are considered under three agro climatic zones: North Western Plateau, North Central Plateau and North Eastern Coastal Plain. However, the three zones are likely to be more or less homogeneous in respect of mean annual rainfall, mean maximum summer and winter temperature, and cropping pattern, crop planning and production. Both North Western Plateau and North Central Plateau are quite similar in respect of their climatic conditions, ecologies and many geographical features, soil type, and irrigation facilities whereas on these grounds North Eastern Coastal Plain is slightly different. North Odisha gets rainfall from the south-west monsoon with an average annual rainfall of 1580 mm and the total rainy day in a year ranging from 50 to 60 days. The most pre-dominant crop in this region is rice covering about 80% of the total farming area. The quantity of annual rainfall received by this region is fairly good. But, its irregular distribution and variation in time and space is a cause of great stress to the farming activities, crop production and crop yield as the agriculture is mostly rain fed. Non availability of assured irrigation facilities further adds uncertainty to the crop production system. An appropriate modelling of the daily rainfall volume is therefore of crucial importance in planning agricultural activities and managing the associated water supply systems at various locations of the study domain.

Our exploration aims to determine the most reliable probability distribution model to fit the observed frequency distribution pattern of daily rainfall amount in the Northern Odisha region by comparing fitting strength of 18 probability distributions. To achieve this objective, we rely on the statistical test score results based on Kolmogorov-Smirnov, Anderson-Darling and Chi-Squared goodness-of-fit tests.

METHODOLOGY

Source and Nature of Data

The study analyses data on daily rainfall amount (mm) of the three meteorological stations – Balasore, Baripada and Kendujhar of the Northern Odisha region for 35 years (1984-2018), obtained from the Meteorological Centre, Bhubaneswar, Odisha. As about 80% of the total annual rainfall occurs during the south western monsoon season (*kharif* season) that is normally limited to four months (June – September), the period considered for the study was taken as 1st June to 30th September (122 days) which also corresponds to the growth season of the paddy crop, the major cash crop in the region. Figures on the daily rainfall amount for the whole study area are obtained by taking average of such figures of the three recording stations. This averaging process will provide proper representative figures on daily rainfall with reduced errors that arising due to random causes [*cf.*, Mooley and Parthasarathy (1984)].

In order to assess distributional pattern of the daily rainfall occurrence, we refer to the computational values of some descriptive statistics such as mean ($\bar{x} = 9.40$ mm), standard deviation ($s = 2.75$ mm), skewness ($\beta_1 = 0.27$) and kurtosis ($\beta_2 = 0.44$) obtained directly from the raw data. The said computational figures give an indication that the distribution of daily rainfall in the study domain during monsoon season is likely to be of positively skewed and platykurtic type with mean 9.40 mm and standard deviation 2.75 mm.

Probability Distributions for Modelling of Rainfall Amount

Selection of a probability distribution that provides an ideal fit to the daily rainfall depth is usually made arbitrarily as there is no unique guideline on this [*cf.*, Murray and Larry (2000), Zeng *et al.* (2015)]. However, exploiting our prior knowledge and past experience gathered from various studies, and based on the shape parametric values computed from the raw data, here we consider 18 probability distributions. Probability density functions ($f(x)$) of the distributions are given below under the assumption that the daily rainfall amount (X) is a continuous random variable.

Beta Distribution

$$f(x) = \frac{1}{B(\alpha_1,\alpha_2)} \frac{(x-a)^{\alpha_1-1}(b-x)^{\alpha_2-1}}{(b-a)^{\alpha_1+\alpha_2-1}};\ \alpha_1, \alpha_2 > 0,\ a \le x \le b;$$

α_1 and α_2 are the shape parameters, and a and b are the boundary parameters ($a < b$).

Dagum Distribution

$$f(x) = \frac{\alpha k \left(\frac{x}{\beta}\right)^{\alpha k-1}}{\beta\left[1+\left(\frac{x}{\beta}\right)^{\alpha}\right]^{k+1}};\ k, \alpha, \beta > 0,\ 0 \le x < \infty;$$

k and α are the shape parameters, and β is the scale parameter.

Erlang Distribution

$$f(x) = \frac{(x-\gamma)^{m-1}}{\beta^m \Gamma(m)} \exp\left[-\frac{x-\gamma}{\beta}\right];\ \gamma \le x < +\infty,\ \beta > 0;$$

m is the shape parameter, β is the scale parameter and γ is the location parameter.

Exponential Distribution

$$f(x) = \lambda \exp[-\lambda(x-\gamma)];\ \lambda > 0,\ \gamma \le x < +\infty;$$

λ is the inverse scale parameter, and γ is the location parameter.

Frechet Distribution

$$f(x) = \frac{\alpha}{\beta}\left(\frac{\beta}{x}\right)^{\alpha+1} \exp\left[-\left(\frac{\beta}{x}\right)^{\alpha}\right];\ \alpha, \beta > 0,\ 0 \le x < +\infty;$$

α is the shape parameter, and β is the scale parameter.

Gamma Distribution

$$f(x) = \frac{x^{\alpha-1}}{\beta^{\alpha}\Gamma(\alpha)} \exp\left[-\left(\frac{x}{\beta}\right)\right];\ \alpha, \beta > 0,\ 0 \le x < +\infty;$$

α is the shape parameter, and β is the scale parameter.

Generalized Extreme Value Distribution

$$f(x) = \begin{cases} \frac{1}{\sigma} \exp\left[-(1+kx)^{-\frac{1}{k}}\right](1+kx)^{-1-\frac{1}{k}}, k \neq 0 \\ \frac{1}{\sigma} \exp[-x - \exp(-x)], k = 0 \end{cases};\ \sigma > 0, 1 + k\frac{(x-\mu)}{\sigma} > 0 \text{ for } k \neq 0 \text{ and}$$

$-\infty < x < +\infty$ for $k = 0$; k is the shape parameter, σ is the scale parameter, and μ is the location parameter.

Gumbel Minimum Distribution

$$f(x) = \frac{1}{\sigma} \exp\left[\frac{x-\mu}{\sigma} - \exp\left(\frac{x-\mu}{\sigma}\right)\right];\ \sigma > 0,\ -\infty < x < +\infty;$$

σ is the scale parameter and μ is the location parameter.

Inverse Gaussian Distribution

$$f(x) = \sqrt{\frac{\lambda}{2\pi x^3}} \exp\left[-\frac{\lambda(x-\mu)^2}{2\mu^2 x}\right];\ \lambda, \mu > 0,\ 0 \le x < +\infty;$$

λ and μ are the parameters.

Laplace Distribution

$$f(x) = \frac{\lambda}{2} \exp(-\lambda|x-\mu|);\ \lambda > 0,\ -\infty < x < +\infty;$$

λ is the inverse scale parameter, and μ is the location parameter.

Log-Logistic Distribution

$$f(x) = \frac{\alpha}{\beta}\left(\frac{x}{\beta}\right)^{\alpha-1}\left[1 + \left(\frac{x}{\beta}\right)^{\alpha}\right]^{-2};\ \alpha, \beta > 0,\ 0 \le x < +\infty;$$

α is the shape parameter, and β is the scale parameter.

Logistic Distribution

$$f(x) = \frac{\exp(-x)}{\sigma[1+\exp(-x)]^2};\ \sigma > 0,\ -\infty < x < +\infty;$$

σ is the scale parameter, and μ is the location parameter.

Lognormal Distribution

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{\ln x - \mu}{\sigma}\right)^2\right];\ \sigma > 0,\ 0 \le x < +\infty;$$

σ and μ are the parameters.

Normal Distribution

$$f(x) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right];\ \sigma > 0,\ -\infty < x < +\infty;$$

σ is the scale parameter, and μ is the location parameter.

Pearson 5 Distribution

$$f(x) = \frac{\exp\left(-\frac{\beta}{x}\right)}{\beta\,\Gamma(\alpha)\left(\frac{x}{\beta}\right)^{\alpha+1}};\ \alpha, \beta > 0,\ 0 \le x < +\infty;$$

α is the shape parameter, and β is the scale parameter.

Pearson 6 Distribution

$$f(x) = \frac{\left(\frac{x}{\beta}\right)^{\alpha_1 - 1}}{\beta\,B(\alpha_1, \alpha_2)\left(1 + \frac{x}{\beta}\right)^{\alpha_1 + \alpha_2}};\ \alpha_1, \alpha_2, \beta > 0,\ 0 \le x < +\infty;$$

$\alpha_1$and α_2 are the shape parameters, and β is the scale parameter.

Rayleigh Distribution

$$f(x) = \frac{x-\gamma}{\sigma^2}\exp\left[-\frac{1}{2}\left(\frac{x-\gamma}{\sigma}\right)^2\right]; \sigma > 0,\ \gamma \le x < +\infty$$

σ is the scale parameter and γ is the location parameter.

Weibull Distribution

$$f(x) = \frac{\alpha}{\beta}\left(\frac{x}{\beta}\right)^{\alpha-1}\exp\left[-\left(\frac{x}{\beta}\right)^{\alpha}\right];\ \alpha, \beta > 0,\ 0 \le x < +\infty;$$

α is the shape parameter, and β is the scale parameter.

Note that thirteen 2-parameter, four 3-parameter and one 4-parameter distributions were taken into account for assessing and comparing their fitting competence.

The Goodness-of-Fit Tests

Three goodness of fit (GOF) tests *viz*., Kolmogorov-Smirnov (KS), Anderson-Darling (AD) and Chi-Square (CS) tests at 5% level of significance have been engaged to identify the best model/distribution among the fitted distributions to the daily rainfall data of the study area. Denoting cumulative distribution function (CDF) of the random variable X by $F(\cdot)$, and $x_1, x_2, \ldots, x_n$ as n sample observations, the GOF test procedures are outlined as follows:

***The Kolmogorov-Smirnov Test*:** The test is based on the largest vertical difference between the theoretical CDF $F_0(x)$ and the observed (empirical) CDF $F_n(x)$ of the random sample of n observations. Under the null hypothesis $H_0: F_0(x) = F_n(x)$ *vs.* alternative hypothesis $H_1: F_0(x) \neq F_n(x)$, the KS test statistic is given by

$$d_{max} = \max_{x}|F_n(x) - F_0(x)|.$$

The null hypothesis is rejected at 5% level of significance if the calculated value of d_{max} exceeds the tabulated value $D_{0.05} = 1.36/\sqrt{n}$.

***The Anderson-Darling Test*:** The AD test measures the fitting strength of an observed CDF to an expected (theoretical) CDF by giving more weight to the tail of the distribution than the KS test. The test statistic (A^2) is defined by

$$A^2 = -n - \frac{1}{n}\sum_{i=1}^{n}(2i-1)[\ln F(x_i) + \ln(1 - F(x_{n-i+1}))].$$

In the present study, the critical values for the distribution of A^2 are not available directly at the whole range of series length. However, we use these values for our purpose as obtained by simulations given in Zeng *et al.* (2015).

***The Chi-Squared Test*:** The CS test conveniently compares to what extent the theoretical distribution fits the empirical distribution. The test statistic is defined by

$$\chi^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i},$$

where $k =$ number of classes, O_i and $E_i = F(x_2) - F(x_1)$ are respectively observed and expected frequencies for bin i with x_1 and x_2 as the lower and upper limits. Critical values for rejecting/accepting the null hypothesis that the agreement between the theoretical and empirical frequency distributions, are readily available in the table.

ANALYSIS OF DATA

For the purpose of analyzing recorded data that includes estimation of relevant parameters, computation of test statistics, fitting of the probability distributions to the data, identification of the most suitable model, and interpretation of the investigation results, we used the Software *EasyFit 5.6 Professional*. According to this Software, analysis is performed after classifying the raw data on daily rainfall figures into seven continuous class intervals and then constructing observed (empirical) frequency distribution for the study region.

Results and Model Selection

Parameters of the candidate distributions are estimated by the method of maximum likelihood and their estimated values are presented in Table-1. On the basis of the estimated parametric values, the aforesaid eighteen distributions were fitted to the data.

Table-1: Estimated Parameters of Fitted Distributions

Distribution	Parameters	Distribution	Parameters
Beta	$\alpha_1 = 27.23, \alpha_2 = 96.50,$ $a = -6.852, b = 67.04$	Laplace	$\lambda = 0.514, \mu = 9.401$
Dagum	$k = 0.434, \alpha = 8.596, \beta =$ 11.01	Log-Logistic	$\alpha = 5.425, \beta = 8.919$
Erlang	$m = 64, \beta = 0.341, \gamma =$ -12.59	Logistic	$\sigma = 1.517, \mu = 9.401$
Exponential	$\lambda = 0.165, \gamma = 3.331$	Lognormal	$\sigma = 0.316, \mu = 2.194$
Frechet	$\alpha = 3.5947, \beta = 7.6272$	Normal	$\sigma = 2.751, \mu = 9.401$
Gamma	$\alpha = 11.68, \beta = 0.8049$	Pearson 5	$\alpha = 9.483, \beta = 80.66$
Gen. Extreme Value	$k = -0.237, \sigma = 2.662,$ $\mu = 8.382$	Pearson 6	$\alpha_1 = 11.47, \alpha_2 = 3447.7,$ $\beta = 2822.7$
Gumbel Min	$\sigma = 2.145, \mu = 10.64$	Rayleigh	$\sigma = 4.810, \gamma = 3.174$
Inv. Gaussian	$\lambda = 109.8, \mu = 9.401$	Weibull	$\alpha = 3.918, \beta = 10.29$

Three GOF tests mentioned above are carried out in accordance with standard procedure. The goal is to evaluate and compare the fitting ability of the comparable distributions by testing the hypothesis that there is no disagreement between observed and theoretical (expected) frequency distributions. The test statistics for all GOF tests were computed and tested at 5% level of significance. Fitted probability distributions were awarded ranks from 1 to 18 according to the minimum value of the calculated test statistic. When a test was failed to fit a distribution to the data set *i.e.*, the computed value of the test statistic exceeded the significant value, the concerned distribution was given no rank. Results on the three GOF tests are summarized in Table-2.

Table-2: Goodness-of-Fit Summary

No.	Distribution	KS Test		AD Test		CS Test	
		Statistic	Rank	Statistic	Rank	Statistic	Rank
1	Beta	0.0411	4	0.20678	3	2.5109	3
2	Dagum	0.03608	1	0.15786	1	1.5966	1
3	Erlang	0.0593	8	0.38177	7	3.5321	8
4	Exponential	0.29225	18	16.264	-	88.366	-
5	Frechet	0.15262	17	5.8792	-	18.208	-
6	Gamma	0.05549	7	0.58205	8	3.2607	7
7	Gen. Extreme Value	0.04311	6	0.25252	6	2.6286	4
8	Gumbel Min	0.10385	14	4.3774	-	8.9474	15
9	Inv. Gaussian	0.06286	10	1.3459	13	5.8881	13
10	Laplace	0.07309	11	1.0313	10	5.3452	12
11	Log-Logistic	0.08014	12	1.2378	12	3.94	10
12	Logistic	0.04066	2	0.22205	4	2.0597	2
13	Lognormal	0.08245	13	1.2303	11	4.3868	11
14	Normal	0.0407	3	0.18874	2	3.189	6
15	Pearson 5	0.10679	15	2.2337	15	9.0035	16
16	Pearson 6	0.05938	9	0.59668	9	2.911	5
17	Rayleigh	0.12306	16	2.1128	14	8.5855	14
18	Weibull	0.04141	5	0.25072	5	3.5903	9

Referring to the results provided in Table-2, we see that it is not only difficult to sort out the best fit distribution but also controversial. Because, the rank of some particular distribution for one test category is different from that of other category. For example, Normal distribution is ranked in the second position under AD test but the distribution is ranked in third and sixth places under the KS and CS tests respectively. Engagement of three different GOF criteria leaves us wondering which test could provide the best possible result. For this scenario, following Suhaila and Jemain (2007a, 2007b), we tried to pick out the best fitting distribution based on the majority of the test procedures as we have no option to investigate which GOF test is the most powerful. With this objective, we have selected three distributions holding the first three ranks with respect to all three tests independently. This leads to prefer 4 distributions: Beta, Dagum, Logistic and Normal as the competing candidate distributions. In Figure-1, we display graphical representations of the said probability distributions fitted to the observed histograms of the daily rainfall amount.

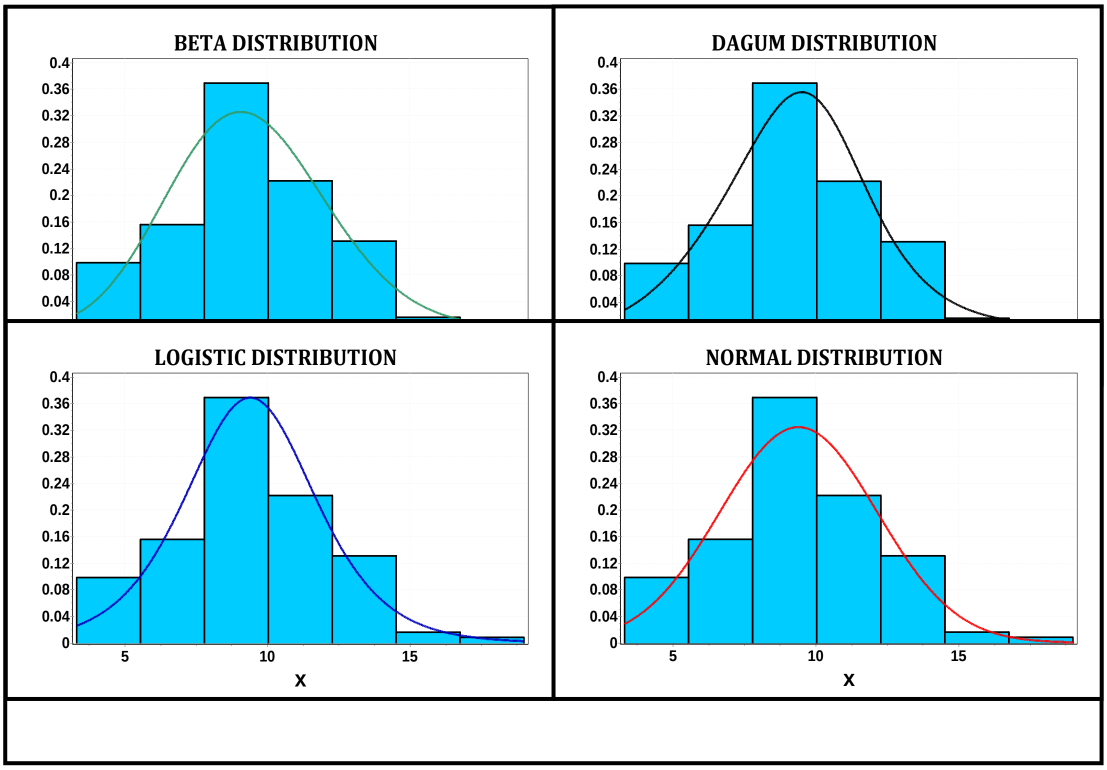

Figure-1: Probability Distribution Curves

Assessment of the Best Probability Model

On consideration of fitting strength in respect of majority of tests as a benchmark of the best fit and closeness of the expected frequency distribution curve to the mid values of the upper sides of the observed histogram rectangles, we select Dagum as the first candidate distribution for modelling daily rainfall volume. On the other hand, Logistic may be sorted out as the second best fitted distribution.

For better reconciliation on the issue of the best fit model identification, further assessment of all probability models were made taking into account the total test score attained from the three GOF tests [*cf.*, Murray and Larry (2000), Olofintoye *et al.* (2009), Sharma and Singh (2010)]. Test scores ranging from 0 to 18 is awarded to each distribution model based on the achieved rank on the concerned test. The distribution best supported by a test *i.e.*, ranked 1 is awarded a score 18, the next best *i.e.*, ranked 2 is awarded 17, and so on in descending order. A distribution is awarded a zero score for a test if the distribution does not have a significant fit to the data. The overall scoring results are presented in Table-3 to recognize the best fit distribution according to the highest score obtained. Based on the tabular results, Dagum and Logistic distributions were found to be the best and second best performed models respectively among the eighteen probability distributions tested for the study domain. It is also interesting to see that these findings based on the total test score are in accordance with those based on the majority of tests and graphic representations.

Table-3: Summary of the Statistical Score Results

Distribution	Total Score	Distribution	Total Score
Beta	47	Laplace	24
Dagum	54	Log-Logistic	23
Erlang	34	Logistic	49
Exponential	01	Lognormal	22
Frechet	02	Normal	46
Gamma	35	Pearson 5	11
Gen. Extreme Value	41	Pearson 6	34
Gumbel Min	09	Rayleigh	13
Inv. Gaussian	21	Weibull	38

CONCLUSION

In the light of our data analysis and three scientific assessment procedures *viz.*, fitting ability in respect of majority of GOF tests, total GOF test scores and graphical method undertaken in this paper, Dagum and Logistic were respectively emerged out as the best and second best fitted distributions for the Northern Odisha region. Although from the graphs a choice between Beta and Normal is not clear cut, on the other two grounds, the Beta distribution may be recommended as the third best fit. However, the present investigation leads to an overall conclusion that the Dagum distribution is one of the best probability models for describing the distribution of daily rainfall volume. Appropriate planning and hydrological design of soil conservation and drainage structures at the Northern Odisha region can therefore be effectively carried out on the basis of predicted amount of daily rainfall using the said distribution.

REFERENCES

[1]. Amin, M. T., Rizwan, M. and Alazba, A.A. (2016). A best-fit probability distribution for the estimation of rainfall in northern regions of Pakistan. *Open Life Science*, 11, 432–440.

[2]. Baghel, H., Mittal, H.K., Singh, P.K., Yadav, K.K. and Jain, S. (2019). Frequency analysis of rainfall data using probability distribution models. *International Journal of Current Microbiology and Applied Sciences*, 8(6), 1390-1396.

[3]. Ghosh, S., Roy, M. K., and Biswas, S. C. (2016). Determination of the best fit probability distribution for monthly rainfall data in Bangladesh. *American Journal of Mathematics and Statistics*, 6(4), 170– 174.

[4]. Lei Ye, Hanson, L.S., Ding, P., Wang, D., and Vogel, R.M. (2018). The probability distribution of daily precipitation at the point and catchment scales in the United States. *Hydrology and Earth System Science*, 22, 6519–6531.

[5]. Moccia, B., Mineo, C., Ridolfi, E., Russo, F. and Napolitano, F. (2021). Probability distributions of daily rainfall extremes in Lazio and Sicily, Italy, and design rainfall inferences. *Journal of Hydrology: Regional Studies*, 33, 100771.

[6]. Mooley, D.A., Parthasarathy, B. (1984). Fluctuations in All-India summer monsoon rainfall during 1871-1978. *Climatic Change*, 6, 287-301.

[7]. Murray, R. S. and Larry, J. S. (2000). *Theory and Problems of Statistics* (3rd Edition). Tata Mc Graw-Hill Publishing Company Ltd., New Delhi.

[8]. Ogarekpe, N.M., Tenebe, I.T., Emenike, P.C., Obianuju A Udodi, O.A. and Antigha, R.E. (2020). Assessment of regional best-fit probability density function of annual maximum rainfall using CFSR precipitation data. *Journal of Earth System Science*, 129:176, 1-18.

[9]. Olofintoye, O. O., Sule, B. F. and Salami, A. W. (2009). Best-fit probability distribution model for peak daily rainfall of selected cities in Nigeria. *New York Science Journal*, 2, 1-12.

[10]. Pina Ximenes, P.S.M., Alves da Silva, A.S., Ashkar, F. and Stosic, T. (2021). Best-fit probability distribution models for monthly rainfall of North-eastern Brazil. *Water Science and Technology*, 84(6), 1541 doi: 10.2166/wst.2021.304.

[11]. Sharma, M. A. and Singh, J. B. (2010). Use of probability distribution in rainfall analysis. *New York Science Journal*, 3, 40-49.

[12]. Sreedhar, B.R. (2019). Fitting of probability distribution for analysing the rainfall data in the state of Andhra Pradesh, India. *International Journal of Applied Engineering Research*, 14(3), 835-839.

[13]. Suhaila, J. and Jemain, A. A. (2007a). Fitting daily rainfall amount in peninsular Malaysia using several types of exponential distributions. *Journal of Applied Sciences Research*, 3, 1027-1036.

[14]. Suhaila, J. and Jemain, A. A. (2007b). Fitting the statistical distributions to the daily rainfall amount in peninsular Malaysia. *Jurnal Teknologi*, 46, 33-48.

[15]. Suhaila, J. and Jemain, A. A. (2008). Fitting the statistical distributions for daily rainfall in peninsular Malaysia based on AIC criterion. *Journal of Applied Science Research*, 4, 1846-1857.

[16]. Suhaila, J., Ching-Yee, K. and Hui-Mean, Y.F.F. (2011). Introducing the mixed distribution in fitting rainfall data. *Open Journal of Modern Hydrology*, 1, 11-22.

[17]. Sukla, M. K., Mangaraj, A. K. and Sahoo, L. N. (2014). An investigation on the stochastic modelling of daily rainfall amount in the Mahanadi Delta region, India. *Research Journal of Mathematical and Statistical Sciences*, 2(9), 1-8.

[18]. Zeng, X., Wang, D. and Wu J. (2015). Evaluating the three methods of goodness of fit test for frequency analysis. *Journal of Risk Analysis and Crisis Response*, 5(3), 178–187.