

Probability Distributions to Describe Rainfall Pattern in North-Western and Eastern Coasts of Sri Lanka

W.W.U.I. Wickramaarachchi¹, T.U.S. Peiris² and S. Samita³

¹Postgraduate Institute of Agriculture,
University of Peradeniya, Peradeniya, 20400, Sri Lanka.
Email: upzwickramaarachchi@gmail.com

² Biostatistics Unit, Faculty of Livestock,
Fisheries and Nutrition, Wayamba University of Sri Lanka,
Makandura, Gonawila, 60170, Sri Lanka.
Email : ursla@wyb.ac.lk

³Postgraduate Institute of Agriculture, University of Peradeniya,
Peradeniya, 20400, Sri Lanka.
Email : ssamita@pdn.ac.lk

ABSTRACT

The changes in the global rainfall have raised concerns over the present and future use of coastal areas in many Asian countries including Sri Lanka. The coasts of the North-western and Eastern provinces of Sri Lanka harbor varied important dynamic industries such as aquaculture, agriculture, and tourism industry. However, no in-depth analysis of rainfall of coastal lines of Sri Lanka has been performed in the past to understand any changes in the pattern in general and to mitigate the effects of those changes on these sectors. In this study, an attempt has been made to determine the best-fit distribution from the monthly rainfall data spanning 30 years of observations from 1987 to 2016 at 19 rain gauge stations as an initial step to identify rainfall patterns and future shifts. To fit the observed data, 45 different distributions were considered. The efficacy of the fits for these distributions was evaluated using three empirical nonparametric goodness-of-fit tests, namely Kolmogorov–Smirnov, Anderson–Darling, and Chi-square test. Results showed that the Johnson SB was the best-fit distributions in stations of North-western province and Gen. Pareto, Generalized Extreme Value distribution, and Frechet (3P) were the best-fitted distributions for Eastern province. The results of this study would be useful to the water resource engineers, policymakers, and planners for the agricultural development and conservation of natural resources of North-western and Eastern coastal lines.

Keywords: Climate change, coastal lines, rainfall analysis, probability distributions.

Mathematics Subject Classification Number: 62P12, 11K99

1. INTRODUCTION

In many countries, detailed statistical studies of rainfall data have been carried out for past decades with fits to multiple probability distributions (Amin et al., 2016, Mayooraan and Laheetharan, 2014, Esberto, 2018 and Sukrutha et al., 2018). In the Sri Lankan context, published literature on fitting probability distribution to rainfall data are limited. Mayooraan and Laheetharan (2014) have performed probability modeling of maximum daily rainfall in Colombo District and revealed that the best-fit probability distributions for different data sets are different. According to them, the log-Pearson 3 (3P) and Burr (4P) distributions were found as the best-fitted probability distribution models for the annual

and first-inter monsoon season, respectively. The Generalized extreme value was observed as the best-fitted probability distribution model for the remaining three rainy seasons.

A large number of statistical studies have been done for rainfall in other neighbouring Asian countries like India (Sukrutha et al., 2018), Bangladesh (Alam et al., 2018 and Khudri and Sadia, 2013), Philippines (Esberto, 2018) and Pakistan (Amin et al., 2016). These studies are not limited to one continent and have been widely distributed and for brevity, it is only mentioned a few selected studies to illustrate the range in the best-fit distribution found from these studies.

Studies also showed that the applicability of different probability distributions is found to be associated with the spatial differences in the study area. For example, Generalized extreme value (GEV) provides better performance for catchments with medium to high mean annual precipitation, while the 3-parameter Lognormal distribution is probably a more appropriate choice for dry intermediate-sized catchments (Salinas et al., 2013). Moreover, the complexity of probability distributions may contribute to model performance. Jamaludin & Jemain, (2007) have assessed the performances of the mixed Gamma, mixed Weibull, and mixed Exponential ones together with single parameter distributions in Peninsular Malaysia and found the mixture distributions to be better than the single distributions in modelling precipitation amounts. Hence, it is clearly demonstrated that no single distribution can accurately describe the rainfall distribution, and the applicability of these probability distributions are location specific. The characteristics of available rainfall data and the statistical tools used for the model selection procedure determine the best-fit-probability distribution.

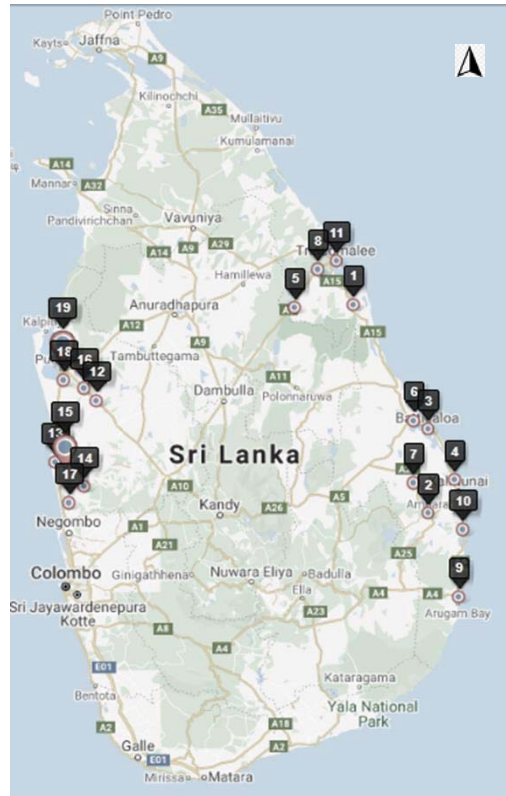
The probability models; generalized extreme value, gamma, Gumbel max, inverse Gaussian, log normal, normal, Pearson, Weibull have been widely used for evaluating the best-fitted probability distribution for rainfall (Sukrutha et al., (2018) and Esberto, (2018)) while others are not much used in similar studies except in a study conducted by Mayooraan & Laheetharan, (2014) to detect the statistical distribution of annual maximum rainfall in Colombo District, Sri Lanka. Each kind of probability distribution has its own applicability and limitations. Assessing the applicability of different distribution models is necessary to provide more accurate precipitation information. Furthermore, the selection of an appropriate probability distribution is of prime importance for further analysis of rainfall events.

The considered two coastal lines; North-western and Eastern harbour many important industries related to agriculture, aquaculture, and tourism all of which rely on the weather. Therefore, determination of underlying probability distribution/s to monthly rainfall data is urgently needed. Further, this is important to reveal the applicability of statistical methods for narrowing the knowledge gap on rainfall distribution in these areas. Thus, the aim of this study was to show that the rainfall distribution is highly localized. Besides, if a climate interdependent industry is planned, it is essential to study the rainfall distribution in that specific location.

2. STUDY AREA AND METHODS

Monthly rainfall values from 19 meteorological and agro-meteorological stations representing the Eastern province (Allai tank¹, Ampara², Batticaloa³, Kallar⁴, Kantale⁵, Mylambaveli⁶, Navakiri Aru⁷,

Palampoddaru⁸, Pottuvil⁹, Sagamam Kulam¹⁰, Trincomalee¹¹) and North-western province (Anamaduwa¹², Chilaw¹³, Horakele¹⁴, Karukkuwa¹⁵, Kottukachchiya¹⁶, Lunuwila¹⁷, Palaviya¹⁸, Puttalam¹⁹) (Plate 1) for the period of 1987–2016 were used in this study.



In this study, the list of probability distributions considered for fitting the rainfall data were beta, Burr (3P, 4P), Cauchy, chi-square (1P, 2P), Dagum (4P), Erlang (2P, 3P), error function, exponential (1P, 2P), fatigue life (3P), Frechet (2P), gamma (2P, 3P), generalized gamma (3P, 4P), generalized extreme value, Gumbel max, inverse Gaussian (2P, 3P), Kumaraswamy, Johnson SB, Laplace, Levy, log-gamma, logistic, log-logistic (2P, 3P), lognormal (2P, 3P), log-Pearson 3 (3P), Nakagami, normal, Pareto(first kind), Pareto(second kind), Pearson type 5 (2P, 3P), Pearson type 6 (3P, 4P), Student's t, Weibull (2P, 3P). The description of the 44 probability distribution models that were used in the study exhibited in Appendix 1. The maximum-likelihood analysis was used to find the best-fit parameters for each of these probability distributions and for each station. To select the best-fit distribution for a given station and to rank each distribution for every station, multiple model comparison techniques were used. We now describe the model comparison techniques used for the study.

2.1. Model comparison tests

Multiple model comparison methods were employed to carry out hypothesis testing and selected the best distribution for the rainfall data. For this purpose, the goodness-of-fit tests used include nonparametric distribution-free tests such as the Kolmogorov–Smirnov test Anderson–Darling test,

and Chi-square test. For each of the probability distributions, the best-fit parameters were estimated for each of the rain gauge stations using least-squares fitting and then model comparison tests were carried out at a 5% significance level for the selection of the best-fit distribution. The goodness of fit tests were used for testing the following hypotheses:

H_0 : The monthly rainfall data follow the specified distribution

H_A : The monthly rainfall data does not follow the specified distribution

(a) Kolmogorov–Smirnov test (K–S)

The Kolmogorov–Smirnov (K–S) test is a goodness-of-fit statistic that compares an empirical distribution function $F_{(x)}$, with a specified distribution function $F_{(y)}$.

The Kolmogorov–Smirnov statistic (D) can be computed as:

$$D = \max |F_x(x) - F_y|$$

The K-S test statistic, the supreme class of EDF statistics, is based on the maximum vertical difference between the theoretical and empirical distributions (Conover, 1999).

(b) Anderson–Darling test (A–D)

The Anderson–Darling (A–D) test was introduced by Anderson and Darling (1954) to place more weight or discriminating power at the tails of the distribution (Farrel et al., 2006). This can be important when the tails of the selected theoretical distribution are of practical significance. The test statistic (AD) is defined as:

$$AD = -n - \frac{1}{n} \sum_{i=1}^n (2i-1)(\ln(x_i) + \ln(1 - (x_{(n+1-i)})))$$

where $\{x_{(1)} < \dots < x_{(n)}\}$ the ordered (from smallest to the largest element) sample of size n , and $F_{(x)}$ is the underlying theoretical cumulative distribution to which the sample is compared. The null hypothesis that $\{x_{(1)} < \dots < x_{(n)}\}$ comes from the underlying distribution $F_{(x)}$ is rejected if AD is larger than the critical value α at a given significance level α .

(c) Chi-square (X^2) test

This is a technique that checks if a specific distribution of a certain observed event's frequency in a sample is suitable for that sample or not. Using O to define “observed count” and E to define “expected count”, the Chi-square test statistic is calculated by:

$$\chi^2 = \frac{\sum (O - E)^2}{E}$$

The calculations were implemented with EasyFit 5.6 software. EasyFit software performs an unconstrained nonlinear minimization of the sum of squared residuals (SSR) with respect to the various parameters and select the best distribution by minimizing the SSR (Schittkowski, 2002).

3. RESULTS

3.1. Effects of the various factors on the coefficient R_p^2

The basic statistics for the annual rainfall data for the selected agro-meteorological stations are summarized in Table 1, where the minimum, maximum, mean, standard deviation (SD), coefficient of variation (CV), coefficient of skewness and kurtosis are shown.

Table 1: Summary statistics of annual rainfall data for the selected 19 stations during the years (1987–2016).

Station	Min Value (mm)	Max Value (mm)	Mean (mm)	Standard Deviation (mm)	Coefficient of Variance (%)	Coefficient of Skewness	Kurtosis
<u>North-western Coastal Line</u>							
Anamaduwa	751	1807	1215.6	270.1	22.22	0.497	-0.228
Chillaw	401	2054	1168.2	393.3	33.66	-0.119	0.031
Horakelle	852	1972	1438.8	260.5	18.10	-0.446	0.355
Karukkuwa	905	2341	1440.7	337.6	23.42	0.621	0.397
Kottikachchuwa	750	1889	1187.8	280.1	23.57	0.472	0.064
Lunuvila	1055	2450	1787.3	316.3	17.69	-0.516	0.352
Palaviya	702	1597	1101.8	256.0	23.23	0.454	-0.750
Puttlam	814	1696	1164.8	241.5	20.73	0.566	-0.301
<u>Eastern Costal Line</u>							
Allai tank	730	3352	1357.6	498.0	36.68	2.198	8.089
Ampara	1118	3255	1771.9	458.6	25.88	1.200	2.462
Batticaloa	1007	3581	1794.2	562.5	31.35	1.489	2.849
Kallar	1000	3192	1565.6	487.7	31.15	1.563	2.859
Kantale	1045	3065	1673.8	443.4	26.49	1.321	2.226
Mylambaveli	1235	3079	1795.6	433.2	24.12	1.303	2.021
Navakiri Aru	589	3885	1752.8	703.7	40.15	1.687	3.606
Palampoddaru	668	2956	2104.3	629.2	29.90	-0.502	-0.479
Pottuvil	857	2196	1433.5	366.0	25.53	0.363	-0.844
SagamamKulam	634	2628	1412.6	459.3	32.51	0.686	0.419
Trincomalee	996	2910	1580.9	389.1	24.61	1.226	3.321

3.2. Determination of the best-fit probability distribution for each rain gauge station

The best-fit test statistics (K–S test (D), Anderson–Darling test (AD) and Chi-square test (X²)) were computed for the 44 probability distributions. In most of the cases, the three goodness of fit tests yielded different types of distribution. To conclude, for selecting the best-fit distribution for every station, each fitted probability distributions (45) were ranked according to the ascending order of test statistics of 03 Goodness of Fit (GoF) tests. Then, the ranks were summed and the probability distribution with the lowest summed rank was selected as the best-fitted probability distribution for

each location (Sukrutha et al., 2017). The best-fit distribution for each station using this ranking technique is given in Table 2.

Table 2: Station-wise best ranked probability distribution using different goodness-of-fit tests

Station	KS	AD	Chi-square
<u>North-Western Coastal Line</u>			
Anamaduwa	Gen. Gamma	Gen. Pareto	Gen. Gamma
Chillaw	Pareto 2	Gen. Extreme Value	Gumbel Max
Horakelle	Johnson SB	Johnson SB	Frechet
Karukkuwa	Johnson SB	Gen. Gamma	Beta
Kottikachchuwa	Johnson SB	Johnson SB	Gamma
Lunuvila	Johnson SB	Johnson SB	Beta
Palavi	Dagum	Log-Pearson 3	Fatigue Life
Puttlam	Frechet (3P)	Log-Pearson 3	Dagum
<u>Eastern Costal Line</u>			
Allai tank	Gen. ExtremeValue	Gen. Pareto	Johnson SB
Ampara	Frechet	Burr	Frechet
Batticaloa	Lognormal(3P)	Burr	Error Function
Kallar	Gamma	Gen. Pareto	Weibull
Kantale	Johnson SB	Johnson SB	Johnson SB
Mylambaveli	Log-Logistic (3P)	Burr	Log-Pearson 3
Navakiri Aru	Lognormal	Log-Pearson 3	Gen. Gamma
Palampoddaru	Gamma (3P)	Log-Pearson 3	Burr
Pottuvil	Log-Pearson 3	Log-Pearson 4	Cauchy
SagamamKulam	Johnson SB	Gen. Pareto	Log-Logistic
Trincomalee	Lognormal	Log-Pearson 3	Levy (2P)

Consequently, the ranking of different probability distributions was obtained from 1 to 45 based on minimum test statistic value. The probability distribution having the first rank was selected for all the three tests independently, which fits the rainfall data using multiple model comparison techniques.

Table 3: Station-wise best fit distribution and their parameters obtained by summing the ranks of each of the distributions from all the model comparison tests considered in Table 2.

Station	Best-fit Probability Distribution	Parameters (MLE)
<u>North-Western Coastal Line</u>		
Anamaduwa	Gen. Gamma	$k=0.9 \quad \alpha=0.7 \quad \beta=76.3$
Chillaw	Gen. Extreme Value	$k=0.4 \quad \sigma=24.2 \quad \mu=15.42$
Horakele	Johnson SB	$\gamma=1.0 \quad \delta=0.55, \lambda=284.21 \quad =-1.61$
Karukkuwa	Johnson SB	$\gamma=0.70 \quad \delta=0.48, \lambda=189.57 \quad =-1.78$
Kottikachchuwa	Johnson SB	$\gamma=1.16 \quad \delta=0.44, \lambda=278.51 \quad =-2.38$
Lunuvila	Johnson SB	$\gamma=0.63 \quad \delta=0.47, \lambda=242.69 \quad =3.67$
Palaviya	Log-Pearson 3	$\alpha=1058.4 \quad \beta=-0.055 \quad \gamma=61.30$
Puttlam	Frechet (3P)	$\alpha=0.809 \quad \beta=10.093 \quad \gamma=-0.752$
<u>Eastern Costal Line</u>		

Allai tank	Gen. Pareto	$k=0.37$ $\sigma=29.032$ $\mu=-7.38$
Ampara	Frechet (3P)	$\alpha=1.01$ $\beta=22.06$ $\gamma=4.22$
Batticaloa	Levy (2P)	$\sigma=9.58$ $\gamma=1.85$
Kallar	Gen. Extreme Value	$k=0.55$ $\sigma=21.01$ $\mu=9.03$
Kantale	Johnson SB	$\gamma=1.01$ $\delta=0.48$ $\lambda=276.71$ $=-0.76$
Mylambaveli	Dagum	$k=27.5$ $\alpha=1.2$ $\beta=2.2$
Navakiri Aru	Beta	$\alpha_1=0.34873$ $\alpha_2=0.23422$ $\alpha=1.0$ $\beta=12.0$
Palampoddaru	Burr	$k=392.12$ $\alpha=0.65$ $\beta=4.32$
Pottuvil	Frechet (3P)	$\alpha=1.14$ $\beta=25.96$ $\gamma=-8.35$
SagamamKulam	Gen. Pareto	$k=0.278$ $\sigma=51.54$ $\mu=-14.8$
Trincomalee	Gen. Pareto	$k=0.28$ $\sigma=36.27$ $\mu=-2.6$

For conciseness in this manuscript, the plots of probability density function for the best-fit distribution for 04 selected stations is given in Figure1 for the Eastern Province and 04 selected stations is given in Figure 2 for the North-western Province.

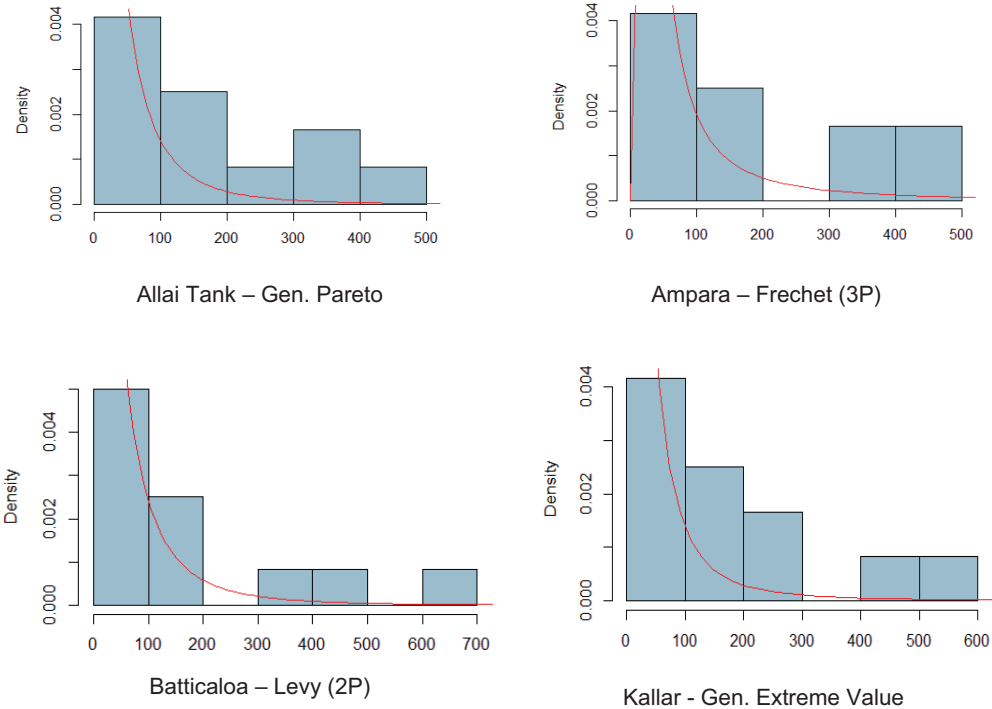


Figure 1. Histogram of the monthly rainfall data at 04 selected rain gauge stations in Eastern Province along with best fit probability distribution for each locations considered.

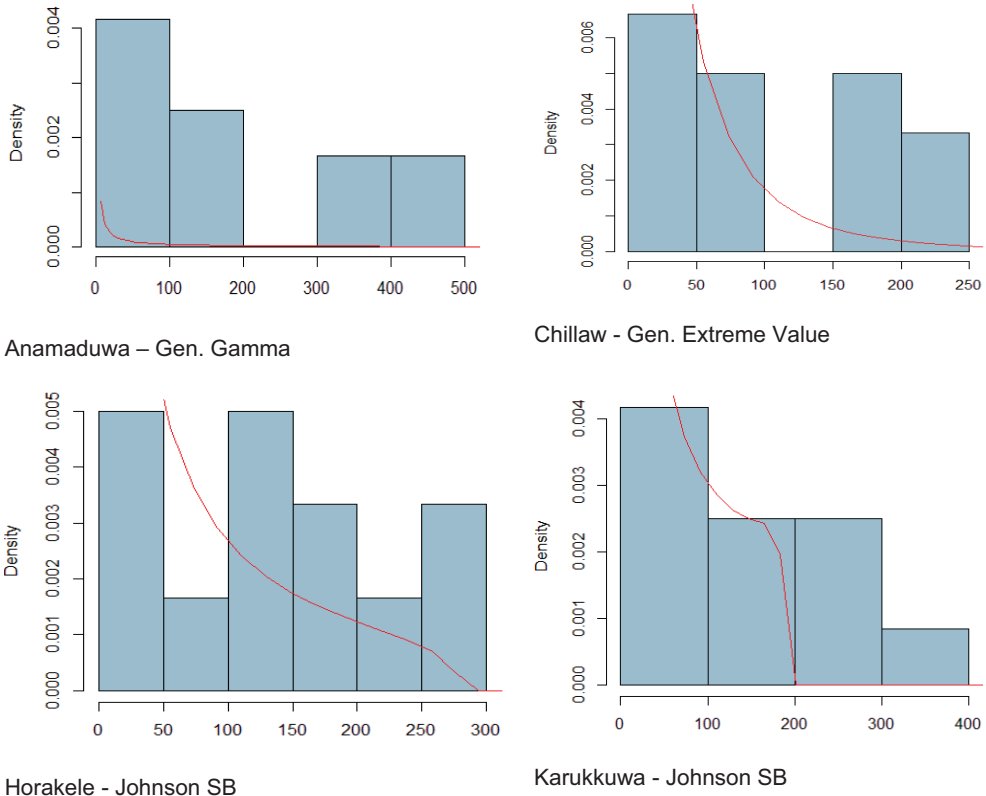


Figure 2. Histogram of the monthly rainfall data at 04 selected rain gauge stations in North-western Province along with best fit probability distribution for each locations considered.

4. DISCUSSION AND CONCLUSION

According to Table 01, the lowest average annual rainfall (1101.8 mm) was recorded at Palaviya (CV = 23.23%) in the North-western coastal line and the highest average rainfall (2104.3 mm) was recorded in Palampoddaru (CV=29.9%) in the Eastern region. In the North-western province the lowest CV (17.69%) was obtained at Lunuvila. Mylambaveli recorded the lowest CV (24.12%) for the Eastern province. However, most stations had recorded high CV (>30%) values in the Eastern coastline buttressing that the rainfall was highly dispersed or there was an inconsistency in the rainfall pattern. According to Hare (2003), CV is used to classify the degree of variability of rainfall events as less (CV <20), moderate (20 < CV < 30), and high (CV > 30). The highest CV was obtained at Navakiri Aru (40.15%) of the Eastern coastal line indicating the existence of high variability in rainfall in the considered location. As table 01 shows, the coefficient of skewness varied from -0.516 to 2.198; kurtosis varied between -0.301 and 8.089. As a rule of thumb, if skewness is in between -1

and -0.5 or between 0.5 and 1, the distribution is considered as moderately skewed. In the North-western coastal line the coefficients of skewness varied from -0.516 to 0.621, indicating moderate skewness. Nevertheless, in the Eastern coastal line coefficients of skewness varied from -0.502 to 2.198 showing that the annual rainfall distribution is highly skewed with more chance of resulting extreme annual rainfall occurrences. Out of 19 stations in the study, 05 rain gauge stations in the North-western coastal line; Anamduwa, Karukkuwa, Kottikachchuwa, Palavi, and Puttlam, and all the stations in Eastern coastal line except Palampoddu showed positive coefficients of skewness values (Table 1). The positive value of the skewness coefficient indicates that the data set is right-skewed. These positive values of skewness indicate that high extreme values of rainfall are rarely seen and annual rainfall amounts, in general, less than averages. Therefore, the frequency of low annual rainfall is high in these stations and the frequency of extreme values above the average is low. Kurtosis is a measure of data peakedness or flatness relative to a normal distribution. For kurtosis, the general guideline is that if the value is greater than +1, the distribution is too peaked. Likewise, a kurtosis of less than -1 indicates a distribution that is too flat. (Hair et al., 2017). Accordingly, flat distributions were observed at Anamaduwa, Palaviya, and Puttlam in the North-western region and Palampoddu and Pottuvil in the Eastern region (Table 1).

According to the K-S test (D), Anderson-Darling test (AD), and Chi-square test (X²) statistics, there is no one distribution that consistently provides the best fit for most of the locations. In the K-S test (D), 04 stations in the North-western province and 02 stations in the Eastern province were best fitted by the Johnson SB distribution. According to D'Adderio et al., (2016) the Johnson SB distribution is capable of estimating the integral parameters of rainfall, especially rainfall rate and reflectivity factor. The same author has mentioned that the Johnson SB distribution reproduces the natural raindrop size distribution quite accurately. In our analysis, Johnson SB distribution was selected as the best-fitted distribution for four stations and these stations have obtained the same result from the K-S test also. It was found that among all distributions, the Johnson SB yielded the most cases of best-fit distributions in North-western province stations. When considering about Eastern province, Gen. Pareto was the best-fitted distribution for three stations. Also, Gen. Extreme Value distribution, Fréchet (3P) were well fitted to 02 and 03 stations respectively within the considered regions. Therefore, although no one distribution provides the best fit for all stations, most of them could be best fitted using Johnson SB distribution. The Generalized Pareto distribution is useful for describing events that exceed a specified lower bound, such as rainfall events above a given threshold. The three-parameters are location, scale, and shape k (Alam et al., 2018). The Fréchet distribution, also known as inverse Weibull distribution (Khan et al., 2008) is a special case of the generalized extreme value distribution. In hydrology, the Fréchet distribution is applied to extreme events such as annually maximum one-day rainfalls and river discharges (Stuart, 2010) as well as it is appropriate for extreme rainfall (Koutsoyiannis, 2004). A well-known three-parameter distribution for maxima is the GEV. In many European countries, such as Austria, Germany, Italy, and Spain, the GEV distribution is recommended to provide the best-fit to flood data (Salinas et al., 2014). The study concludes that Johnson SB is a more common and acceptable distribution for North-western coastal line rainfall as this model was fitted for most stations. But for the rainfall in

Eastern Coastal Line, several distributions were commonly selected. Accordingly, Lognormal, Gamma, and Johnson SB were more common and acceptable distributions for Eastern Coastal Line. Further it shows that climate change in terms of rainfall either in amount and pattern is location-specific in Sri Lanka. Hence, a thorough locational analysis and understanding of climate transition are critical for formulating effective adaptation strategies and implementing proper mitigation strategies. Also, authors would like to suggest to spur future studies to consider these distributions for fitting rainfall data in the remaining provinces of Sri Lanka and carry an island wide study for further analysis. However, if the availability of rainfall data at present is not sufficient to satisfy the above needs, in long run, we need to establish weather stations in close proximities. This findings can be used in crop cultivation timetables. Therefore, in light of the evolving climate change impacts, to achieve growth and sustainability in the agriculture, aquaculture, and tourism industries, rigorous efforts are necessary to mainstream climate change adaptation into national development policies and ensure that they are implemented at national, regional, and local levels.

5. REFERENCES

- Alam, M. A., Farnham, C., & Emura, K, 2018, Best-fit probability models for maximum monthly rainfall in Bangladesh using gaussian mixture distributions. *Geosciences* (Switzerland), 8(4). <https://doi.org/10.3390/geosciences8040138>
- Amin, M. T., Izwan, M., & Alazba, A. A., 2016, A best-fit probability distribution for the estimation of rainfall in northern regions of Pakistan. *Open Life Sciences*, **11**(1), 432–440. <https://doi.org/10.1515/biol-2016-0057>
- D'Adderio, L. P., Cugerone, K., Porcù, F., De Michele, C., & Tokay, A., 2016), Capabilities of the Johnson SB distribution in estimating rain variables. *Advances in Water Resources*, **97**, 241–250. <https://doi.org/10.1016/j.advwatres.2016.09.017>
- Hair, J. F., Hult, G. T. M., Ringle, C. M., and Sarstedt, M., 2017. *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)*. 2nd Ed. Thousand Oaks, CA: Sage.
- Jamaludin, S., & Jemain, A., 2007, Fitting daily rainfall amount in Peninsular Malaysia using several types of exponential distributions. *Journal of Applied Sciences Research*.
- Khan M.S., Pasha G.R., Pasha A.H., 2008, Theoretical Analysis of Inverse Weibull Distribution. *WSEAS TRANSACTIONS on MATHEMATICS*. **7** (2), 30–38
- Khudri, M., & Sadia, F., 2013, Determination of the Best Fit Probability Distribution for Annual Extreme Precipitation in Bangladesh. *European Journal of Scientific Research*, **103**(JUNE), 391–404. <https://doi.org/10.5923/j.ajms.20160604.05>
- Koutsoyiannis, K., 2004, Statistics of extremes and estimation of extreme rainfall: II. Empirical investigation of long rainfall records / Statistiques de valeurs extrêmes et estimation de précipitations extrêmes: II. Recherche empirique sur de longues séries de précipitations. *Hydrological Sciences Journal*. **49**, 4, -610, DOI: 10.1623/hysj.49.4.591.54424
- Mayooran, T., & Laheetharan, A., 2014, The Statistical Distribution of Annual Maximum Rainfall in Colombo District. *Sri Lankan Journal of Applied Statistics*, **15**(2), 107. <https://doi.org/10.4038/sljastats.v15i2.7411>
- P. Esberto, M. D., 2018, Probability Distribution Fitting of Rainfall Patterns in Philippine Regions for Effective Risk Management. *Environment and Ecology Research*, **6**(3), 178–186. <https://doi.org/10.13189/eer.2018.060305>

Salinas, J. L., Castellarin, A., Kohnová, S., & Kjeldsen, T. R., 2013, On the quest for a pan-European flood frequency distribution: effect of scale and climate. *Hydrology and Earth System Sciences Discussions*. <https://doi.org/10.5194/hessd-10-6321-2013>

Salinas, J.L., Castellarin, A., Kohnová, S., & Kjeldsen, T.R., 2014, Regional parent flood frequency distributions in Europe—Part 2: Climate and scale controls. *Hydrology and Earth System Sciences Discussions*. **18**, 4391–4401.

Schittkowski, K., 2002, *Data Fitting in Dynamical Systems with EASY-FIT - User's Guide*. 1–326.

Stuart, C., 2001, *An Introduction to Statistical Modeling of Extreme Values*. Springer-Verlag. ISBN 978-1-85233-459-8.

Sukrutha, A., Dyuthi, S. R., & Desai, S., 2018, Multimodel response assessment for monthly rainfall distribution in some selected Indian cities using best-fit probability as a tool. *Applied Water Science*, **8**(5), 1–10. <https://doi.org/10.1007/s13201-018-0789-4>

.

Annexure: Probability distributions and their probability density functions used to fit – best fit probability distribution for the monthly rainfall data

No	Distribution	Probability Density Function	Range	Parameters
1	Beta	$f(x) = \frac{1}{B(\alpha_1, \alpha_2)} \frac{(x-a)^{\alpha_1-1} (b-x)^{\alpha_2-1}}{(b-a)^{\alpha_1+\alpha_2-1}}$ Where B is the beta function	$a \leq x \leq b$	α_1 continuous shape parameter ($\alpha_1 > 0$) α_2 continuous shape parameter ($\alpha_2 > 0$) a, b continuous boundary parameters ($a < b$)
2	Burr (4P)	$f(x) = \frac{\alpha k \left(\frac{x-\gamma}{\beta}\right)^{\alpha-1}}{\beta \left(1 + \left(\frac{x-\gamma}{\beta}\right)^\alpha\right)^{k+1}}$	$\gamma \leq x < +\infty$	k continuous shape parameter α continuous shape parameter β continuous scale parameter γ continuous location parameter ($\gamma = 0$ yields the three-parameter Burr distribution)
3	Burr (3P)	$f(x) = \frac{\alpha k \left(\frac{x-\gamma}{\beta}\right)^{\alpha-1}}{\beta \left(1 + \left(\frac{x}{\beta}\right)^\alpha\right)^{k+1}}$		
4	Cauchy	$f(x) = \frac{1}{\pi \sigma \left(1 + \left(\frac{x-\mu}{\sigma}\right)^2\right)}$	$-\infty < x < +\infty$	σ continuous scale parameter ($\sigma > 0$) μ continuous location parameter
5	Chi-Square (2P)	$f(x) = \frac{(x-\gamma)^{\frac{\nu}{2}-1} \exp\left(-\frac{(x-\gamma)}{2}\right)}{2^{\nu/2} \Gamma(\frac{\nu}{2})}$	$\gamma \leq x < +\infty$	ν degrees of freedom (positive integer) γ continuous location parameter ($\gamma=0$ yields the one-parameter Chi squared distribution)
6	Chi-Square (1P)	$f(x) = \frac{(x-\gamma)^{\frac{\nu}{2}-1} \exp\left(-\frac{(x-\gamma)}{2}\right)}{2^{\nu/2} \Gamma(\frac{\nu}{2})}$		

No	Distribution	Probability Density Function	Range	Parameters
7	Dagum (4P)	$f(x) = \frac{\alpha k \left(\frac{x-\gamma}{\beta}\right)^{\alpha k-1}}{\beta \left(1 + \left(\frac{x-\gamma}{\beta}\right)^{\alpha k+1}\right)}$	$\gamma \leq x < +\infty$	k continuous shape parameter α continuous shape parameter β continuous scale parameter γ continuous location parameter (γ =0 yields the three-parameter Dagum distribution)
8	Erlang (3P)	$f(x) = \frac{(x-\gamma)^{m-1}}{\beta^m \Gamma(m)} \exp\left(-\left(\frac{x-\gamma}{\beta}\right)\right)$	$\gamma \leq x < +\infty$	m shape parameter (positive integer) β continuous scale parameter γ continuous location parameter (γ =0 yields the two-parameter Erlang distribution)
9	Erlang (2P)	$f(x) = \frac{(x-\gamma)^{m-1}}{\beta^m \Gamma(m)} \exp\left(-\left(\frac{x}{\beta}\right)\right)$		
10	Error function	$f(x) = \frac{h}{\sqrt{\pi}} \exp(- (hx)^2)$	$-\infty < x < +\infty$	k continuous shape parameter σ continuous scale parameter (σ<0) μ continuous location parameter
12	Exponential (2P)	$f(x) = \lambda \exp(-\lambda(x-\gamma))$	$\gamma \leq x < +\infty$	λ continuous shape parameter γ continuous location parameter (γ =0 yields the one-parameter Exponential distribution)
13	Exponential (1P)	$f(x) = \lambda \exp(-\lambda x)$		
14	Frechet (2P) (Maximum extreme Value type 2)	$f(x) = \frac{\alpha}{\beta} \left(\frac{\beta}{x}\right)^{\alpha+1} \exp\left(1\left(\frac{\beta}{x}\right)^\alpha\right)$	$0 \leq x < +\infty$	α continuous shape parameter (α>0) β continuous scale parameter (β>0)
15	Gamma (3P)	$f(x) = \frac{x - \gamma^{\alpha-1}}{\beta^\alpha \Gamma(\alpha)} \exp\left(-1\left(\frac{x-\gamma}{\beta}\right)\right)$	$\gamma \leq x < +\infty$	α continuous shape parameter (α>0) β continuous scale parameter (β>0) γ continuous location parameter (γ =0 yields the two-parameter Gamma distribution)
16	Gamma(2P)	$f(x) = \frac{(x)^{\alpha-1}}{\beta^\alpha \Gamma(\alpha)} \exp\left(-1\left(\frac{x}{\beta}\right)\right)$		
17	Generalized Gamma (4P)	$f(x) = \frac{k(x-\gamma)^{k\alpha-1}}{\beta^{k\alpha} \Gamma(\alpha)} \exp\left(-1\left(\left(\frac{x-\gamma}{\beta}\right)^k\right)\right)$	$\gamma \leq x < +\infty$	k continuous shape parameter α continuous shape parameter β continuous scale parameter γ continuous location parameter (γ =0 yields the three-parameter Generalized Gamma distribution)
18	Generalized Gamma (3P)	$f(x) = \frac{k(x)^{k\alpha-1}}{\beta^{k\alpha} \Gamma(\alpha)} \exp\left(-1\left(\left(\frac{x}{\beta}\right)^k\right)\right)$		

No	Distribution	Probability Density Function	Range	Parameters
19	Generalized Extreme Value	$f(x) = \frac{1}{\alpha} \exp(-(1 - kz)^{-1/k})((1 + kz)^{-(k+1)/k})$ If $k \neq 0$	Where $z = \frac{(x-\pi)}{\sigma} \cdot 1 + k\left(\frac{x-\pi}{\sigma}\right)$ > 0 For $k \neq 0$, $-\infty < x < +\infty$	k continuous shape parameter σ continuous scale parameter ($\sigma > 0$) μ continuous location parameter
20	Gumble Max (Maximum extreme value Type 1)	$f_{(x)} = 1/\beta \exp(-z + \exp(-z)), z = \frac{(x - \pi)}{\sigma}$	$-\infty < x < +\infty$	σ continuous scale parameter ($\sigma > 0$) μ continuous location parameter
21	Inverse Gaussian (3P)	$f_{(x)} = \sqrt{\frac{\lambda}{2\pi(x - \gamma)}} \exp\left(-\frac{\lambda(x - \gamma - \mu)^2}{2\mu^2(x - \gamma)}\right)$	$\gamma \leq x < +\infty$	λ continuous shape parameter ($\lambda > 0$) μ continuous parameter ($\mu > 0$) γ continuous location parameter ($\gamma = 0$ yields the two-parameter Inverse Gaussian distribution)
22.	Inverse Gaussian (2P)	$f_{(x)} = \sqrt{\frac{\lambda}{2\pi x^3}} \exp\left(-\frac{\lambda(x - \gamma - \mu)^2}{2\mu^2(x - \gamma)}\right)$		
23.	Kumaraswamy	$f_{(x)} = \frac{\alpha_1 \alpha_2 Z^{(\alpha_1)} (1 - Z^{\alpha_1})^{(\alpha_2 - 1)}}{b - a}$	$a \leq x \leq b$	α_1 continuous shape parameter ($\alpha_1 > 0$) α_2 continuous shape parameter ($\alpha_2 > 0$) a, b continuous boundary parameters ($a < b$)
24.	Johnson SB	$f_{(x)} = \frac{\delta}{\lambda \sqrt{2\pi z(1 - z)}} \exp\left(-\frac{1}{2}\left(\gamma + \delta \ln\left(\frac{z}{1 - z}\right)\right)^2\right)$	$\xi \leq x \leq \xi + \lambda$	γ continuous shape parameter δ continuous shape parameter ($\delta > 0$) λ continuous scale parameter ($\lambda = 0$) ξ continuous location parameter
25.	Laplace (Double Exponential)	$f_{(x)} = \frac{\lambda}{2} \exp(-\lambda x - \lambda)$	$-\infty < x < +\infty$	λ continuous inverse shape parameter ($\lambda > 0$) μ continuous parameter

No	Distribution	Probability Density Function	Range	Parameters
26.	Levy	$f_{(x)} = \sqrt{\frac{\sigma}{2\pi}} \frac{\exp\left(-0.5\frac{x}{(x)^{3/2}}\right)}{(x)^{3/2}}$	$0 < x < +\infty$	σ continuous scale parameter ($\sigma > 0$)
27.	Log-Gamma	$f_{(x)} = \frac{(\ln(x))^{\alpha-1}}{x\beta\Gamma(\alpha)} \exp\left(\ln\left(\frac{x}{\beta}\right)\right)$	$0 < x < +\infty$	α continuous parameter ($\alpha > 0$) β continuous parameter ($\beta > 0$)
28.	Logistic	$f_{(x)} = \frac{\exp\left(-\left(\frac{x-\mu}{\sigma}\right)\right)}{\sigma\left(1 + \exp\left(-\left(\frac{x-\mu}{\sigma}\right)\right)\right)^2}$	$-\infty < x < +\infty$	σ continuous scale parameter ($\sigma > 0$) μ continuous location parameter
29.	Log-Logistic (3P)	$f_{(x)} = \frac{\alpha}{\beta} \left(\frac{x-\gamma}{\beta}\right)^{\alpha-1} \left(1 + \left(\frac{x-\gamma}{\beta}\right)^{\alpha}\right)^{-2}$	$\gamma \leq x < +\infty$	α continuous shape parameter β continuous scale parameter γ continuous location parameter ($\gamma = 0$ yields the two-parameter Log-Logistic distribution)
30.	Log-Logistic (2P)	$f_{(x)} = \frac{\alpha}{\beta} \left(\frac{x}{\beta}\right)^{\alpha-1} \left(1 + \left(\frac{x}{\beta}\right)^{\alpha}\right)^{-2}$		
31.	Lognormal (3P)	$f_{(x)} = \frac{\exp\left(-\frac{1}{2}\left(\frac{\ln(x-\gamma)-\mu}{\sigma}\right)^2\right)}{(x-\gamma)\sigma\sqrt{2\pi}}$	$\gamma \leq x < +\infty$	α continuous parameter μ continuous parameter γ continuous location parameter ($\gamma = 0$ yields the two-parameter Lognormal distribution)
32.	Lognormal (2P)	$f_{(x)} = \frac{\exp\left(-\frac{1}{2}\left(\frac{\ln(x)-\mu}{\sigma}\right)^2\right)}{(x)\sigma\sqrt{2\pi}}$		
33.	Log-Pearson 3 (3P)	$f_{(x)} = \frac{1}{x \beta \Gamma(\alpha)} \left(\frac{\ln(x)-\gamma}{\beta}\right)^{\alpha-1} \exp\left(\frac{\ln(x)-\gamma}{\beta}\right)$	$0 < x \leq e^{\gamma}$ for $\beta < 0$ and $e^{\gamma} \leq x < +\infty$ For $\beta > 0$	α continuous parameter ($\alpha > 0$) β continuous parameter ($\beta > 0$) γ continuous parameter

No	Distribution	Probability Density Function	Range	Parameters
34.	Nakagami	$f_{(x)} = \frac{2m^m}{\Gamma(m)\Omega^m} x^{2m-1} \exp\left(-\frac{m}{\Omega} x^2\right)$	$0 < x < +\infty$	m continuous parameter $m \geq 0.5$ Ω continuous parameter ($\Omega > 0$)
35.	Normal	$f_{(x)} = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right)$	$-\infty < x < +\infty$	σ continuous scale parameter ($\sigma > 0$) μ continuous location parameter
36.	Pareto (1 st Kind)	$f_{(x)} = \frac{\alpha\beta^\alpha}{(x)^\alpha}$	$\beta < x < +\infty$	α continuous parameter ($\alpha > 0$) β continuous parameter ($\beta > 0$)
37.	Pareto (2 nd Kind)	$f_{(x)} = \frac{\alpha\beta^\alpha}{(x+\beta)^{\alpha+1}}$	$0 < x < +\infty$	α continuous parameter ($\alpha > 0$) β continuous parameter ($\beta > 0$)
No	Distribution	Probability Density Function	Range	Parameters
38.	Pearson type 5 (3P)	$f_{(x)} = \frac{\exp(-\beta/(x-\gamma))}{\beta\Gamma(\alpha)((x-\gamma)/\beta)^{\alpha+1}}$	$\gamma < x < +\infty$	α continuous shape parameter ($\alpha > 0$) β continuous scale parameter ($\beta > 0$) γ continuous location parameter ($\gamma = 0$ yields the two-parameter Pearson 5 distribution)
39.	Pearson type 5 (2P)	$f_{(x)} = \frac{\exp(-\beta/x)}{\beta\Gamma(\alpha)(x/\beta)^{\alpha+1}}$		
40.	Pearson type 6 (4P)	$f_{(x)} = \frac{((x-\gamma)/\beta)^{\alpha_1-1}}{\beta B(\alpha_1, \alpha_2)(1+(x-\gamma)/\beta)^{\alpha_1+\alpha_2}}$	$\gamma \leq x < +\infty$	α_1 continuous shape parameter ($\alpha_1 > 0$) α_2 continuous shape parameter ($\alpha_2 > 0$) β continuous scale parameter ($\beta > 0$) γ continuous location parameter ($\gamma = 0$ yields the three-parameter Pearson 6 distribution)
41.	Pearson type 6 (3P)	$f_{(x)} = \frac{((x/\beta)^{\alpha_1-1}}{\beta B(\alpha_1, \alpha_2)(1+x-\beta)^{\alpha_1+\alpha_2}}$		
42.	Student's t	$f_{(x)} = \frac{1}{\sqrt{\pi}\nu} \left(\frac{\nu+1}{2}\right) \left(\frac{\nu}{\nu+x^2}\right)^{\frac{\nu+1}{2}}$	$-\infty < x < +\infty$	ν degrees of freedom (positive integer)
43.	Weibull (3P)	$f_{(x)} = \frac{\alpha}{\beta} \left(\frac{x-\gamma}{\beta}\right)^{\alpha-1} \exp\left(-\left(\frac{x-\gamma}{\beta}\right)^\alpha\right)$	$\gamma \leq x < +\infty$	α continuous shape parameter ($\alpha > 0$) β continuous scale parameter ($\beta > 0$) γ continuous location parameter ($\gamma = 0$ yields the two-parameter Weibull distribution)
44.	Weibull (2P)	$f_{(x)} = \frac{\alpha}{\beta} \left(\frac{x}{\beta}\right)^{\alpha-1} \exp\left(-\left(\frac{x}{\beta}\right)^\alpha\right)$		