Applied Mathematics and Computational Intelligence
Volume 10, No.1, Dec 2021 [288-300]

Application of the Length-Biased Weibull-Rayleigh Distribution to Fit the Rainy Season Rainfall for the Upper Ping River in Northern Thailand

Tanachot Chaito[1], Nawapon Nakharutai[2], Manad Khamkong[2*]

[1]Ph.D. Program in Applied Statistics, Department of Statistics, Faculty of Science, Chiang Mai University, Chiang Mai, 50200 Thailand
[2]Data Science Research Center, Department of Statistics, Faculty of Science, Chiang Mai University, Chiang Mai, 50200 Thailand

[*]Corresponding author : manad.k@cmu.ac.th

Received: 25 October 2021; Accepted: 8 November 2021; Available online: 23 November 2021

ABSTRACT

Statistical distributions are important and useful to determine the appropriate distributions for rainfall data and predict the return levels of rainfall data. The objectives of this study are to apply the length-biased Weibull-Rayleigh (LBWR) distribution for fitting the rainy season rainfall data and to predict the return levels of the rainy season rainfall data. The LBWR distribution is compared with the Rayleigh, Weibull, and Weibull-Rayleigh distributions to determine the best - fit for the rainy season rainfall data from 1957 to 2020 of Samoeng (Chiang Mai) and Mae Tha (Lamphun) stations in Thailand. From the Kolmogorov-Smirnov test, Anderson-Darling test and Akaike information criteria, the results show that the LBWR distribution is the best-fit distribution of the rainy season rainfall data for both stations. Based on the prediction of the return levels of the rainy season rainfall data, it shows that the rainy season rainfall at Samoeng station had higher return level than the rainy season rainfall at Mae Tha station. This could conclude that Samoeng district having a higher risk of flood compared to Mae Tha district. Therefore, the results from this study could be useful to formulate guidelines and strategies for flood irrigation and water management in Samoeng district.

Keywords: Length-biased distribution, Rainfall data, Ping River, Return level, Profile likelihood.

1 INTRODUCTION

Nowadays, the world is facing with climate change, including flood and drought. The flood could give damage of life, habitat and economy, and they are expected to become more severe in the future [1]. Intergovernmental Panel on Climate Change [2] reported that in the 21st century, heavy rainfall will occur more frequently in many areas of the world cause the increasing risk of flooding that contributes damaging infrastructure and economy. In many areas of Thailand, there were several extreme floods occurred. For example, the northern and central regions that had heavily flooded in 2011. These floods resulted in damage of agricultural, industry and economy sectors [3]. Upper Ping

River basin consists of Chiang Mai and Lamphun provinces. Both areas are still experiencing continuous flooding, which severely affects the agricultural and industry in the areas. The major factor causing of flood is extreme rainfall. Therefore, an approach that will prevent or reduce the severity of flood is to monitor the areas using flood irrigation [4] or to introduce suitable models for determining the return levels of the highest rainfall during the return period [5].

Rainfall data is often right–skewed and in some situations, outlier or extreme values can occur because of heavy rain events [6]. Thus, various researchers have tried to fit several models for rainfall data [7, 8, 9]. Common statistical distributions that have been applied to model rainfall data are Gumbel, Weibull, gamma, lognormal, Pearson type III, Frechet, and generalized extreme value distributions, among others [8]. Several authors [7, 9] claimed that the Weibull distribution could be a best choice for fitting rainfall data because it is a heavy–tailed distribution. However, the Weibull distribution has been developed in order to fit hydrological data. For example, Ganji et al. [10] developed the Weibull–Rayleigh distribution, which is mixed distribution, and suggested that it could provide a better fitting for flood data compared with the beta–Pareto, generalized exponential, Weibull, three–parameter Weibull, and Pareto distributions. In their work, they did not apply the Weibull–Rayleigh distribution to fit rainfall data. Thus, we would like to investigate whether the Weibull–Rayleigh distribution could be further developed to fit rainfall data.

Rainfall data is categorised as environment data that are usually non–random and non–replicated which can lead to bias recorded observations [11]. A weighted distribution is a common method using when the probabilities of observations recorded from a random process are not equal. The weighted distribution was first proposed by Fisher [12] and further extended by Rao [13] as the length–biased distribution. Various researchers have used the length–biased distribution to improve the fitting of a distribution such as the length–biased weighted generalized Rayleigh distribution [14], the length–biased weighted Weibull distribution [15] and the length–biased beta distribution [16] were generalized from the length–biased distribution. These improved distributions were widely applied to fit data in a variety of fields, such as reliability, lifetime, and engineering [14, 16]. However, as far as we know, the length–biased distribution is rarely used to hydrological data. There are a few studies on applying the length–biased distribution to hydrological data e.g., the length–biased weighted Weibull and the length–biased Weibull–Rayleigh (LBWR) distributions [17]. In our previous study [17], we modified the Weibull–Rayleigh distribution using the length–biased distribution and suggested that the LBWR distribution could provide more efficiency of fitting to flood datasets than the Rayleigh, Weibull, Pareto, and Weibull–Rayleigh distributions. Therefore, it might be potential to apply the LBWR distribution to fit rainfall data.

Among four distributions, Rayleigh, Weibull, Weibull–Rayleigh and LBWR distributions, we are going to find the best–fit distribution for the rainy season rainfall data, i.e. collected during June to September in every year, from the Hydrology and Water Management Center for the Upper Northern Region of Thailand. The data were collected from 1957 to 2020 at Samoeng and Mae Tha stations. Moreover, we will predict the return levels of the rainy season rainfall data to identify the risk of flooding in particular areas. The article is organized as follow. Section 2 describes the study area and the methodology that we use in studying. Section 3 shows the results and discussion of an appropriate distribution to the rainy season rainfall data and the return levels of the rainfall data. Conclusion of this study is presented in Section 4.

2 MATERIAL AND METHODS

2.1 Study Area

Samoeng District, Chiang Mai Province, and Mae Tha District, Lamphun Province, (Figure 1) were chosen at study cases, since floods were frequently occurred in both areas. Monthly rainfall data of Samoeng and Mae Tha stations from January 1957 to December 2020 were obtained from the Hydrology and Water Management Center for the Upper Northern Region of Thailand [18]. This study handled the missing monthly rainfall data by replacing it with the average of the past five years of such month. Then, the monthly rainfall data were classified into seasonal rainfall data using the Thai seasonal criterion determined by Meteorological Department of Thailand [19]: the summer season (February to May), the rainy season (June to September) and the winter season (October to January of the following year). Therefore, this study uses only the rainy season rainfall data to determine appropriate distributions for rainfall data and predict the return levels of rainfall data.

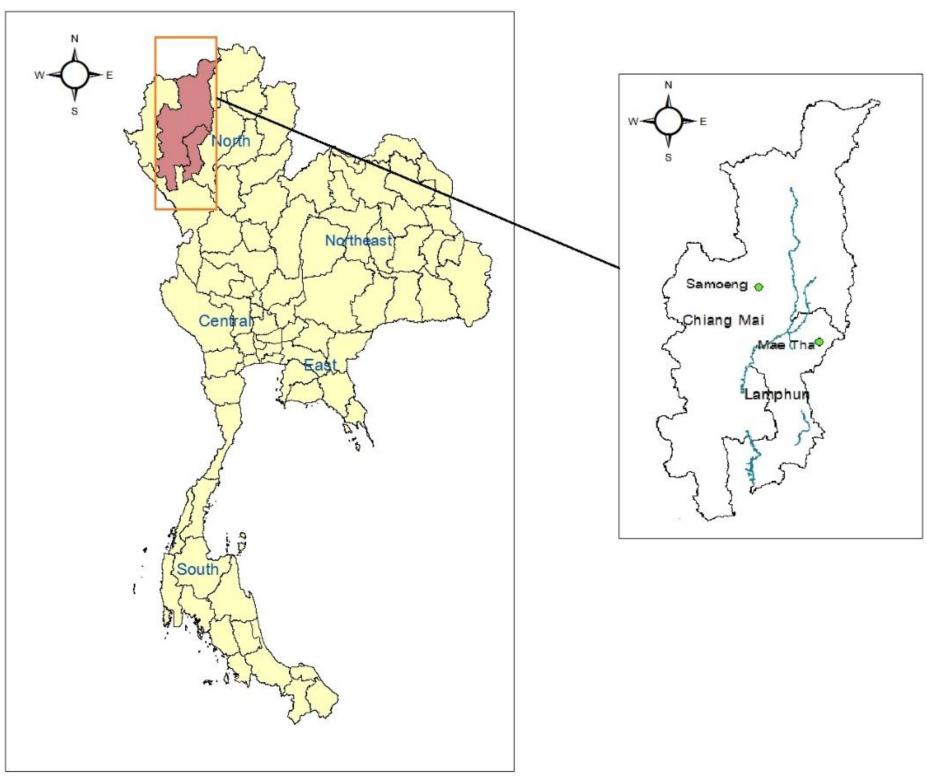

Figure 1: Location of two stations on upper Ping River in northern Thailand.

2.2 The Length–Biased Weibull–Rayleigh (LBWR) Distribution

The LBWR distribution was presented by Chaito and Khamkong [17]. This distribution improved Weibull–Rayleigh distribution by using the length–biased distribution, which is special case of weighted distribution. The probability density function (pdf) and cumulative distribution function (cdf) of the LBWR distribution are given by

$$f_L(x) = \frac{\alpha x^2}{\beta\delta^2\sqrt{2\beta\delta^2}\,\Gamma\left(1+\frac{1}{2\alpha}\right)}\left(\frac{x^2}{2\beta\delta^2}\right)^{\alpha-1}\exp\left[-\left(\frac{x^2}{2\beta\delta^2}\right)^{\alpha}\right], \quad x>0, \quad \alpha,\beta,\delta>0, \tag{1}$$

$$F_L(x) = \frac{\gamma\left(1+\frac{1}{2\alpha},\left(\frac{x^2}{2\beta\delta^2}\right)^{\alpha}\right)}{\Gamma\left(1+\frac{1}{2\alpha}\right)}, \tag{2}$$

where α is a shape parameter, β and δ are scale parameters, $\Gamma(\alpha)=\int_0^\infty u^{\alpha-1}e^{-u}du$ is a gamma function and $\gamma(\alpha,x)=\int_0^x u^{\alpha-1}e^{-u}du$ is the lower incomplete gamma function.

2.3 Maximum Likelihood Estimation

This study was conducted by using maximum likelihood estimation for estimating parameters of the LBWR distribution. Maximum likelihood estimation considers random sample. Let $X_1,X_2,\dots,X_n$ be a random sample from the LBWR distribution with parameter vector $\mathbf{\Theta}=(\alpha,\beta,\delta)$, $x_1,x_2,\dots,x_n$ be the sample values. The likelihood and log–likelihood functions are given by

$$L(\mathbf{\Theta}) = \prod_{i=1}^{n}\left\{\frac{\alpha x_i^2}{\beta\delta^2\sqrt{2\beta\delta^2}\,\Gamma\left(1+\frac{1}{2\alpha}\right)}\left(\frac{x_i^2}{2\beta\delta^2}\right)^{\alpha-1}\exp\left[-\left(\frac{x_i^2}{2\beta\delta^2}\right)^{\alpha}\right]\right\}, \tag{3}$$

$$\begin{aligned}\log L(\mathbf{\Theta}) = \; & n\log\alpha + \sum_{i=1}^{n}\log x_i^2 - n\log\beta - 2n\log\delta - \frac{1}{2}n\log(2\beta\delta^2) - n\log\Gamma\left(1+\frac{1}{2\alpha}\right) \\ & + (\alpha-1)\sum_{i=1}^{n}\log\left(\frac{x_i^2}{2\beta\delta^2}\right) - \sum_{i=1}^{n}\log\left(\frac{x_i^2}{2\beta\delta^2}\right)^{\alpha}.\end{aligned} \tag{4}$$

By differentiating (4) with respect to α, β, and δ, we obtain

$$\frac{\partial\log L(\mathbf{\Theta})}{\partial\alpha} = \frac{n}{\alpha} - n\psi\Gamma\left(1+\frac{1}{2\alpha}\right) + \sum_{i=1}^{n}\log\left(\frac{x_i^2}{2\beta\delta^2}\right) - \sum_{i=1}^{n}\left(\frac{x_i^2}{2\beta\delta^2}\right)^{\alpha}\log\left(\frac{x_i^2}{2\beta\delta^2}\right), \tag{5}$$

$$\frac{\partial\log L(\mathbf{\Theta})}{\partial\beta} = -\frac{n\alpha}{\beta} - \frac{n}{2\beta} + \frac{2\alpha\delta^2}{(2\beta\delta^2)^{\alpha+1}}\sum_{i=1}^{n}(x_i)^{2\alpha}, \tag{6}$$

$$\frac{\partial \log L(\mathbf{\Theta})}{\partial \delta} = -\frac{n(2\alpha-1)}{\delta} + \frac{4\alpha\beta\delta}{(2\beta\delta^2)^{\alpha+1}}\sum_{i=1}^{n}(x_i)^{2\alpha}, \tag{7}$$

where $\psi(y) = \frac{d}{dy}\log\Gamma(y) = \frac{\Gamma^{/}(y)}{\Gamma(y)}$ is a digamma function. The maximum likelihood estimators of parameters α, β, and δ can be estimated by setting (5), (6) and (7) equal to zero and solving them simultaneously. To obtain the maximum likelihood estimation, we used ***mle*** function in ***stats4*** package in the ***R*** statistical software [20].

2.4 Model Selection Criteria

Determine the appropriate distribution for the rainy season rainfall data is difficult because the data depend on spatial and temporal factors. For the upper Ping River, the rainy season rainfall data was found to be right-skewed with outliers [6]. We consider four different distributions, namely, Rayleigh, Weibull, Weibull-Rayleigh and LBWR distributions. To select the appropriate distribution for the rainy season rainfall data, we use Kolmogorov-Smirnov (KS) test, Anderson-Darling (AD) test [21] and Akaike information criterion (AIC) [22]. The best-fit distribution for the rainy season rainfall data can be selected from minimum values of KS test, AD test and AIC. The KS test is defined as

$$\mathrm{KS} = \operatorname*{Sup}_{x}[G_0(x) - G(x)], \tag{8}$$

where $G_0(x)$ is empirical distribution function of the observed data and $G(x)$ is the cdf of the hypothesized distribution. The AD test is given by

$$\mathrm{AD} = -n - \frac{1}{n}\sum_{i=1}^{n}(2i-1)\left[\ln G(x_i) + \ln\left(1 - G(x_{n-i+1})\right)\right], \tag{9}$$

where $G(x)$ is the cdf of the hypothesized distribution, n is the sample size and x_i are the ordered data. The AIC is written as

$$\mathrm{AIC} = 2k - 2\log L(\widehat{\mathbf{\Theta}}), \tag{10}$$

where k is the number of parameters and $L(\widehat{\mathbf{\Theta}})$ is the maximized value of the likelihood function.

2.5 Return Level

Once we find the appropriate distribution for the rainy season rainfall data, we also use it to predict the return levels of the data. If the appropriate distributions of the rainy season rainfall data are known, then the return levels can be calculated via these cumulative distribution functions [23]. The return period (T) is defined as the value which is exceeded average once every interval of time T with a probability $\frac{1}{T}$ [23]. The $T - year$ return level of the LBWR distribution can be calculated as follows:

$$x_T = \sqrt{2\hat{\beta}\hat{\delta}^2 A^{\frac{1}{\hat{\alpha}}}}, \tag{11}$$

where $A = \gamma^{-1}\left[1 + \frac{1}{2\hat{\alpha}}, \Gamma\left(1 + \frac{1}{2\hat{\alpha}}\right)\left(\frac{1}{T}\right)\right]$, when γ^{-1} is inverted of the lower incomplete gamma function and Γ is the gamma function, T is return period, $\hat{\alpha}$ is a shape parameter, and $\hat{\beta}$ and $\hat{\delta}$ are scale parameters, which were estimated via maximum likelihood estimation method.

Profile likelihood method is used for calculating the confident interval of return levels in this study. The profile likelihood method assumes that let $X_1, X_2, \ldots, X_n$ are the independent random variables, $\widehat{\boldsymbol{\theta}}_0$ is the maximum likelihood estimator of the d-dimensional parameters of model $\boldsymbol{\theta}_0 = \left(\boldsymbol{\theta}^{(1)}, \boldsymbol{\theta}^{(2)}\right)$, where $\boldsymbol{\theta}^{(1)}$ corresponds the interesting component in $\boldsymbol{\theta}_0$ and $\boldsymbol{\theta}^{(2)}$ corresponds the remaining component in $\boldsymbol{\theta}_0$ [24]. Then, under suitable regularity conditions and for large n, the deviance function is

$$D_p\left(\boldsymbol{\theta}^{(1)}\right) = 2\left\{\ell\left(\widehat{\boldsymbol{\theta}}_0\right) - \ell_P\left(\boldsymbol{\theta}^{(1)}\right)\right\} \sim \chi_k^2, \tag{12}$$

where $\ell\left(\widehat{\boldsymbol{\theta}}_0\right)$ is the maximized value of log–likelihood for the model and $\ell_P\left(\boldsymbol{\theta}^{(1)}\right)$ is the maximized value of log–likelihood for $\boldsymbol{\theta}^{(1)}$, $\ell_P\left(\boldsymbol{\theta}^{(1)}\right) = \max_{\boldsymbol{\theta}^{(2)}} \ell\left(\boldsymbol{\theta}^{(1)}, \boldsymbol{\theta}^{(2)}\right)$ is the profile log–likelihood for $\boldsymbol{\theta}^{(1)}$. For the profile likelihood confidence interval of return levels of the LBWR distribution, we partition the vector $\boldsymbol{\theta} = (x_T, \alpha, \delta)$ into two components $\left(\boldsymbol{\theta}^{(1)}, \boldsymbol{\theta}^{(2)}\right)$, where $\boldsymbol{\theta}^{(1)}$ represents x_T and $\boldsymbol{\theta}^{(2)}$ represents α and δ. Then, the profile log–likelihood is given by

$$\ell_P\left(\boldsymbol{\theta}^{(1)}\right) = \max_{\alpha,\delta} \ell\,(x_T, \alpha, \delta). \tag{13}$$

Therefore, the profile likelihood of $1 - \omega$ confidence interval of x_T can be written as

$$C_\omega = \left\{x_T : 2\left[\ell(\alpha, \beta, \delta) - \max_{\alpha,\delta} \ell\,(x_T, \alpha, \delta)\right] \leq C_{1-\omega}\right\}, \tag{14}$$

where $C_{1-\omega}$ is 1- ω quantile of the chi–square distribution with one degree of freedom and ω is the significance level.

3 RESULTS AND DISCUSSION

3.1 Model Selection Criteria of Rainy Season Rainfall Data

The results (in Table 1) on an analysis of the rainy season rainfall data from 1957 to 2020 of two stations found that the Samoeng gauge station has an average of the rainy season rainfall at 773.60 mm. and a minimum and maximum of the rainy season rainfall data are 272.40 and 1920.40 mm., respectively. For the Mae Tha gauge station, an average of the rainy season rainfall data is 698.20 mm. and a minimum and maximum are 381.50 and 1933.60 mm., respectively. The rainy season rainfall data of both stations are right–skewed. Figure 2 presents the outlier of the rainy season rainfall data.

Samoeng station has outlier in 1963 (1562.50 mm.) and 1994 (1920.40 mm.) while the Mae Tha station has outlier in 1958 (1933.60 mm.).

Table 1 : Descriptive statistics of the rainy season rainfall data for Samoeng and Mae Tha stations.

Station	Rainfall (mm.)								
	Min.	**Max.**	**Q_1**	**Q_2**	**Q_3**	**Mean**	**SD**	**Skewness**	**Kurtosis**
Samoeng	272.40	1920.40	608.10	747.60	896.60	773.60	266.47	1.5241	7.7506
Mae Tha	381.50	1933.60	567.0	674.90	800.00	698.20	221.20	2.6457	16.2500

mm. denotes millimeter; Q_i denotes the i^{th} quartile of data and SD denotes the standard deviation

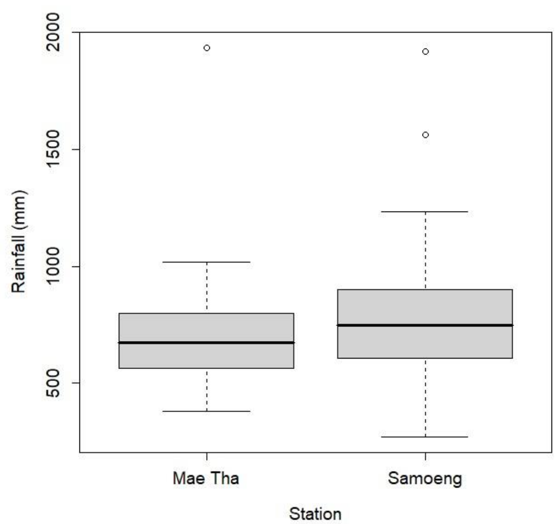

Figure 2: Boxplot of the rainy season rainfall data for Samoeng and Mae Tha stations.

Table 2 and Figure 3 show analysis of the appropriate distributions for the rainy season rainfall data of Samoeng and Mae Tha stations. The result shows that the LBWR distribution is the best-fit distribution for the rainy season rainfall data of both stations due to minimum values of the KS test, AD test and AIC.

Table 2: Summary of selected distributions using the KS test, AD test and AIC for the rainy season rainfall data in Samoeng and Mae Tha stations.

Station	Distribution	KS	AD	AIC
Samoeng	Rayleigh	0.2392	5.3347	913.6498
	Weibull	0.1203	1.7352	900.4815
	Weibull–Rayleigh	0.1206	1.7364	902.4815
	LBWR	**0.1098**	**1.3798**	**898.6661**
Mae Tha	Rayleigh	0.2470	7.3541	896.6363
	Weibull	0.1402	2.8046	880.8026
	Weibull–Rayleigh	0.1403	2.8083	882.8026
	LBWR	**0.1284**	**2.2462**	**876.8741**

Based on the minimum values of the KS test, AD test and AIC in Table 2, the results present that the LBWR distribution is the best–fit distribution for this data. In contrast, due to higher values of the KS test, AD test and AIC, the three other distributions are not suitable for this data. Moreover, in Figure 3, the plots for theoretical densities of four distributions also supported the KS test, AD test and AIC values in Table 2. Specifically, the theoretical density (green line) of the Rayleigh distribution is the least suitable fitting and the theoretical densities of the Weibull distribution (red line) and the Weibull–Rayleigh distribution (blue line) which are similar, are also not suitable in fitting this data. On the other hand, theoretical density of the LBWR distribution (purple line) is the most suitable in fitting this data because the histograms of the rainy season rainfall data from two stations are similar to its plots. Overall, the LBWR distribution outperforms the three distributions in fitting rainy season rainfall data. Furthermore, estimated parameters of the LBWR distribution using maximum likelihood estimation for the rainy season rainfall data are presented in Table 3. All estimated parameters will be used to predict the return levels of the rainy season rainfall data.

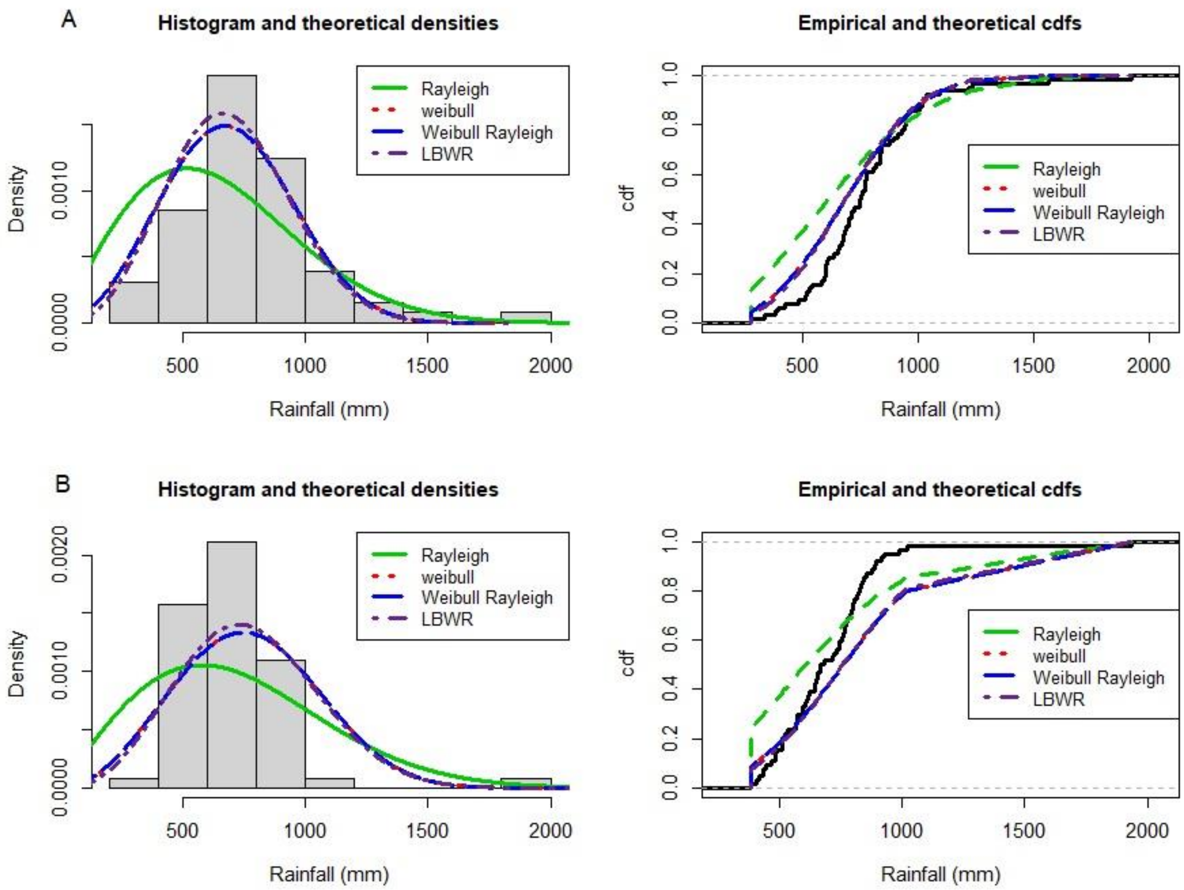

Figure 3: Histograms and theoretical densities (the left column) and empirical and theoretical cdfs (the right column) for the rainy season rainfall data. Figure 1(A) represents results for the data collected from Samoeng station and Figure 1(b) represents results for the data collected from Mae Tha station. The labels indicate different distributions. The horizontal axis presents the rainy season rainfall data (mm.). The vertical axis presents densities in the left column while the vertical axis presents cdfs in the right column.

Table 3: Estimated parameters of the LBWR distribution for the rainy season rainfall data in Samoeng and Mae Tha stations.

Station	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\delta}$
Samoeng	1.2298	151.8907	41.9395
Mae Tha	1.2713	147.1836	38.6006

3.2 Return Levels of Rainy Season Rainfall Data

The predictions for the return levels at 1, 2, 3, 4, 5, 10, 15, 20, 30, 40, 50 and 100 return periods (years) of the rainy season rainfall data in Samoeng and Mae Tha stations are reported in Table 4. The return levels can be predicted via estimated parameters using the maximum likelihood estimation in Table 3. Let T be a return period. The equation to predict the return level (x_T) of Samoeng station is

$$x_T = \sqrt{2 \times 151.8907 \times (41.9395)^2 \times A^{\left(\frac{1}{1.2298}\right)}}, \tag{15}$$

where $A = \gamma^{-1}\left[1 + \frac{1}{(2\times 1.2298)}, \Gamma\left(1 + \frac{1}{(2\times 1.2298)}\right)\left(\frac{1}{T}\right)\right]$. For Mae Tha, the equation to predict the return level (x_T) is

$$x_T = \sqrt{2 \times 147.1836 \times (38.6006)^2 \times A^{\left(\frac{1}{1.2713}\right)}}, \tag{16}$$

where $A = \gamma^{-1}\left[1 + \frac{1}{(2\times 1.2713)}, \Gamma\left(1 + \frac{1}{(2\times 1.2713)}\right)\left(\frac{1}{T}\right)\right]$.

Based on the results in section 3.1, the estimated parameter values of the LBWR distribution in Table 3 are used to calculate the return levels of the rainy season rainfall data collected at Samoeng and Mae Tha stations. The results on an analysis of the return levels for this data in Table 4 present that the return levels of rainy season rainfall collected at Samoeng station are higher than the return levels at Mae Tha station in all considered return periods. Furthermore, the result shows that the return levels of the rainy season rainfall data from two stations are increasing significantly from periods 1 to 10 and they are increasing slowly after period 10. Among all the periods that we considered, at period 2, the return levels of the rainy season rainfall data from two station are closest to the average of the rainy season rainfall data (773.60 mm. and 698.20 mm., respectively). This could suggest that this model is a suitable choice for prediction of the rainy season rainfall data at period 2.

Table 4: The estimates of return levels and 95% confidence intervals of return levels based on the profile likelihood method for the rainy season rainfall data in Samoeng and Mae Tha stations.

Return period	Samoeng			Mae Tha		
	Lower confidence limit	Return level	Upper confidence limit	Lower confidence limit	Return level	Upper confidence limit
1	306.82	331.01	363.45	283.02	304.56	333.66
2	701.93	757.28	831.51	633.83	682.07	742.24
3	818.35	882.88	969.42	735.85	791.86	867.52
4	885.59	955.42	1049.07	794.56	885.04	936.74
5	932.11	1005.62	1104.19	835.10	898.67	984.53
10	1055.82	1139.08	1250.73	942.58	1014.33	1111.25
15	1118.04	1206.20	1324.43	996.47	1072.33	1174.79
20	1158.90	1250.28	1372.84	1031.82	1110.36	1216.45
30	1212.74	1308.37	1436.61	1078.32	1160.41	1271.28
40	1248.68	1347.15	1479.19	1109.33	1193.78	1307.84
50	1275.45	1376.02	1510.90	1132.40	1218.60	1335.04
100	1353.28	1459.10	1603.10	1199.41	1290.71	1414.03

4 CONCLUSION

Determine appropriate distributions for rainfall data during the rainy season can predict the return levels of the rainy season rainfall data which can be useful for flood irrigation planning and efficient water management. This study aims to find an appropriate distribution for the rainy season rainfall data. Comparing the LBWR distribution with Rayleigh, Weibull and Weibull-Rayleigh distributions, the results showed that the LBWR distribution is outperformed and suitable in fitting the rainy season rainfall data. Using the LBWR distribution, the return levels of the rainy season rainfall data showed that the levels at Samoeng station might be higher than the levels at Mae Tha station. This can imply that Samoeng district may have a higher risk compared to Mae Tha district. Therefore, the LBWR distribution can be another choice for modelling of the rainy season rainfall data.

ACKNOWLEDGEMENT

This study was supported by a teaching assistant and research assistant (TA/RA) scholarship from the Graduate School, Chiang Mai University, Chiang Mai, Thailand.

REFERENCES

[1] Rittima, "Hydrology." http://irre.ku.ac.th/books/pdf/48.pdf. (accessed Feb. 5, 2021).

[2] Intergovernmental Panel on Climate Change, Managing the risks of extreme events and disasters to advance climate change adaptation. Cambridge, NY, USA: Cambridge University Press, 2012.

[3] Hydro - Informatics Institute. "Flood event record." National Hydroinformatics Data Center. https://www.thaiwater.net/ web/index.php/archive.Html. (accessed Feb. 5, 2021).

[4] T. Paladkong, P. Nanthaprut, and K. Chaisee, "Extreme value analysis and return estimation of rainfall data in Bangkok metropolitan region," in Proceedings - the National Undergraduate Conference on Statistics, 2019, pp. 119 -127.

[5] C. Chaleeraktrakoon, Analysis of the flood frequency. Bangkok, Thailand: Thammasat Printing house, 2008.

[6] C. Pudprommara and M. Khamkong, "Statistical modeling to fit seasonal rainfall data from the Doisaket rain gauge station in Thailand," Journal of Applied Statistics and Information Technology, vol. 5, no. 1, pp. 1–9, 2020.

[7] F. Yusof and F. Hui-Mean, "Use of statistical distribution for drought analysis," Applied Mathematical Sciences, vol. 6, no. 21, pp. 1031–1051, 2012.

[8] G., M. Mansoor and S. Provost, "The Harris extended Lindley distribution for modeling hydrological data," Chilean Journal of Statistics, vol. 10, no. 1, pp. 77–94, 2019.

[9] T. Hussain, H. Bakouch, and C. Chesneau, "A new probability model with application to heavy-tailed hydrological data," Environmental and Ecological Statistics, vol. 26, no. 1, pp. 127–151, 2019.

[10] M. Ganji, H. Bevrani, G. Hami, and S. Zabihi, "The Weibull–Rayleigh distribution, some properties, and applications," Journal of Mathematical Sciences, vol. 218, no. 3, pp. 269–277, 2016.

[11] N. Nanuwong, "The generalized of mixture Pareto distributions," Ph.D. dissertation, Dept. Statist., Kasetsart Univ., Bangkok, Thailand, 2015.

[12] R. Fisher, "The effect of methods of ascertainment upon the estimation of frequencies," Annals of Eugenics, vol. 6, no. 1, pp. 13–25, 1934.

[13] C. Rao, "On discrete distributions arising out of methods of ascertainment," Sankhya: The Indian Journal of Statistics, Series A, vol. 27, no. 2, pp. 311–324, 1965.

[14] K. Das and T. Roy, "Applicability of length biased weighted generalized Rayleigh distribution," Advances in Applied Science Research, vol. 2, no. 4, pp. 320–327, 2011.

[15] K. Das, D. Tanusree, and R. Deb, "On some length-biased weighted Weibull distribution," Advances in Applied Science Research, vol. 2, no. 5, pp. 465–475, 2011.

[16] K. Mir, A. Ahmed, and J. Reshi, "Structural properties of length biased beta distribution of first kind," American Journal of Engineering Research, vol. 2, no. 2, pp. 1–6, 2016.

[17] T. Chaito and M. Khamkong, "The length-biased Weibull-Rayleigh distribution for application to hydrological data," Lobachevskii Journal of Mathematics, in press.

[18] Hydrology and Water Management Center for the Upper Northern Region of Thailand. "Monthly and annual rainfall data." Rainfall data. https://www.hydro-1.net/. (accessed Feb. 1, 2021).

[19] Meteorological Department of Thailand. "Thailand season." Seasons, seasons of the world, seasons of Thailand. https://www.tmd.go.th/info/info.php?FileID=. (accessed Feb. 15, 2021).

[20] R Core Team. "R: Language and environment for statistical computing." R. https://www.R-project.org/. (accessed Mar. 12, 2021).

[21] T. Anderson and D. Darling, "Asymptotic theory of certain "goodness-of-fit" criteria based on stochastic processes," Annals of Mathematical Statistics, vol. 23, no. 2, pp.193–212, 1952.

[22] H. Akaike, "Information theory and an extension of the maximum likelihood principle," in Proceedings - 2nd International Symposium on Information Theory, 1973, pp. 267 -281.

[23] T. Chow, Applied hydrology. Hyderabad, India: Tata McGraw-Hill Education, 2010.

[24] H. Paldynski, "Modelling large claims in probability and home insurance-extreme value analysis," M.S. thesis, Dept. Math. Sci., Lund Univ., Lund, Sweden, 2015.