ตัวแบบทางสถิติที่เหมาะสมสำหรับปริมาณฝนรายฤดูกาลจากสถานีตรวจวัดปริมาณฝนอำเภอดอยสะเก็ด ประเทศไทย

ชูเกียรติ ผุดพรมราช[1] และ มานัดถุ์ คำกอง[2*]

Received: 26 February 2020; Revised: 6 April 2020; Accepted: 14 April 2020

บทคัดย่อ

ตัวแบบทางสถิติสำหรับข้อมูลปริมาณฝนรายฤดูกาลของประเทศไทยมีความสำคัญต่อการวางแผนและการบริหารจัดการด้านอุทกวิทยาของประเทศ การศึกษาครั้งนี้มีวัตถุประสงค์คือ 1. เพื่อหาการแจกแจงที่มี 2 พารามิเตอร์ที่เหมาะสมสำหรับปริมาณฝนรายฤดูกาลของสถานีตรวจวัดปริมาณฝนอำเภอดอยสะเก็ด 2. เพื่อวิเคราะห์อิทธิพลของฤดูกาลที่มีต่อการแจกแจงปริมาณฝน การแจกแจงที่ใช้ในงานศึกษาในครั้งนี้ ได้แก่ การแจกแจงไวบูล การแจก แจงแกมมา การแจกแจงล็อกนอร์มอล การแจกแจงปรกติ การแจกแจงแบบเลขชี้กำลังลินด์เลย์ และการแจกแจงแบบเลขชี้กำลังนัยทั่วไป ซึ่งการแจกแจงที่กล่าวไปข้างต้นจะถูกนำมาใช้เพื่อหาการแจกแจงที่เหมาะสมที่สุดสำหรับข้อมูลปริมาณฝนรายฤดูกาลของสถานีดอยสะเก็ด ผลการศึกษาพบว่าการแจกแจงแกมมาเป็นการแจกแจงที่เหมาะสมสำหรับข้อมูลปริมาณฝนรายฤดูกาลของสถานีดอยสะเก็ด เนื่องจากมีเกณฑ์สารสนเทศของอะกะอิเกะหรือเกณฑ์เอไอซี (Akaike Information Criterion: AIC) และค่าสถิติทดสอบแอนเดอร์สัน-ดาร์ลิ่ง (Anderson-Darling Test: AD) มีค่าต่ำที่สุด นอกจากนี้การวิเคราะห์การถดถอยแกมมาแสดงให้เห็นว่าปริมาณฝนในฤดูฝนมีผลต่อปริมาณฝนในฤดูหนาว ผลการศึกษาในครั้งนี้มีประโยชน์ต่อภาครัฐและหน่วยงานที่เกี่ยวข้อง ในการวางแผนเชิงกลยุทธ์และการบริหารจัดการน้ำเพื่อป้องกันภัยพิบัติที่เกิดจากฝนในพื้นที่อำเภอดอยสะเก็ดได้อย่างมีประสิทธิภาพ

คำสำคัญ: การคัดเลือกตัวแบบ, เบ้เชิงบวก, ข้อมูลปริมาณฝนรายฤดูกาลของประเทศไทย, การวิเคราะห์การถดถอยแกมมา, ความสัมพันธ์แกมมา

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

[1] *สาขาวิชาสถิติประยุกต์, คณะวิทยาศาสตร์และเทคโนโลยี, มหาวิทยาลัยราชภัฏสวนสุนันทา*

[2] *ภาควิชาสถิติ, คณะวิทยาศาสตร์, มหาวิทยาลัยเชียงใหม่*

Statistical Modeling to Fit Seasonal Rainfall Data from the Doisaket Rain Gauge Station in Thailand

Chookait Pudprommarat[1] and Manad Khamkong[2,*]

Received: 26 February 2020; Revised: 6 April 2020; Accepted: 14 April 2020

Abstract

The statistical modeling of Thai seasonal precipitation data is crucial for the effective planning and management of the country's hydrological operations. There are two aims of this study: 1) to find an appropriate two-parameter statistical distribution to represent seasonal rainfall data from the Doisaket rain gauge station in northern Thailand and 2) to analyze the effect of seasonality on the rainfall data distribution. Two-parameter distributions, namely Weibull, gamma, lognormal, normal, Lindley exponential, and generalized exponential, were used to determine the best-fitting model of seasonal rainfall data from the Doisaket rain gauge station in Thailand. It was found that the gamma distribution with two parameters was the best fit, as indicated by the minimum values for the Akaike information criterion and the Anderson-Darling goodness-of-fit criterion. In addition, gamma regression showed that the precipitation amount during the rainy season affects that in the cold season. The approach and outcomes of this study could be useful for involved government agencies to strategically plan and manage water resources and to effectively prevent rain-related disasters in the Doisaket area.

Keywords: Model selection, Positively skewed, Thai seasonal rainfall data, Gamma regression, Correlated gamma.

[*] *Corresponding author: manad.k@cmu.ac.th*
[1]*Applied Statistics, Faculty of Science and Technology, Suan Sunandha Rajabhat University*
[2] *Department of Statistics, Faculty of Science, Chiang Mai University*

1. Introduction

In Thailand, the selection of an appropriate statistical distribution for seasonal precipitation data is important for water management. The characteristics of rainfall data for Thailand are influenced by the southwest monsoon in the rainy season (mid-May to October) and the northeast monsoon in the cold season (November to February) (Chaowiwat et al., 2016). Statistical distributions for precipitation data are selected based on statistical information criteria, goodness-of-fit testing, or by graphical methods for convenience. Many hydrological researchers have determined suitable statistical distributions to fit the rainfall data for a variety of study areas and durations. McKee et al. (1993) used a gamma distribution to evaluate drought based on standardized precipitation. Markovic recommended two-parameter lognormal and gamma distributions for fitting the annual precipitation data in the western USA and southwestern Canada (Hydrology Papers no. 8, Colorado State University, 1965). Yue and Hashino (2007) investigated the statistical distribution of annual, seasonal, and monthly rainfall data in Japan and concluded that the Pearson type-III, log-Pearson type-III, and lognormal distributions are suitable for fitting precipitation data. Yusof and Hui-Mean (2012) considered the exponential, gamma, and Weibull distributions to fit the rainfall data in the state of Johor, Malaysia, and found that the Weibull distribution was the best.

In Thailand, Khamkong and Bookkamana (2015) investigated the generalized extreme value (GEV) distribution for the annual maxima of daily (AMR1) and two-day (AMR2) rainfall data in upper northern Thailand. They concluded that the data fit the GEV model for location parameter changes depending on a quadratic trend for only one station whereas the data from the others were successfully fitted with stationary GEV models from AMR1. Chaito and Khamkong (2018) and Chaito et al. (2019) investigated an appropriate transformation to determine the standardized precipitation index under gamma, Weibull, and Pearson Type-III distributions to best fit the monthly rainfall data in upper Northern Thailand. To effectively select an appropriate statistical model, there have been many studies that have compared a variety of methods for estimating parameters in a Weibull distribution (Chang, 2011; Werapun et al., 2015). The maximum likelihood estimation)MLE(method seems to be the best choice for parameter estimation for large sample sizes. Pakoksung and Takagi (2016) applied MLE methods to estimate the parameters of zero-inflated Weibull and gamma distributions and investigated the best-fitting statistical distribution for monthly rainfall data in Thailand .Furthermore, Jaithun and Khamkong (2017) evaluated the performance of three methods for estimating the parameters for a zero-inflated gamma distribution using the precipitation data from six rain gauge stations on the Yom River in northern Thailand. They concluded that the MLE and expectation-maximization algorithm methods were suitable for the parameter estimation method whereas the moment method was not. However, studies on the comparison of different statistical models for two-parameter distributions in fitting seasonal rainfall data and the effect of seasonal rainfall data variation in Thailand are relatively rare.

Therefore, there are two objectives of this study: 1) to find an appropriate two-parameter statistical distribution to fit the seasonal rainfall data from the Doisaket station in northern Thailand and 2) to analyze the effect of seasonality on the rainfall data distribution.

2. Methodology

In this section, we describe the characteristics of the data and the methodology for appropriate modeling of the seasonal rainfall data from the Doisaket rain gauge station.

2.1 Study area and data

The Doisaket rain gauge station measures the amount of precipitation around the Mae Kuang Udom Thara Dam, which is a sub-source of the Ping River in northern Thailand. It is important for precipitation data analysis toward the effective planning and management of water operations in this part of northern Thailand. In this study, monthly precipitation data were obtained during the period from February 1957 to January 2014 from the Hydrology and Water Management Center for the Upper Northern Region of Thailand (Hydrology and Water Management Center for Upper Northern Region Chiang Mai Thailand, 2015; in mm). The classification of Thai seasons by collected monthly rainfall data into four months periods follows the criterion of the Meteorological Department of Thailand (Agro-Meteorological Academic Group Meteorological Development, 2015): the summer season (February to May), the rainy season (June to September), and the winter season (October to January of the following year).

2.2 Statistical modeling of seasonal rainfall

An exponential distribution, Exp(λ), with scale parameter λ is the first-choice statistical distribution for analyzing lifetime data. Its probability density function (pdf) is given by

$$f(x \mid \lambda) = \lambda e^{-x\lambda}, \quad x > 0,\ \lambda > 0. \tag{1}$$

The mean and variance of an Exp(λ) distribution are $E(X) = \lambda^{-1}$ and $V(X) = \lambda^{-2}$. The cumulative distribution function (cdf) of Exp(λ) is as follows:

$$F(x) = 1 - e^{-x\lambda}, \quad x > 0,\ \lambda > 0. \tag{2}$$

Since some real datasets are complex and their distributions are not exponential, many researchers have tried to develop an appropriate statistical distribution by adding shape parameter α used in distributions such as Weibull(α, λ) with a cdf given by

$$F(x) = 1 - e^{-(x\lambda)^{\alpha}}, \quad x > 0,\ \lambda,\ \alpha > 0. \tag{3}$$

Furthermore, the generalized exponential distribution (Gupta and Kundu, 1999), E(α, λ), has a cdf as follows:

$$F(x) = \left(1 - e^{-x\lambda}\right)^{\alpha}, \quad x > 0,\ \lambda,\ \alpha > 0. \tag{4}$$

For a gamma distribution, gamma (α, λ) is the sum of the independent exponential random variables. Thus, the pdf of Gamma(α, λ) distribution can be denoted as

$$f(x \mid \alpha, \lambda) = \frac{\lambda^{\alpha} x^{\alpha-1} e^{-x\lambda}}{\Gamma(\alpha)}, \quad x > 0,\ \alpha, \lambda > 0, \tag{5}$$

where $\Gamma(.)$ is the gamma function. Note that a gamma distribution where $\alpha = 1$ is reduced to an exponential distribution. Gamma(α, λ), has a cdf as follows:

$$F\left(x\right) = \Gamma_{x/\lambda}\left(a\right) / \Gamma\left(\alpha\right),\ \alpha, \lambda > 0 \tag{6}$$

and $\Gamma_{x}(.) = \int_{0}^{x} t^{\alpha-1} e^{-t} dt$ is the incomplete gamma function.

The Lindley exponential distribution (Bhati et al., 2015), Lindley(α, λ), has a cdf that can be written as

$$F(x) = \frac{\left(1 - e^{-x\lambda}\right)^{\alpha}\left[1 + \alpha - \alpha \ln\left(1 - e^{-x\lambda}\right)\right]}{1 + \alpha},\ x > 0,\ \lambda,\ \alpha > 0. \tag{7}$$

For a statistical distribution, more parameters make the model more complex and the statistical inference in the model selection can become ambiguous (Burnham and Anderson, 2002). Moreover, when analyzing seasonal precipitation, we are interested in two-parameter positively skewed distributions. However, a large sample size can limit the distributed tend to the normal distribution, norm (μ, σ^2), with location parameter μ and scale parameter σ^2. The cdf for this can be written as

$$F(x) = \frac{1}{2}\left[1 + erf\left(\frac{x - \mu}{\sigma\sqrt{2}}\right)\right], \quad \sigma > 0\ , -\infty < x, \mu < +\infty, \tag{8}$$

where $erf\left(x\right) = \frac{2}{\sqrt{\pi}} \int_{0}^{x} \exp\left(-t^{2}\right) dt$ is an error function. The mean and variance of norm(μ, σ^2) are given by $E(X) = \mu$ and $V(X) = \sigma^2$, respectively. Alternatively, when of a normal distribution is positively skewed, we can take the logarithm to attain the lognormal distribution, lnorm(μ,σ^2). The cdf for this distribution is given by

$$F(x)=\frac{1}{2}\left[1+erf\left(\frac{\ln(x)-\mu}{\sigma\sqrt{2}}\right)\right],\quad x>0,\ \sigma>0\ ,-\infty<\mu<+\infty. \tag{9}$$

The parameters for all of the statistical distributions mentioned can be estimated by using the MLE.

2.3 Modeling the selection criteria

The choice of the most appropriate statistical distribution is important in hydrological events to rationalize the use of a particular function. Generalization for model selection based on information criteria, goodness-of-fit tests, etc., has also been considered. In practice, Stephens (1974) reported that the Anderson-Darling (AD) (Anderson and Darling, 1952), goodness-of-fit test statistic performed better than other goodness-of-fit tests and thus recommended its use for testing the fit of a positively skewed distribution. The AD test statistic can be written in the form

$$A^2=-n-\frac{1}{n}\sum_{i=1}^{n}(2i-1)\left[\ln\left(F(x_i)\right)+\ln(1-F(x_{n+1-i}))\right], \tag{10}$$

where n is the sample size, $F(\cdot)$ is the expected cdf, and x_i are the ordered data. For the model selection criteria are used the Akaike information criterion (AIC) (Akaike, 1973), can also be used as a model selection criterion for nested statistical distributions. It can be defined as

$$AIC=2k-2\ln LL, \tag{11}$$

where k is the number of parameters in the model and LL is the log-likelihood function for the model. Laio et al. (2009) and Baldassarre et al. (2009) mentioned that the AD test outperformed AIC in cases of statistical distributions with three parameters whereas AIC was better for two-parameter statistical distributions. Additionally, Dey and Kunde (2009) suggested that the largest maximized likelihood value is suitable to discriminate among three statistical distributions.

Consequently, the best-fitting model for seasonal precipitation data based on the maximized likelihood function and the smallest AIC and AD values indicates that the tested statistical distribution is appropriate (Ghitany et al., 2017; Hussain et al., 2017).

2.4 Generalized linear models

Normally, rainfall data form a positively skewed distribution (Ashkar, 2017). For this reason, a nonparametric correlation measure called Spearman's rank correlation coefficient (r_S) is appropriated for measuring the strength of the relationship between variables:

$$r_S=1-\frac{6\sum_{i=1}^{n}d_i^2}{n(n^2-1)}, \tag{12}$$

where n is the pair sample size and $d_i,\ i=1,\ 2,\ \ldots\ ,\ n$, for each pair of observations. Moreover, generalization to determine how a response variable depends on the value of another one can be determined by using a generalized linear model (GLM) (Nelder and Wedderburn, 1972). A GLM includes three components:

1) A random component which is the response variable ($Y_i,\ i=1,\ 2,\ \ldots\ ,\ n$) when given the values of the explanatory variables in the model and the associated probability distribution in an exponential family:

$$f(y\,|\,\theta,\phi)=\exp\left(\frac{y\theta-b(\theta)}{a(\phi)}+c(y,\phi)\right), \tag{13}$$

where θ is a natural parameter; ϕ is a dispersion parameter; and $a(.)$, $b(.)$, $c(.)$ are functions.

2) A linear predictor or systematic component which includes p explanatory variables ($X_1, X_2, \ldots, X_p$) and the relationships among them:

$$\eta = \beta^t X = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_p X_p . \tag{14}$$

3) A link function which specifies the relationship between the linear predictor and the expectation of the response given the values of the explanatory variables:

$$g(\mu) = \eta = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_p X_p . \tag{15}$$

In practice, the regression coefficients (β_i, $i = 0, 1, 2, \ldots, p$) are estimated by applying ordinary least squares regression. The gamma regression is based on the assumptions: 1) the dependent variable (Y) is gamma distributed with which probability density function (pdf) is given in Equation (5), 2) a linear predictor which includes p explanatory variables with unknown coefficients and the relationships among them in Equation (14) and 3) a link function is given in Equation (15).

3. Results and Discussion

To assess when the statistical modeling of seasonal rainfall data from the Doisaket rain gauge station is appropriate, the data were analyzed using the R statistical program (R Core Team, 2015). The results of the analysis are listed in Table 1.

3.1 Model selection criteria for seasonal rainfall data

For the Doisaket rain gauge station in the summer season, the precipitation had a maximum/minimum cumulative seasonal rainfall of 504.80/40.90 mm with a central tendency of 210.19 mm. The slight outlier (in 1990) in the maximum of the precipitation indicates a summer storm (Figure 1).

Table 1 Descriptive statistics of the precipitation data from the Doisaket rain gauge station

Season	Minimum	Maximum	Mean	Std. Deviation	Q_1	Q_2	Q_3
Summer	40.90	504.80	210.19	96.06	128.05	208.00	273.15
Rainy	394.90	1,158.30	746.50	191.26	616.15	734.50	887.75
Winter	22.20	328.00	159.24	76.28	113.85	136.30	209.35

Note. Precipitation in mm. Q_i, the i^{th} quartile of the data.

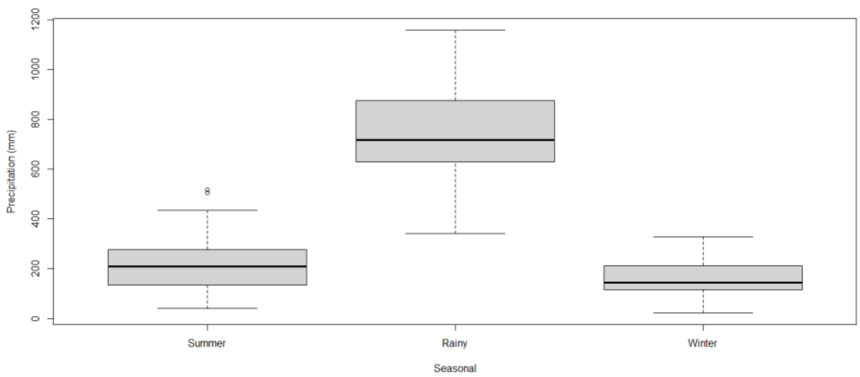

Figure 1 A box plot of the precipitation data from the Doisaket rain gauge station during different seasons

The maximum/minimum cumulative rainfall amounts were 1,158.30/394.90 mm with a central tendency of 746.50 mm for the rainy season (the highest) and 328.00 mm/22.20 mm with a central tendency of 159.24 mm for the winter season (the lowest). Table 2 and Figure 2 demonstrate that a Weibull distribution for the summer season and a gamma distribution for the rainy and winter seasons provided the best fit.

Table 2 Summary of the tested statistical distributions for the Doisaket rain gauge station precipitation data

Distribution	Summer			Rainy			Winter		
	LL	**AIC**	**AD**	**LL**	**AIC**	**AD**	**LL**	**AIC**	**AD**
Weibull	**-338.39**	**680.78**	**0.25**	-380.22	764.43	0.36	-324.69	653.38	0.67
Gamma	-338.43	680.85	0.35	**-379.49**	**762.98**	**0.22**	**-324.66**	**653.32**	**0.47**
lnorm	-340.69	685.38	0.71	-380.13	764.25	0.33	-326.91	657.83	0.59
norm	-340.58	685.15	0.42	-379.83	763.66	0.27	-327.44	658.88	1.24
Lindley	-339.39	682.78	0.52	-400.51	805.02	0.57	-325.40	654.80	0.48
GE	-339.29	682.58	0.50	-381.15	766.31	0.48	-325.30	654.59	0.47

Note. The values in bold indicate the best-fitting distribution.

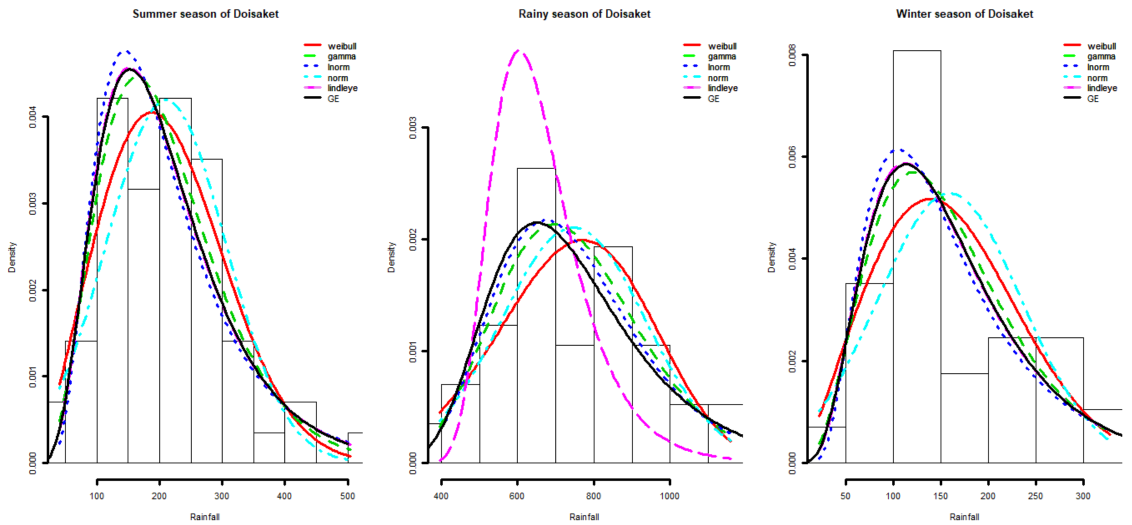

Figure 2 Comparison of histograms and theoretical densities for the tested statistical distributions for Doisaket rain gauge station precipitation data

3.2 Two correlated gamma random variables

Table 3 reports the precipitation amounts during the winter season. There is correlation between the precipitation amounts during the winter and rainy seasons at the 0.05 significance level (p-value < 0.05). The coefficient of determination ($R^2 = (0.28)^2 = 0.078$), equivalent to 7.8% variation in the precipitation amount in the winter season, can be accounted for by the variation in the precipitation amount during the rainy season.

Table 3 Spearman's correlation (p-value) for seasonal rainfall data measured at the Doisaket rain gauge station

Season	Summer	Rainy	Winter
Summer	1.00	-0.11 (0.41)	-0.02 (0.89)
Rainy		1.00	0.28 (0.03)
Winter			1.00

Gamma regression was applied to analyze the effect of the precipitation amount in the rainy season on the precipitation amount in the winter season (Table 4).

Table 4. Gamma regression estimates for seasonal rainfall data measured at the Doisaket rain gauge station

Source	Coefficient	Std. Error	t value	p-value
Intercept	4.498	0.254	17.699	0.000
Rainy	0.001	0.000	2.286	0.026

Outcome: Winter (log link).

The results in Table 4 indicate that there is a significant relationship between the precipitation amounts in the winter and rainy seasons. Moreover, the estimated gamma regression to predict the precipitation amount in the winter season ($\hat{y}$) given the precipitation amount in the rainy season (x) is

$$\hat{y} = \exp(4.498 + 0.001x). \tag{16}$$

It was found that an increase (decrease) in the precipitation amount in the winter season is associated with an increase (decrease) in the precipitation amount in the rainy season.

4. Conclusions

Determining the appropriate statistical distribution for seasonal precipitation is important for forecasting extreme events and for effective water management. The results based on the model selection criteria of the appropriate statistical distribution indicate that the gamma distribution is the best-fitting distribution for the seasonal precipitation data from the Doisaket rain gauge station in Thailand. Furthermore, gamma regression showed that the precipitation amount during the winter season is correlated with that in the rainy season, indicating that less precipitation in the rainy season was linked to more precipitation in the winter season and vice versa, a pattern that is expected to increase over time. The information obtained from our results could offer help to the agencies involved in planning and managing water resources in the Doisaket area to understand the effects of seasonal precipitation.

Acknowledgements

C. Pudprommarat is grateful to the Suan Sunandha Rajabhat University for financial support.

References

Agro-Meteorological Academic Group Meteorological Development. (2015). **Study on drought index in Thailand**. cited 2019 Jul 13 Retrieved from http//:www.tmd.go.th/info /info.php? FileID=71.

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. **Proceedings of the 2nd International Symposium on Information Theory**, 267 – 281.

Anderson, W., & Darling, D.A. (1952). Asymptotic theory of certain "goodness-of-fit" criteria based on stochastic processes. **The Annals of Mathematical Statistics,** 23(2), 193-212.

Ashkar, F. (2017). Model selection tools for hydrological frequency analysis: some new results. **Proceedings of the World Environmental and Water Resources Congress**, 384 – 389.

Baldassarre, G.D., Laio, F., & Montanari, A. (2009). Design flood estimation using model selection criteria. **Physics and Chemistry of the Earth**, 34(10-12), 606-611.

Bhati, D., Malik, M.A., & Vaman, H.J. (2015). Lindley–exponential distribution: properties and applications. **Metron**, 73(3), 335-357.

Burnham ,K.P., & Anderson, D. (2002). **Model selection and multimodel inference**. 2nd ed. Berlin-Heidelberg, New York: Springer.

Chaito, T., & Khamkong, M. (2018). A modified box and cox power transformation to determine the standardized precipitation index. **Songklanakarin Journal of Science and Technology,** 40(4), 867-877.

Chaito, T., Khamkong, M., & Murnta, P. (2019). Appropriate transformation techniques to determine a modified standardized precipitation index for the Ping River in northern Thailand. **EnvironmentAsia,** 12(3), 32-42.

Chang,T,P. (2011). Performance comparison of six numerical methods in estimating Weibull parameters for wind energy application. **Applied Energy,** 88(1), 272-282.

Chaowiwat, W., Sarinnapakprn, K., & Boonya-aroonnet, S. (2016). Bias correction of seasonal rainfall forecasts of Thailand from general circulation model by using the ratio of gamma CDF parameter method. **Naresuan University Engineering Journal**, 11(1), 7-13.

Dey, A.K., & Kundu, D. (2009). Discriminating among the log-normal, Weibull and generalized exponential distributions. **IEEE Transactions on Reliability**, 58(3), 416-424.

Ghitany, M.E., Song, P., & Wang, S. (2017). New modified moment estimators for the two-parameter weighted Lindley distribution. **Journal of Statistical Computation and Simulation,** 87(16), 3225-3240.

Gupta, R.D, & Kundu, D. (1999). Generalized exponential distributions. **Australian and New Zealand Journal of Statistics**, 44(2), 173 - 188

Hussain, T., Bakouch, H.S., & Iqbal, Z.A. (2017). New probability model for hydrologic event: properties and applications. Journal of Agricultural. **Biological and Environmental Statistics**, 23(1), 63-82.

Hydrology and Water Management Center for Upper Northern Region Chiang Mai Thailand. (2015). **Bureau of Water Management and Hydrology Royal Irrigation Department Thailand, Rainfall Data**. cited 2019 July 13 Retrieved from: http://hydro-1.net.

Hydrology Papers no. 8, Colorado State University, Fort Collins, Colorado, USA. (1965). **Probability of best fit to distributions of annual precipitation and runoff**. cited 2019 July 13 Retrieved from https://moun tain scholar. org/bitstream/ handle/10217/61285/HydrologyPapers _n8. pdf?sequence=1&isAllowed=y.

Jaithun, M., & Khamkong, M. (2017). Optimal parameter estimation for zero-inflated gamma distributions with application to rainfall data of Yom River in northern Thailand. **Proceedings of the 13th IMT-GT International Conference on Mathematics, Statistics and their Applications,** o50021-1-6.

Khamkong, M., & Bookkamana, P. (2015). Development of statistical models for maximum daily rainfall in upper northern region of Thailand. **Chiang Mai Journal of Science**, 42(4), 1044-1053.

Laio, F., Baldassarre, G.D., & Montanari, A. (2009). Model selection techniques for the frequency analysis of hydrological extremes. **Water Resources Research,** 45(7), 29-40.

Mckee, T.B., Doesken, N.J., & Kleist, J. (1993). The relationship of drought frequency and duration on time scale. **Proceedings of the 8th Conference on Applied Climatology**, 179-184.

Nelder, J.A., & Wedderburn, R.W.M. (1972). Generalized linear models. **Journal of the Royal Statistical Society: Series A**, 1972; 135(3): 370-384.

Pakoksung, K., & Takagi, M. (2016). Mixed Zero-Inflation Method and Probability Distribution in Fitting Daily Rainfall Data. **Engineering Journal**, 21(2), 63-80.

R Core Team. (2015). **R: Language and Environment for Statistical Computing**. cited 2019 July 13 Retrieved from https://www.r-project.org/foundation.

Stephens, M.A. (1974). EDF statistics for goodness of fit and some comparisons. **Journal of the American Statistical Association**, 69(34), 730-737.

Werapun, W., Tirawanichakul, Y., & Waewsak, J. (2015). Comparative study of five methods to estimate Weibull parameters for wind speed on Phangan Island, Thailand. **Energy Procedia.** 79, 976-981.

Yue, S., & Hashino, M. (2007). Probability distribution of annual, seasonal and monthly precipitation in Japan. **Hydrological Sciences Journal**, 52(5), 863 - 877.

Yusof, F., & Hui-Mean, F. (2012). Use of statistical distribution for drought analysis. **Applied Mathematical Sciences**, 6(21), 1031 - 1051.