Conference Paper No: PF-08

Copula-based drought severity-duration-frequency analysis for Anuradhapura and Puttalam in the dry zone of Sri Lanka

W.R.P.M.S.S. Wijesundara[1*] and K. Perera[2]

[1]Computing Center, Faculty of Engineering, University of Peradeniya, Sri Lanka
[2]Department of Engineering Mathematics, Faculty of Engineering, University of Peradeniya, Sri Lanka
sachinisw1515@gmail.com*

Abstract

Drought is a severe problem in many areas of Sri Lanka, where rainfall amounts are low and extremely high due to climatic changes. To reduce the negative consequences of droughts, it is important to understand the drought characteristics (drought duration and drought severity) and their associations. Therefore, to build a drought severity-duration-frequency (SDF) relationship, a probabilistic technique is proposed using rainfall data from 1996 to 2018 in the two districts Anuradhapura and Puttalam in the dry zone of Sri Lanka. Drought characteristics were defined using 3-month standardized precipitation index (SPI) and Copulas are employed to derive the joint distribution function. Occurrences of 41 droughts from both stations were identified. The derived SDF relationship is a function of marginal distribution functions of drought characteristics linked by a copula. Log-normal distribution and Gamma distribution were identified as the best marginal distribution to represent drought duration and drought severity, respectively, using AIC, BIC and Kolmogorov-Smirnov test. Gaussian copula and Frank copula were identified as the best among the other five copulas based on AIC, BIC and Cramer-Von Mises statistics. Both Log-normal distribution and Gamma distribution along with the Gaussian copula and Frank copula combined to derive the joint distribution of Anuradhapura and Puttalam, respectively. Joint return periods in terms of recurrence intervals were calculated and derived the SDF curves. According to the derived SDF curves, drought severity in Puttalam is greater than those in Anuradhapura for a given recurrence interval and drought duration.

Keywords

Copula, Drought, Joint distribution, Marginal distributions, SPI

Introduction

In Sri Lanka, drought is one of the most injurious natural hazards that has islandwide effects. It can cause severe damages to the agriculture, socioeconomic structure of humans, and also the ecosystem. Mainly there are three climatic zones in Sri Lanka as dry, wet, and intermediate which are categorized according to the total annual rainfall received by those climatic regions. The dry zone is the main source that provides both irrigated and rainfed crops for the Sri Lankan food industry. Dryzone receives less than 1,800 mm of rainfall per year throughout both the south-western and north-eastern monsoon seasons. Since there are inadequate irrigation systems and high rainfall fluctuations prevail within the dry zone, areas like Anuradhapura and Puttalam are experienced a prolonged drought stemming from reduced precipitation levels. Anuradhapura is one of the largest crop production areas in Sri Lanka which is frequently affected by droughts. Since Puttalam and Anuradhapura are situated very close to each other in the dry zone it is important to identify the dependence nature of drought characteristics of those two areas. As it is impossible to evade droughts, drought preparedness can be developed while managing drought impacts to identify the drought severity as it plays a major role in risk management.

Each drought seasons can be categorized by two attributes that correlate to each other: Severity and Duration (Shiau & Shen, 2001). Many indices have been developed for modeling a drought event, such as Palmer Drought Severity Index (Palmer, 1965) and the Standardized Precipitation Index (SPI) (Mckee et al., 1993). Since SPI is simple, spatially invariant, and just requires monthly precipitation data to calculate, (Mckee et al., 1993) drought duration and drought severity derived from SPI used to describe the droughts.

However joint multivariate models of droughts are difficult to develop because different distribution functions are frequently used to fit various drought attributes such as severity and duration. The use of copulas to link different marginal distributions to generate the joint multivariate distribution function alleviates this challenge as it has the advantage of not requiring any assumptions about the variables being independent, normal or having the same type of marginal distributions (Zhang & Singh, 2007b). Most research on multivariate distributions using copulas have been focused on rainfall and floods (Shiau, 2006; Shiau & Modarres, 2009; Zhang & Singh, 2007a, 2007b) which have been used for rainfall frequency analysis, flood frequency analysis and drought frequency analysis. Shiau in 2009 has developed a Copula-based drought severity-duration frequency analysis to assess the simultaneous multi-attributes of droughts occurred in two regions of Iran (Shiau & Modarres, 2009). Dalezios in 2000 developed drought severity-duration frequency (SDF) curves to analyze droughts and wet periods in Greece (Dalezios et al., 2000).

This study aims to derive copula-based drought severity-duration-frequency (SDF) curves for Anuradhapura and Puttalam by identifying the best joint probability distribution function of drought duration and drought severity based on the best mariginal distributions and best Copula. Identified best joint distribution of Anuradhapura and Puttalam will be used to generate the joint return periods of drought events for managing the impacts of droughts.

Methodology

Description of data

Monthly rainfall data for 23 years (1996–2018) were collected from two meteorological stations, Anuradhapura and Puttalam in the dry zone of Sri Lanka.

Defining drought using SPI

SPI values were often estimated over time intervals of 1, 3, 6, 9, and 12 months (Ganguli & Reddy, 2012). SPI on 3-months was employed in this study to identify short-term seasonal drought occurrences. As drought is defined when the values of SPI fall below zero, a period with negative SPI values is considered a drought event while it ends when SPI reaches a value of 0. Drought duration is the continuous negative SPI periods in a drought event. Whereas drought severity is the sum of cumulative SPI periods within a drought event (Shiau & Modarres, 2009).

Modeling the joint cumulative distribution function (JCDF)

Pearson correlation coefficient and Kendall's tau (τ) measure were used to identify the dependence nature of the drought characteristics. Exponential, Gamma, Normal, Log-normal, Weibull and Logistic distributions were fitted and identified the best marginal distribution out of them using Akaike's information criteria (AIC) and Bayesian information criteria (BIC). The Kolmogorov-Smirnov (K-S) test was then employed to confirm the results. Parameters of the distributions were estimated using the maximum likelihood method.

The copula method which allows to study and measure the relationships between random variables is first developed by Sklar (Sklar, 1959) to describe the function that links univariate distribution functions to multivariate distribution function. Tail dependency of drought characteristics can be checked by ranked scatter plots. Five copulas were tested (Gumbel-Hourgaard, Joe, Clayton, Frank, and Gaussian) and identified the best copula using AIC and BIC. Cramer-von Mises S_n statistic was used for further confirmation of the goodness of fit test while the estimation of S_n was done by the methods inverting the tau parameter and parametric bootstrapping (Genest & Rémillard, 2008).

Then generated the Joint Cumulative Distribution Function (JCDF) of drought severity and drought duration using the best fitted marginal distributions and the best copula.

Joint return periods

The joint return period $T_{X,Y}(x,y)$ of a bivariate random variable (X,Y) corresponding to a value (x,y) is given by;

$$T_{X,Y}(x,y) = \left(\frac{1}{1-F_{XY}(x,y)}\right) \tag{1}$$

Where, $F_{XY}(x,y)$ is the joint cumulative distribution function and $F_{XY}(x,y) = P(X \leq x, Y \leq y)$ is the probability of two events $X \leq x, Y \leq y$ (Zhang & Singh, 2007a).

Copula-based drought severity-duration-frequency (SDF) relationship

Since each drought event is treated as a bivariate random variable, joint recurrence intervals can be used to show the relationship between drought severity, drought duration, and frequency (in terms of recurrence interval) (Shiau & Modarres, 2009).

Results and Discussion

During the study period, Anuradhapura has experienced the longest drought duration in 2002 while Puttalam in 2010 & 2012. The most severe droughts in Anuradhapura occurred in 2002 and 2017 and in Puttalam it was in 2008, 2012 and 2017.

The mean annual rainfall for Anuradhapura and Puttalam stations are 117 mm and 97 mm, respectively. The number of droughts that occurred during the study period for both stations is 41. Although it is the same for both stations, the standard deviation of drought duration and severity for Puttalam are larger than those of Anuradhapura. This fact indicates that drought characteristics are highly fluctuating for Puttalam.

Calculated Pearson's correlation coefficients for drought duration and severity of Anuradhapura and Puttalam are 0.876 and 0.844, respectively, while Kendall's tau values for the drought characteristics of Anuradhapura and Puttalam are 0.770 and 0.768, respectively. Results showed a strong positive relationship. Thus the drought duration and drought severity of each station should be modeled jointly.

Drought duration and drought severity of the two stations can be fitted to Lognormal and Gamma distributions, respectively, as they have low AIC and BIC values while the K-S test result confirmed it further. Parameters of the selected distributions are estimated by the maximum likelihood method and are summarized in Table 1. The critical values for a sample size of 43 (Anuradhapura) and 34 (Puttalam) are 0.00045 and 0.01388 respectively, at the 10% significance level. The maximum deviations between observed data and proposed distributions of drought characteristics for Anuradhapura and Puttalam are also showed in Table 1.

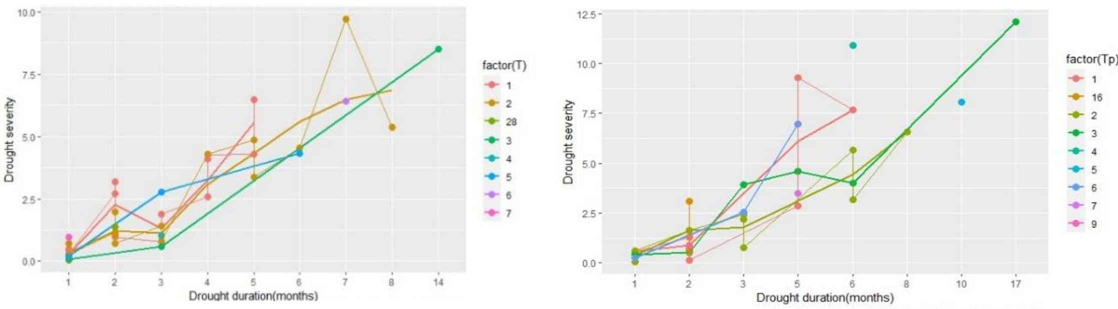
Table 1. *Estimated parameters of the proposed distributions.*

Character	**Estimates**	**Anuradhapura**	**D max**	**Puttalam**	**D max**
Duration	meanlog/θ	0.870	0.208	1.078	0.172
	sdlog/α	0.733		0.728	
Severity	shape/α	0.773	0.099	0.873	0.112
	scale/β	0.333		0.269	

Gaussian copula and Frank copula are the best among other copulas to represent two stations as they have low AIC and BIC values and no upper or lower tail dependency. Calculated Cramer-von Mises S_n statistics for the Gaussian copula and Frank copula are 0.028 and 0.022, respectively**.** Inverting Kendall's tau and parametric bootstrapping is used to obtain S_n values. The values of the copula parameters of Anuradhapura and Puttalam are 0.876 and 8.081, respectively.

Then Gaussian copula was used to combine the identified Gamma marginal distribution and Log-normal distribution to create the joint distribution of Anuradhapura while the Frank copula was used to combine the identified Gamma marginal distribution and Log-normal distribution to create the joint distribution of Puttalam.

The developed drought SDF curves for Anuradhapura and Puttalam with selected recurrence intervals generated using equation 1 are sketched in Figure 1. These SDF curves show an increasing pattern in drought severity for the increasing stages of drought duration.



(a) (b)

Figure 1. *The drought SDF curves for Anuradhapura (a) and Puttalam (b).*

According to Figure 2, we can identify that drought severity at Puttalam is greater than the value at Anuradhapura for the same duration and recurrence interval. Therefore, more severe droughts can occur in Puttalam than in Anuradhapura.

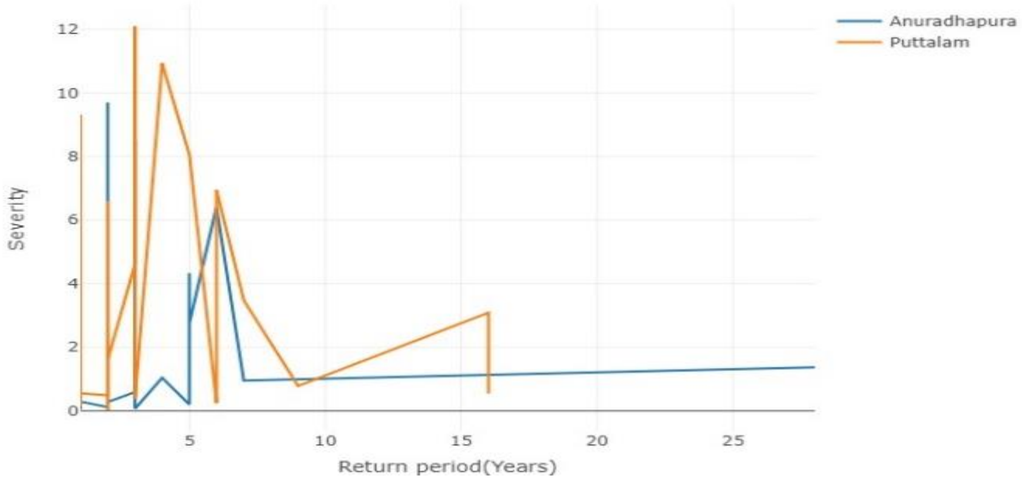


Figure 2. *The drought severity limits with various recurrence intervals for Anuradhapura and Puttalam.*

Conclusion

Understanding drought characteristics is a vital step in developing effective mitigation measures for droughts-related problems. In this study, 3-month SPI series was used and occurrences of 41 drought events from both stations were identified. Anuradhapura region has experienced 2 most severe droughts in 2002 and 2017 while the Puttalam region has experienced 3 most severe droughts in 2008,2012 and 2017. The Kendall's tau correlation and Pearson correlation test results revealed that there is a positive relationship between each drought variable. For both Anuradhapura and Puttalam marginal distributions of drought duration fit to a Log-normal distribution while marginal distributions of drought severity fit to a Gamma distribution quite well among other distributions.

The Gaussian copula and the Frank copula were found to be the best copulas for Anuradhapura and Puttalam, respectively. The Gaussian copula was used to combine the identified Gamma marginal distribution and Log-normal distribution to create the joint distribution of Anuradhapura while the Frank copula was used to combine the identified Gamma marginal distribution and Log-normal distribution to create the joint distribution of Puttalam.

Proposed method is used to create drought SDF curves with various recurrence intervals for the two rain gauge stations in the dry zone of Sri Lanka. These SDF curves show an increasing pattern. That is at most drought severity increases with drought duration, but not at a continuous rate. When comparing the generated drought SDF curves, drought severity at Puttalam is often greater than the value in Anuradhapura for the same period and recurrence interval. Therefore, highly fluctuating rainfall exists in the Puttalam region leads to the occurrences of more severe droughts than those in Anuradhapura. Thus based on the proven results government can develop drought preparedness while managing drought impacts within the dry zone of Sri Lanka.

References

Dalezios, N. R., Loukas, A., Vasiliades, L., & Liakopoulos, E. (2000). Severity-duration-frequency analysis of droughts and wet periods in Greece. *Hydrological Sciences Journal*, *45*(5), 751–769. https://doi.org/10.1080/02626660009492375

Ganguli, P., & Reddy, M. J. (2012). Risk Assessment of Droughts in Gujarat Using Bivariate Copulas. *Water Resources Management*, *26*(11), 3301–3327. https://doi.org/10.1007/s11269-012-0073-6

Genest, C., & Rémillard, B. (2008). Validity of the parametric bootstrap for goodness-of-fit testing in semiparametric models. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, *44*(6), 1096–1127. https://doi.org/10.1214/07-AIHP148

Mckee, T. B., Doesken, N. J., & Kleist, J. (1993). *The relationship of drought frequency and duration to time scales*. *January*, 17–22.

Palmer, W. C. (1965). Meteorological Drought. In *U.S. Weather Bureau, Res. Pap. No. 45* (p. 58). https://www.ncdc.noaa.gov/temp-and-precip/drought/docs/palmer.pdf

Shiau, J. T. (2006). *Fitting Drought Duration and Severity with Two-Dimensional Copulas*. 795–815. https://doi.org/10.1007/s11269-005-9008-9

Shiau, J. T., & Modarres, R. (2009). *Copula-based drought severity-duration-frequency analysis in Iran*. *489*(June), 481–489. https://doi.org/10.1002/met

Shiau, J. T., & Shen, H. W. (2001). R a h d d s. *Water Resources*, *February*, 30–40.

Sklar, M. (1959). Fonctions de Répartition à n Dimensions et Leurs Marges. *Open Access,Publications de l'Institut Statistique de l'Université de Paris*, *229–231*(8).

Zhang, L., & Singh, V. P. (2007a). *Bivariate rainfall frequency distributions using Archimedean copulas*. 93–109. https://doi.org/10.1016/j.jhydrol.2006.06.033

Zhang, L., & Singh, V. P. (2007b). *Gumbel – Hougaard Copula for Trivariate Rainfall Frequency Analysis*. *August*, 409–419.