JOINT DISTRIBUTION OF RAINFALL AND TEMPERATURE IN ANURADHAPURA, SRI LANKA USING COPULAS

D. G. T. C. Keerthirathne[1], K. Perera[2]

[1]Postgraduate Institute of Science, University of Peradeniya, Peradeniya, Email:thilinkeerthirathne@gmail.com
[2]Department of Engineering Mathematics, Faculty of Engineering, University of Peradeniya, Peradeniya, Email:kanthip@pdn.ac.lk

ABSTRACT

This study aims at deriving the best joint distribution between rainfall and temperature for monsoon seasons and months using the copula method for Anuradhapura region in Sri Lanka. The totals of average values of maximum and minimum of daily temperature and the totals of daily rainfall were used for this study. Kendall's tau(τ) was used to find the correlation between two variables for monsoon seasons and months. The distributions Gamma, Lognormal, Weibull, Exponential, Logistic, t and Gaussian were used to obtain the marginal distributions for both variables, while Frank, Gaussian and Student-t copulas were used since there was a negative relationship between the variables. The Akaike Information Criteria(AIC) and the Bayesian Information Criteria(BIC) were used to identify the best marginal distributions and the best copulas. The correlations were negative and significant (p-value < 0.05) for the First Inter Monsoon Season (FIMS), South West Monsoon Season (SWMS) and for the months February, April, May, June, July, August and September. Hence, the copula method was used to identify the joint distribution. The Gamma distribution for rainfall and the Weibull distribution for temperature were identified as the best marginal distributions for the FIMS, while the Weibull distribution for both variables in SWMS. The Exponential distribution for rainfall and the Weibull distribution for temperature were the best distributions for the above mentioned months except April. Gamma distribution and Weibull distribution were the best fitted marginal distributions for both variables in April. The Frank copula was identified as the best copula for both monsoon seasons and the months. The joint distributions between the two variables were fitted using the identified marginal distributions and the Frank copula. Using the fitted joint distributions, the return periods were calculated for maximum rainfall and corresponding temperature values and the maximum temperature and corresponding rainfall values for both seasons and months.

Key words: Copula, Joint Distribution, Marginal Distribution, Return Periods, Bayesian Information Criteria, Akaike Information Criteria

1. INTRODUCTION

Agriculture in Sri Lanka mainly depends on the rice production. Although there are many rice cultivated districts in Sri Lanka, Anuradhapura district acts a strong role among them. Accurate modeling of multivariate weather distributions would allow farmers to make better decisions on their cultivations. Rainfall and Temperature are important climatic inputs for Agriculture.

Sri Lanka has two main monsoon seasons namely south west monsoon season from May to September and north east monsoon season from December to February. Between these two monsoon seasons, there are two inter monsoon periods. The first inter monsoon season is from March to April and the second inter monsoon season is from October to November.

This research was focused to obtain the joint distributions for rainfall and temperature in Anuradhapura, Sri Lanka for monsoon seasons and the months using copulas.

Cong and Brady have found the interdependency between rainfall and temperature using copula analysis. They have observed the seasonal and non-seasonal variations of rainfall and temperature. They have found the relationship between monthly temperature and the rainfall and they selected negatively related months to fit the suitable copula and have found that the student copula is the most suitable copula for these months [3].

Aghakouchak *et al.* have found that Copulas are advantageous because they can model joint distributions of random variables with greater flexibility both in terms of marginal distributions and the dependence structure [1]. AghaKouchak *et al.* applied two different elliptical copula families to temperature and rainfall, namely, Gaussian and t-copula, to simulate the spatial dependence of rainfall and found that using the t-copula might have significant advantages over the well-known Gaussian copula particularly with respect to extremes [2].

Using annual data Huang *et al.* also found a negative correlation between rainfall and temperature [4].

2. METHODOLOGY

2.1. Data Used

The daily rainfall (mm) and maximum/minimum temperature ($^{\circ}C$) data for Anuradhapura district from 1951 to 2012 were obtained from Meteorological department, Colombo were used for this analysis.

2.2. Procedure for Analysis

By using the above mentioned data total rainfall in particular months and seasons and the temperatures were calculated. Kendall's tau(τ) was used to determine the correlation between the two variables for each monsoon seasons and months. The Gamma, Lognormal, Weibull and Exponential distributions were used to fit the marginal distributions for the variables. Since there were some zero values in monthly rainfall data, Exponential, t and Logistic distributions were used to fit the marginal distributions Best fitted marginal distributions for each monsoon seasons and months were selected using Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC). Copula parameters (θ) were determined using Kendall's tau correlation values. The best copula was selected using AIC and BIC values. The joint distributions were obtained by combining the identified marginal distributions using the selected best copula function. Using the fitted joint distributions, the return periods were calculated.

2.2.1. Definition of Copula

Copulas are joint distributions of uniform variables. Thus copulas provide a method for describing joint distribution for correlated variables. Consider p uniform variables $U_1, U_2, \dots\dots, U_p$ on the interval [0,1]. The copula function C is defined as,

$$C(u_1, u_2, \dots u_p) = Pr(U_1 \le u_1, \dots, U_p \le u_p)$$

2.2.2. Frank Copula

The frank copula is a symmetric Archimedean copula. Its function is given by eq. (01).

$$C_\theta(u, v) = -\frac{1}{\theta} \ln\left(1 + \frac{(e^{-\theta u}-1)(e^{-\theta v}-1)}{e^{-\theta}-1}\right) \quad (01)$$

2.2.3. Joint distribution

Let a joint distribution function $H(x, y)$ with the marginal distributions $F_X(x)$ and $F_Y(y)$ which are marginal distribution function of X and Y. Then there exists a function such that eq. (02).

$$H(x, y) = C(F_X(x), F_Y(y)) \quad (02)$$

where, C is the copula.

To construct the copula, let $U = F_X(x)$ and $V = F_Y(y)$. If U and V are the marginal distributions then C can be defined by,

$$C(u, v) = H(F_X^{-1}(u), F_Y^{-1}(v)) \quad (03)$$

Using equations

$$C(F_X(x), F_Y(y)) = H\left(F_X^{-1}(F_X(x)), F_Y^{-1}(F_Y(y))\right)$$

$$= H(x, y) \quad (04)$$

where, function of $C(u, v)$ is the copula.

2.2.5. Return Periods

Return period is known as a recurrence interval time between many kinds of events like weather events (Rainfall and temperature etc.).

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 eq. (05).

$$T_{x,y}(x, y) = \left(\frac{1}{1 - F_{XY}(x, y)}\right) \quad (05)$$

where,

$F_{XY}(x, y)$ is the joint cumulative distribution function and

$F_{XY}(x, y) = P(X \le x, Y \le y)$ is the probability of two events $X \le x$ and $Y \le y$.

The joint return periods of an event expressed by the eq. (05) represents that either X or Y or both values are exceeded.

3. RESULTS

There is a negative relation between between two variable for both monsoon seasons and for the months (Kendall's tau p-value <0.05). The fitted marginal distributions and their parameters for both monsoon seasons are shown in Table 1. Table 2 and Table 3 show the AIC and BIC values

Table 1: Parameters of marginal distributions of rainfall and temperature in FIMS and SWMS

Distribution	Parameter	Rainfall		Temperature	
		FIMS	SWMS	FIMS	SWMS
Gamma	Shape	5.67	4.06	796.53	30.13
	Rate	0.02	0.01	0.45	1.25
Lognormal	Mean	5.38	5.28	7.47	8.39
	SD	0.44	0.61	0.03	0.02
Weibull	Shape	2.53	2.14	33.99	55.47
	Scale	270.7	259.4	1782.8	4463.4
Exponential	Rate	0.004	0.004	0.104	0.041

Table 2: AIC and BIC values of fitted distributions for FIMS

Probability distribution	Rainfall		Temperature	
	AIC	BIC	AIC	BIC
Gamma	732	737	682	687
Lognormal	736	740	684	688
Weibull	735	739	671	675
Exponential	792	794	1035	1037

Table 3: AIC and BIC values of fitted distributions for SWMS

Probability distribution	Rainfall		Temperature	
	AIC	BIC	AIC	BIC
Gamma	668	672	670	674
Lognormal	675	679	671	675
Weibull	664	668	648	652
Exponential	697	699	1016	1018

By using minimum AIC and BIC values the Gamma distribution for rainfall and the Weibull distribution for temperature were identified as the best marginal distributions for the FIMS, while the Weibull distribution for both variables in SWMS.

Similarly the Exponential distribution for rainfall and the Weibull distribution for temperature were identified as the best marginal distributions for the months except April. Gamma distribution and Weibull distribution were the best fitted marginal distributions for both variables in April.

Then copula parameters were calculated for all months and monsoon seasons using Kendall's tau correlation values. Table 4 shows the copula parameters for FIMS and SWMS.

Table 4: Copula parameters for FIMS and SWMS

Copula	Parameter (θ)	
	FIMS	SWMS
Frank	-1.6092	-3.8771
Gaussian	-0.2704	-0.5606
Student-t	-0.3705	-0.5610

Table 5: AIC and BIC values of different copulas for FIMS and SWMS

Copula	FIMS		SWMS	
	AIC	BIC	AIC	BIC
Frank	-1.12	0.98	10.87	12.98
Gaussian	1.36	3.47	51.56	53.67
Student-t	12.02	16.24	44.93	49.15

According to AIC and BIC values in table 5, the best copula for FIMS and SWMS is the Frank copula. Similarly, the Frank copula fits well for all the months. By combining the identified best marginal distributions using the selected copulas, joint distributions for monsoon seasons and months were obtained. The joint distributions for monsoon season are given below.

The joint distribution for FIMS is:

$$H(x,y) = \frac{1}{1.6092}\ln\left(1 + \frac{(e^{1.6092F_X(x)}-1)(e^{1.6092F_Y(y)}-1)}{e^{1.6092}-1}\right) \quad (06)$$

The joint distribution for SWMS is:

$$H(x,y) = \frac{1}{3.8771}\ln\left(1 + \frac{(e^{3.8771F_X(x)}-1)(e^{3.8771F_Y(y)}-1)}{e^{3.8771}-1}\right) \quad (07)$$

where

x- Rainfall and y-Temperature

Using the fitted joint distributions for monsoon seasons and months return periods were calculated. According to the table 6, the pair of 494.3 mm and 4482.0 ℃ will occur once in 4 SWMS. Other three pairs occurred in every relevant season.

Table 6: Return Periods for FIMS and SWMS

Maximum value of Rainfall and corresponding temperature value for FIMS			
Rainfall (x/mm)	Temperature $(y/℃)$	CDF $P(X \leq x, Y$	Return Period (Seasons)
551.4	1472.4	0.0015	1
Maximum value of temperature and corresponding rainfall value for FIMS			
172.9	1896.4	0.2733	1
Maximum value of Rainfall and corresponding temperature value for SWMS			
494.3	4482.0	0.7161	4
Maximum value of Rainfall and corresponding temperature value for SWMS			
40.3	4603.8	0.0183	1

4. CONCLUSION

The Gamma distribution for the rainfall and the Weibull distribution for the temperature were identified as the best marginal distributions for the FIMS and for the SWMS the Weibull distribution is identified as the best marginal distributions for both variables. The exponential distribution is the best marginal distribution for rainfall for the selected months except April, and for April the gamma distribution is identified as the best marginal distribution for rainfall. The Weibull distribution is the best marginal distribution for temperature for the above mentioned months. Frank copula is the best copula to represent the joint distribution for the FIMS and SWMS and for the above mentioned. It should be noted that only Frank, Gaussian and t copulas are used for this study because of the negative correlation exists between the two variables.

The distribution function is evaluated using the joint distribution with the best fitted Frank copula. Then the return periods were calculated. the pair of 494.3 mm and 4482.0 ℃ will occur once in 4 SWMS. Other three pairs occurred in every relevant season.

5. REFERENCES

[1] A. AghaKouchak, A. Bardossy, and E. Habib, "*Conditional simulation of remotely sensed rainfall data using a non-Gaussian v transformed copula*", Advances in Water Resources, vol. 33, pp. 624-634, 2010.

[2] A. AghaKouchak, A. Bardossy, and E. Habib, *"Copula- based uncertainty modeling: application to multisensor precipitation estimates"*, Hydrological Processes**,** vol. 24, pp. 2111-2124, 2011.

[3] R. Cong and M. Brady*, "The interdependence between rainfall and temperature :Copula Analysis",* The Scientific World Journal, 2012.

[4] Y. Huang, J. Cai, H. Yin and M. Cai, *"Correlation of Precipitation to temperature variation in the Huanghe River (Yellow River) basin during 1957-2006,"* Journal of Hydrology, vol. 372, pp. 1-8, 2009.