

# Joint Modeling of Wet Condition Characteristics of Makassar City

Wahidah Sanusi<sup>1</sup>, Sahlan Sidjara<sup>2</sup>, Sudarmin<sup>3</sup> and Muhammad Abdy<sup>4</sup>  
{wahidah.sanusi@unm.ac.id<sup>1</sup>, sahlansidjara@unm.ac.id<sup>2</sup>, sudarmin70@gmail.com<sup>3</sup>}

Department of Mathematics, Universitas Negeri Makassar, Parangtambung, Makassar, 90224, Indonesia<sup>1, 2</sup>, Department of Statistics, Universitas Negeri Makassar, Parangtambung, Makassar, 90224, Indonesia<sup>3</sup>

**Abstract.** A multivariate analysis of rainfall variables is needed to understand relationships among characteristics of rainfall variables. The monthly rainfall amount of records from three stations during the period of 1985-2014 in Makassar is used. The Standardized Precipitation Index (SPI) is employed to classify data into the wet category. The study aims are to investigate the most appropriate joint distribution function of wet duration and wet severity based on bivariate Archimedean copulas. The copulas are Gumbel-Hougaard, Frank, Joe, and Clayton. Parameter of copulas is estimated by Kendall's tau correlation coefficient. The study result shows that the Frank copula was identified as the best copula in joint modeling between wet duration and wet severity at Meteorologi Maritim Paotere (MMP) and Biring Romang Panakkukang (BRP) stations, meanwhile the Gumbel-Hougaard copula at Balai Besar Meteorologi, Klimatologi, and Geofisika (BBMKG) station. This result could be useful information for determining return periods of wet condition characteristics.

**Keywords:** Bivariate Distributions, Copula, Wet Duration, Wet Severity

## 1 Introduction

Analysis of rainfall variables as rainfall intensity, duration and frequency (IDF) is an important multivariate tool in flood management and the urban drainage design. The relationships rainfall IDF are usually drawn in a single diagram or IDF curves. In practice, the curves are obtained empirically [1]. Copulas are a tool for understanding relationships among multivariate variables, and also for describing the dependence structure between random variables. The copula method does not require assumptions such as normality, independent or the same marginal distribution types [2], [3].

Karmakar and Simonovic [4] used copula to analyze flood frequency, while Chen et al. [5] applied copula to develop the joint probability distribution of flood magnitudes. For drought frequency analysis, Shiau and Moderras [1] employed Clayton copula to derive the joint distribution function of drought duration and drought severity in Iran. Meanwhile, Chen et al. [6] applied the Archimedean and metaelliptical copulas to model the joint distributions of drought duration, interval time, drought severity, and minimum SPI value.

SPI is one of the indices to evaluate the water supply deficit. Sanusi et al. [7] applied SPI to evaluate drought characteristics in Peninsular Malaysia. In addition to identifying drought events, SPI could also be applied to detect wet conditions [8].

The goal of this study is to identify the best joint distribution function between wet duration and wet severity using bivariate Archimedean copulas.

## 2 Methodology

### 2.1 Definition of the wet characteristics

SPI values are calculated from the monthly rainfall amount of data. For more information about SPI could be found in [9]. Further, SPI values are classified into wet and dry categories. A wet category for the SPI values  $\geq 1$ , while a dry category for the SPI values  $\leq -1$ . This study will only focus on wet conditions based on the one-month SPI series. The wet characteristics include wet duration and wet severity. The wet duration ( $D$ ) is defined as the number of consecutive months of wet condition, while wet severity ( $S$ ) is defined as a cumulative value of SPI within the wet duration [6].

### 2.2 Copulas for bivariate distribution

Sklar in [2] developed the copula method to describe a function that links univariate marginal distribution functions to a multivariate distribution function. Sklar's Theorem in equation (1) states that if  $X$  and  $Y$  are two correlated random variables with  $F_X(x)$  and  $F_Y(y)$  are the marginal cumulative distribution functions, respectively,  $H(x,y)$  is a bivariate distribution function of  $X$  and  $Y$ , then there exists a copula  $C$  such that:

$$H(x, y) = C(F_X(x), F_Y(y)) = C(u, v) \quad (1)$$

where  $F_X(x) = u$ ,  $F_Y(y) = v$ ,  $U$ , and  $V$  follow uniform distribution on the interval (0,1), respectively.

For modeling, firstly, we identify the correlation between wet duration and wet severity using Kendall's tau correlation coefficient. After that, we identify the marginal distributions for a wet duration and wet severity. Four distributions namely exponential, gamma, normal, and Weibull distributions were fitted to investigate the best marginal distribution for both wet duration and wet severity. Akaike's information criteria (AIC) is employed to determine the best fitted marginal distribution [3].

Furthermore, four families of bivariate Archimedean copulas namely Gumbel-Hougaard, Frank, Joe, and Clayton (Table 1) are used and then compared in the study to perform the best-fitted copula models of wet duration and wet severity based on AIC and log-likelihood values of the model. The Kendall's tau correlation coefficient ( $\tau$ ) is used to estimate the parameter of the copula ( $\alpha$ ). Details on the estimation of copula parameter from Kendall's tau correlation coefficient could be found in [2], [3] and references therein. Finally, we evaluate the best joint distribution function for the wet duration and wet severity using the best copula and the best marginal distributions in [10], [1], [5].

**Table 1.** One-parameter families of bivariate Archimedean copulas.

| Family          | Generator ( $\varphi_\alpha(t)$ )   | Bivariate copula ( $C_\alpha(u, v)$ )  |
|-----------------|---|--|
| Gumbel-Hougaard | $(-\ln(t))^\alpha, \alpha \in [1, \infty)$  | $\exp\{-[(-\ln u)^\alpha + (-\ln v)^\alpha]^{\frac{1}{\alpha}}\}$  |
| Frank           | $-\ln\left[\frac{e^{-\alpha t} - 1}{e^{-\alpha} - 1}\right], \alpha \in \mathbb{R} \setminus \{0\}$ | $-\frac{1}{\alpha} \ln\left(1 + \frac{(e^{-\alpha u} - 1)(e^{-\alpha v} - 1)}{(e^{-\alpha} - 1)}\right)$ |
| Joe             | $-\ln[1 - (1 - t)^\alpha], \alpha \in [1, \infty)$  | $1 - [(1 - u)^\alpha + (1 - v)^\alpha - (1 - u)^\alpha(1 - v)^\alpha]^{\frac{1}{\alpha}}$                |
| Clayton         | $\frac{(t)^{-\alpha} - 1}{\alpha}, \alpha \in [-1, \infty) \setminus \{0\}$                         | $(u^{-\alpha} + v^{-\alpha} - 1)^{-\left(\frac{1}{\alpha}\right)}$                                       |

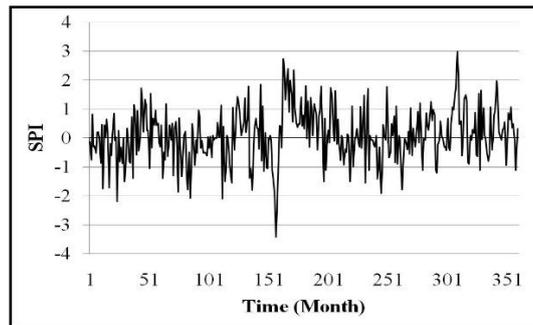
Sources: [2], [3].

### 2.3 Data source

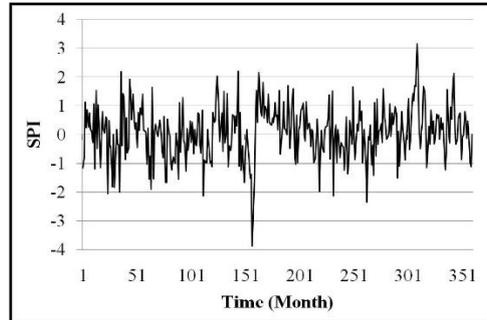
The monthly precipitation amount data (in mm) are obtained from 3 stations in Makassar city, i.e. *Meteorologi Maritim Paotere (MMP)*, *Biring Romang Panakkukang (BRP)* stations, and *Balai Besar Meteorologi, Klimatologi, and Geofisika (BBMKG)* station., for the period January 1985 to December 2014.

### 3 Results and discussions

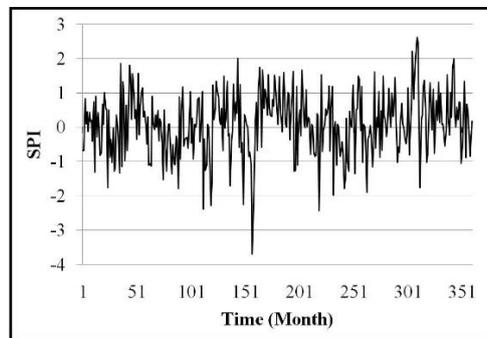
**Figure 1-3** present that the SPI values varying from -3.87 to 3.15. The results also show that although the distribution pattern of SPI values of all station are almost similar, however, the extreme wet events ( $SPI \geq 2$ ) for BRP station is varied than the other stations, that is about 2.04 – 3.15.



**Fig 1.** Temporal distribution of one-month SPI values for MMP station.



**Fig 2.** Temporal distribution of one-month SPI values for BRP station.



**Fig 3.** Temporal distribution of one-month SPI values for BBMKG station.

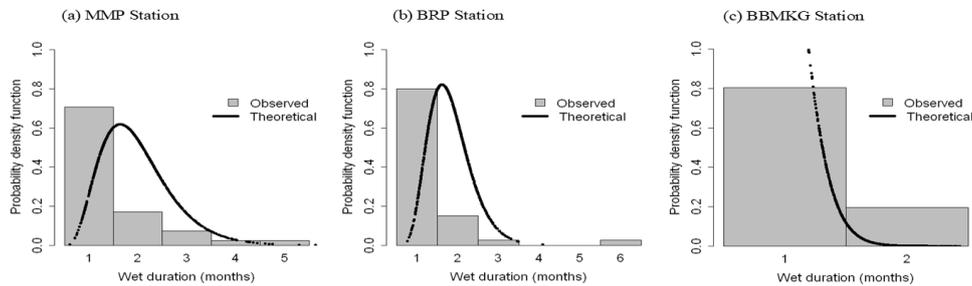
**Table 2.** Performance of four probability distributions based on the AIC values.

| Wet Variable             | Station | Exponential | Gamma | Normal | Weibull |
|--------------------------|---------|-------------|-------|--------|---------|
| Duration<br>( <i>D</i> ) | MMP     | 116.6       | 87.65 | 112.9  | 97.31   |
|                          | BRP     | 104.5       | 70.63 | 107.03 | 86.09   |
|                          | BBMKG   | 98.62       | 31.44 | 44.54  | 44.32   |
| Severity<br>( <i>S</i> ) | MMP     | 148.6       | 134.6 | 173.2  | 142.1   |
|                          | BRP     | 138.3       | 119.3 | 161.4  | 129.7   |
|                          | BBMKG   | 122.2       | 66.59 | 77.69  | 75.55   |

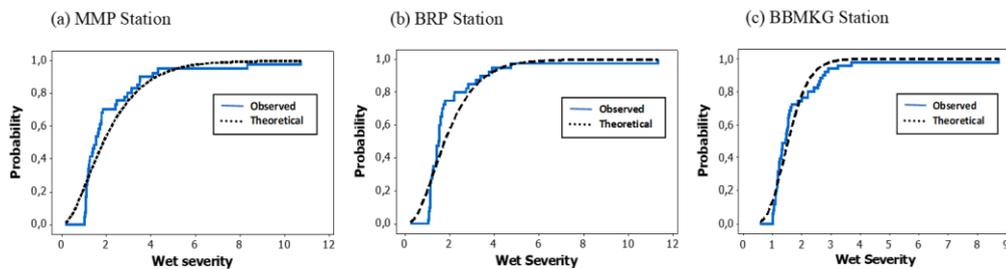
Further, the SPI values are used to determine wet duration and wet severity. In Table 2, four probability distributions were used to fit the suitable marginal distribution model for both wet duration and wet severity. Table 2 shows that for all stations, gamma distribution was the preferred marginal probability distribution with a lower AIC value as compared to the other probability distributions for both wet duration and wet severity. Table 3 presents the estimation of gamma parameters for the wet duration and wet severity. Meanwhile, Figure 4 and Figure 5 display comparing between observed and theoretical gamma distribution for the wet duration and wet severity, respectively, which also illustrates good agreement between fitted gamma model and observed data for all stations.

**Table 3.** The estimated parameters of gamma distributions.

| Wet Variable | Parameter | Rainfall station |       |       |
|--------------|-----------|------------------|-------|-------|
|              |           | MMP              | BRP   | BBMKG |
| $D$          | Alpha     | 2.586            | 2.225 | 8.873 |
|              | Beta      | 0.575            | 0.596 | 0.135 |
| $S$          | Alpha     | 1.298            | 1.332 | 7.015 |
|              | Beta      | 1.694            | 1.518 | 0.227 |



**Fig 4.** Frequency curves of gamma distribution for the wet duration.



**Fig 5.** Frequency curves of gamma distribution for wet severity

Table 4 presents the values of Kendall's tau correlation coefficient for wet characteristics. The result reveals that there is a positive correlation between wet duration and wet severity for all stations. Therefore, the copula method could be applied to evaluate the joint probability distribution of wet characteristics, and then the most suitable copula was investigated. Further, the copula parameter is estimated by using Kendall's- tau correlation value as in Table 5. The most suitable copula model was selected according to the log-likelihood and AIC values and the study results find that the Frank copula gave the highest log-likelihood value and the smallest AIC value for both the MMP and BRP stations as in Table 6. Meanwhile, the Gumbel- Hougaard copula was identified as the best copula at BBMKG station as displayed in Table 6.

**Table 4.** Correlation test between wet duration and wet severity.

| Station | Kendall-tau ( $\tau$ ) | Z<br>value | p-value               |
|---------|------------------------|------------|-----------------------|
| MMP     | 0.691                  | 5.495      | $3.88 \times 10^{-8}$ |
| BRP     | 0.589                  | 4.525      | $6.03 \times 10^{-6}$ |
| BBMKG   | 0.569                  | 4.343      | $1.40 \times 10^{-5}$ |

**Table 5.** The estimated parameters of bivariate Archimedean copulas.

| Station | Gumbel-H | Frank   | Joe   | Clayton |
|---------|----------|---------|-------|---------|
| MMP     | 3.236    | 11.0114 | 5.272 | 4.472   |
| BRP     | 2.433    | 7.643   | 3.696 | 2.866   |
| BBMKG   | 2.320    | 7.156   | 3.476 | 2.640   |

**Table 6.** The log-likelihood and AIC values of bivariate Archimedean copulas.

| Station | Copula                  | Log-likelihood  | AIC            |
|---------|-------------------------|-----------------|----------------|
| MMP     | Gumbel-Hougaard         | -41.366         | 92.732         |
|         | <b>Frank</b>            | <b>-38.481</b>  | <b>86.962</b>  |
|         | Joe                     | -47.127         | 104.254        |
|         | Clayton                 | -79.621         | 169.242        |
| BRP     | Gumbel- Hougaard        | -38.149         | 86.298         |
|         | <b>Frank</b>            | <b>-34.383</b>  | <b>78.766</b>  |
|         | Joe                     | -41.716         | 93.432         |
|         | Clayton                 | -48.515         | 107.03         |
| BBMKG   | <b>Gumbel- Hougaard</b> | <b>-160.873</b> | <b>331.746</b> |
|         | Frank                   | -192.512        | 395.024        |
|         | Joe                     | -214.238        | 438.476        |
|         | Clayton                 | -386.479        | 782.958        |

The copula-based the joint cumulative distribution function (JCDF) of wet duration and wet severity for MMP station, BRP station, and BBMKG station are given in equation (2), equation (3) and equation (4), respectively.

$$C(F_D, F_S) = \frac{-1}{11.0114} \ln \left( 1 + \frac{(\exp(-11.0114F_D)-1)(\exp(-11.0114F_S)-1)}{(\exp(-11.0114)-1)} \right), \quad (2)$$

where

$$F_D(d) = \frac{1}{(0.575)^{(2.586)}\Gamma(2.586)} \int_0^d x^{1.586} \exp\left(-\frac{x}{0.575}\right) dx,$$

$$F_S(s) = \frac{1}{(1.694)^{(1.298)}\Gamma(1.298)} \int_0^s y^{0.298} \exp\left(-\frac{y}{1.694}\right) dy.$$

$$C(F_D, F_S) = \frac{-1}{7.643} \ln \left( 1 + \frac{(\exp(-7.643F_D)-1)(\exp(-7.643F_S)-1)}{(\exp(-7.643)-1)} \right), \quad (3)$$

where

$$F_D(d) = \frac{1}{(0.586)^{(2.225)}\Gamma(2.225)} \int_0^d x^{1.225} \exp\left(-\frac{x}{0.596}\right) dx,$$

$$F_S(s) = \frac{1}{(1.518)^{(1.332)}\Gamma(1.332)} \int_0^s y^{0.332} \exp\left(-\frac{y}{1.518}\right) dy.$$

$$C(F_D, F_S) = \exp\left(-\left[(-\ln(F_D))^{2.32} + (-\ln(F_S))^{2.32}\right]^{0.431}\right) \quad (4)$$

where

$$F_D(d) = \frac{1}{(0.135)^{(8.873)}\Gamma(8.873)} \int_0^d x^{7.873} \exp\left(-\frac{x}{0.135}\right) dx,$$

$$F_S(s) = \frac{1}{(0.227)^{(7.015)}\Gamma(7.015)} \int_0^s y^{6.015} \exp\left(-\frac{y}{0.227}\right) dy.$$

### 3 Results and discussions

The one-month time scale of SPI values was determined based on monthly readings of rainfall data for three selected stations from 1985 to 2014. The results show that the occurrence of wet and dry events for all stations is nearly coincident, but Makassar city has experienced the extreme wet in the last 30 years.

Based on the result of the study, it is found that the duration and severity of wet events have the same probability distribution that is the two parameters gamma distribution. The results

also show that the most appropriate Archimedean copula model for MMP and BRP stations are Frank copula, while the Gumbal-Hougaard copula is the most suitable model for BBMKG station. Future, the best joint distribution will be employed to identify the return period of wet events. The return period information of the wet events is useful for assessing the risk of drought or flood.

**Acknowledgments.** The authors would like to thank the staff of the Meteorological, Climatological, and Geophysical Agency (BMKG) of Makassar for supplying the monthly rainfall data. The first three authors are also grateful to the Ministry of Research, Technology and Higher Education of Indonesia for providing funding with contract number: 36/UN36.9/PL/2018 and 130/UN36.9/PL/2019.

## References

- [1] Shiau, J.T. and Moderras, R.: Copula-based drought severity-duration-frequency analysis in Iran. *Meteorol. App.* Vol. 16, issue number 4, pp. 481-489 (2009)
- [2] Kpanzou, T.A.: *Copulas in Statistics*, African Institute for Mathematical Sciences (AIMS). University of Stellenbosch (2007)
- [3] Zhang, L and Singh, V.P.: Bivariate rainfall frequency distributions using Archimedean copulas. *Journal of Hidrology*. Vol. 332, pp. 93 – 109 (2007)
- [4] Karmakar, S. and Simonovic, S. P.: Flood frequency analysis using copula with mixed marginal distributions. *Water Resources Research report No.059*. The University of Western Ontario, Department of Civil and Environmental Engineering, London, Ontario. 142 pages. ISBN: (print) 978-0-7714-2658-2; (online) 978-0-7714-2659-9 (2007)
- [5] Chen, L., Singh, V.P., Shenglian, G., Hao, Z. & Li, T.: Flood Coincidence Risk Analysis Using Multivariate Copula Functions. *J. Hydrol. Eng.* Vol. 17, pp. 742-755 (2012)
- [6] Chen, L., Singh, V.P., Guo, S., Mishra, A. K. & Gou, J.: Drought analysis using copulas. *J. Hydrol. Eng.* Vol. 18, pp. 797-808 (2013)
- [7] Sanusi, W., Jemain, A. A. and Zin, W.Z.W.: The drought characteristics using the first-order homogeneous Markov chain of monthly rainfall data in Peninsular Malaysia. *Water Resources Management*. Vol. 29, issue number 5, pp. 1523-1539 (2015)
- [8] Sanusi, W. and Ibrahim, K.: Application of Loglinear Models in Estimating wet category in Monthly Rainfall. *Sains Malaysiana*. Vol. 41, issue number 11, pp. 1345-1353 (2012)
- [9] Mishra, A.K., Singh, V.P., and Desai, V.R.: Drought forecasting using stochastic models. *Stoch Environ Res Risk Assess.* Vol. 19, pp. 326 – 339 (2005)
- [10] Ekanayake, E.M.R.S.B. and Perera, K.: Analysis of drought severity and duration using copulas in Anuradhapura, Sri Lanka. *British Journal of Environment & Climate Change*. Vol. 4, issue 3, pp. 312-327 (2014)