CHAPTER 9

Copula Application in Rainfall Analysis

Mohd Aftar Abu Bakar, Noratiqah Mohd Ariff, Abdul Aziz Jemain and Adina Najwa Kamarudin

9.1 INTRODUCTION

Hydrological data such as rainfall are complex in nature. Therefore, hydrological events mostly appear to be multivariate. In addition, previous studies have also found that single variable frequency analysis would only offer inadequate evaluations and representations of these complex events (Yue et al., 2001). Hence, several attempts have been explored to derive joint distributions of rainfall characteristics. Common rainfall characteristics which are used as random variables in multivariate distributions of precipitations are rainfall duration, depth, and intensity.

Previous studies have assumed some premises to obtain multivariate distributions based on conventional statistical methods. For example, the random variables in the distribution are assumed to be independent of each other, which is said to be inappropriate and unrealistic (Cordova & Rodriguez-Iturbe, 1985). Thus, many rainfall models incorporate the relationship between random variables but with assumptions on the random variables' marginal distributions. Yue (2000) assumed that the marginals are normally distributed. Some studies assumed similar probability distributions for the marginals. This led to the use of bivariate exponential (Favre et al., 2002; Goel et al., 2000), bivariate Gamma (Yue et al., 2001), bivariate lognormal (Yue, 2002) and bivariate extreme value distribution (Shiau, 2003) to represent the joint probability distributions of random variables. However, these marginals are not necessarily normal or similar since rainfall is a complicated phenomenon.

Copulas reduce the complexity of obtaining multivariate distributions by separating the fitting of marginal distributions and the analysis of the structure of dependence for multivariate distribution (Nelson, 2006). Hence, copulas provide a manageable approach that allows more selections and options for the probability distributions of the marginals and the dependence structures between random variables (Kao & Govindaraju, 2008). Various types of copulas have been used in hydrology to build rainfall models (Favre et al., 2004; Renard & Lang, 2007; Salvadori & De Michele, 2004; Zhang & Singh, 2006).

9.2 COPULA METHOD

A copula is regularly described as a function that connects a multivariate distribution function with the marginal distribution of each of its random variables or a function made up of cumulative uniform distributions in a closed interval of [0,1] (Nelson, 2006). Copulas are designed based on uniform marginals (Dzupire et al., 2020) and have emerged as viable and efficient functions to model the dependency of multivariable data (Tahroudi et al., 2020). For simplicity purposes, the copula method is explained in terms of a bivariate distribution. However, the same concept can be easily extended to trivariate and any other multivariate distributions.

A bivariate distribution function, $H_{X,Y}$, of two continuous random variables X and Y have marginal distributions $U = F_x$ and $V = F_y$ respectively. Hence, U and V are uniform distributions in

$[0,1]$, i.e. $U \sim U[0,1]$ and, $V \sim U[0,1]$ based on the probability integral transform. Thus, realizations of U and V can be written in terms of realizations of X and Y as $u = F_x(x)$ and $\upsilon = F_y(y)$ respectively. Based on a theorem by Sklar (1959), a distinct copula, $C_{U,V}$, exists where

$$\begin{aligned} C_{U,V}(u,v) &= P(U \le u, V \le v) \\ &= P\left(X \le F_X^{-1}(u), Y \le F_Y^{-1}(v)\right) \\ &= H_{X,Y}\left(F_X^{-1}(u), F_Y^{-1}(v)\right) = H_{X,Y}(x,y) \end{aligned} \tag{9.1}$$

for all x and y in $\mathbb{R}$. As can be seen in the equation for $C_{U,V}$, copula maps random observations (x,y) from $\mathbb{R}^2$ to a bounded domain $[0,1]^2$.

A one-parameter copula where the parameter relates to the dependence measure is usually used to represent a bivariate distribution. It can be obtained using the following steps:

(1) Find the marginals U and V of the random variables X and Y, respectively.
(2) Calculate the measure of the relationship between X and Y.
(3) Calculate the copula parameter, θ, through its relationship with the correlation. This association between copula parameters and correlation differs for different copulas.
(4) Build the copula function by taking the estimated value θ and the marginals U and V into the respective copula equation.