Available online at www.sciencedirect.com

ScienceDirect

Journal of Hydro-environment Research 11 (2016) 113–122

Journal of Hydro-environment Research

www.elsevier.com/locate/jher

Research paper

Application of copula functions to construct confidence intervals of bivariate drought frequency curve

Jiyoung Yoo [a], Dongkyun Kim [b], Hungsoo Kim [c], Tae-Woong Kim [d,*]

[a] Department of Civil Engineering, Chonbuk National University, Jeonju 561-756, Republic of Korea

[b] School of Urban and Civil Engineering, Hongik University, Seoul 121-791, Republic of Korea

[c] Department of Civil Engineering, Inha University, Incheon 402-751, Republic of Korea

[d] Department of Civil and Environmental Engineering, Hanyang University, Ansan 426-791, Republic of Korea

Received 1 November 2013; revised 2 October 2014; accepted 6 October 2014

Abstract

Drought frequency analysis provides a comprehensive view point by simultaneously considering duration and severity when assessing drought risk. This study presents the applications of copula functions to construct confidence intervals of bivariate drought frequency curve. The bivariate drought frequency curves were constructed using observed monthly rainfall after extracting drought properties such as severity and duration. In order to construct the confidence intervals of bivariate drought frequency curve, 100 realizations of 100-year long monthly rainfall were generated using the Copula-GARCH rainfall generation model, after investigating model performance of various copula functions. The quantiles of drought severity for different drought durations were then calculated in accordance with bivariate frequency analysis. The application results achieved in this study illustrates that the proposed method would be available for quantifying the uncertainty of bivariate drought frequency curves in practice.

Keywords: Bivariate; Confidence interval; Copula; Drought; Frequency; Severity

1. Introduction

Droughts gradually occur over a long period of time without noticeable immediate damage. However, the negative impact of drought on the human and natural environment is substantial, because spatial coverage and temporal duration of drought are significantly greater than those of other natural disasters such as floods and hurricanes. In addition, the adverse impacts of drought are exacerbated by intensifying hydrologic cycle associated with the process of global climatic change (World Meteorological Organization (WMO), 2006). Unprecedented patterns of drought and their corresponding socio-economic impacts are constantly observed across the world (Cook et al., 2004; Leblanc et al., 2009; Fang et al., 2010; Parry et al., 2012).

One of the most widely used methods to make preparedness plans for drought is drought frequency analysis, in which the probability distributions of drought properties, such as duration and severity, are determined and their recurrence intervals are estimated based on extreme value theory. It is also important to note that the results of drought frequency analysis based on a single property alone have been inconsistent (Fernández and Salas, 1999; Chung and Salas, 2000; Cancelliere and Salas, 2004). In other words, the recurrence interval of the drought duration and the drought severity can be different, even if both values are obtained from the same drought event. To resolve this issue, several studies suggested to consider multiple drought properties simultaneously. In particular, joint probability density functions that combine the two primary properties of drought such as duration and severity have been studied extensively (Kim et al., 2003; Salas et al., 2005; Mirakbari et al., 2010; Yoo et al., 2012b). Several studies have also employed copula functions to develop bivariate probability functions based on joined-variable marginal probability distributions (Kao and Govindaraju, 2008; Chebana and Ouarda, 2011; Lee and Salas, 2011; Lee et al., 2013).

* Corresponding author. Tel.: +82 31 400 5184.
E-mail address: twkim72@hanyang.ac.kr (T.-W. Kim).

http://dx.doi.org/10.1016/j.jher.2014.10.002
1570-6443/

While bivariate drought frequency analysis includes both drought duration and severity in a joint probability distribution to develop a comprehensive prediction, values located in marginal probability distributions are often unreliable. Drought with a given recurrence interval and severity can have different duration estimates, and drought with a given recurrence interval and duration can have different severity estimates. This is mainly due to the lack of sufficient recorded data, which is limited to less than 100 years in most studies. Droughts recorded over such limited time periods, often do not have duration and severity values corresponding to the marginal position of the bivariate drought frequency curve (or simply when the drought of interest has a high recurrence interval). To ensure the reliability of bivariate drought frequency analysis, a sufficient amount of data must be obtained (Scott, 1992; Kim et al., 2006). To overcome this issue, this study suggested confidence intervals within which the values from the bivariate frequency curve could be considered reliable. The confidence intervals were derived from time series of rainfall generated with the copula-GARCH (Generalized Autoregressive Conditional Heteroscedasticity) rainfall simulation model. The greatest strength of the copula-GARCH model is that it integrates the conditional dependence of variability between variables in the simulation. Heteroscedasticity in hydrologic time series including rainfall is widely reported across the world (Wang et al., 2005; Modarres and Ouarda, 2012; Yusof and Kane, 2012), thus the copula-GARCH model is more appropriate to simulate a rainfall time series than other conventional approaches which are usually based on the assumption of homoscedasticity.

The primary purpose of this study is to provide the confidence intervals for bivariate frequency curve considering drought duration and severity. This study suggests the confidence intervals based on the simulated rainfall time series using the copula-GARCH modeling approach. This study is described as follows: Section 2 presents the theories, methodologies, and application results of the copula-GARCH rainfall generation model; Section 3 explains the methodology to construct bivariate drought frequency curves using copula functions; Section 4 presents methodology to estimate confidence intervals of bivariate drought frequency curve and discusses the corresponding results; and Section 5 draws conclusions from the overall results presented in this study.

2. Generating monthly rainfall

2.1. Copula-GARCH model

This chapter presents theories, methodologies, and application results of the copula-GARCH rainfall generation model. The GARCH model was developed as an alternative to models based on the assumption of linearity between variables at different time steps, which cannot account for the conditional dependence of the variance or heteroscedasticity. It has been widely applied, particularly in the field of finance and economics, because of its strength in modeling variables of which the variation is significant (Duan, 1996; Tse and Tsui, 2002; Floros et al., 2007; Watanabe, 2012). Recently, the GARCH model has been applied in simulating hydrologic time series (Wang et al., 2005; Modarres and Ouarda, 2012; Yusof and Kane, 2012).

In this study, Engle's ARCH test was performed to detect heteroscedasticity in rainfall time series collected at twelve sites shown in Fig. 1. The null hypothesis of Engle's ARCH test is that a series of residuals exhibit no conditional heteroscedasticity (Engle, 1988). The null hypothesis was rejected at the 5% significance level for 7 of 12 sites. In addition, Fig. 2 shows the autocorrelogram for the square of the residual (long-term mean subtracted from monthly rainfall) for site #156, as representative test case. The lag-1 autocorrelation of the square of the residual was 0.18. This means the monthly rainfall deviation of any given month from its long term mean was similar to that of the previous month. In this way the GARCH model provided more realistic model results compared to conventional approaches.

The GARCH (p, q) model is defined by Eq. (1).

$$\begin{cases} X_t = \sigma_t \varepsilon_t \\ \sigma_t^2 = \alpha_0 + \sum_{k=1}^{p} \alpha_k X_{t-k}^2 + \sum_{k=1}^{q} \beta_k \sigma_{t-k}^2 \end{cases} \tag{1}$$

where subscript t represent the time step; X_t represents residual series of the variable being modeled subtracted from its mean, σ_t is the standard deviation of X_t, ε_t is the random variable drawn from a distribution representing X_t, p and q represents GARCH model dimensions; α_0, α_k, and β_k represents GARCH model parameters to be estimated.

After investigating various values for p and q, we concluded that the simple GARCH (1, 1) model provided good representation for a variety of volatility processes. This is in accordance with Bollerslev et al. (1992). The parameters of the GARCH (1, 1) model were estimated using the maximum likelihood method. The random variable ε_t in Eq. (1) was drawn from a skewed student-t distribution, as Hansen (1994) suggested. The skewed student-t distribution is known to well fit the heavy-tailed data, and it can also handle slight skewness; such a tendency was observed in rainfall data used in this study. A copula function links the marginal distributions of correlated variables and develop a single joint probability distribution function, it is often incorporated with the GARCH model.

Two types of copula function were considered in this study; the rotated Gumbel copula and the Student-t copula, as given by Eqs. (2) and (3), respectively.

$$C(u, v; \theta) = exp\left(-\left[(-log(u))^{\theta} + (-log(v))^{\theta}\right]^{1/\theta}\right) \tag{2}$$

$$C(u, v; \theta) = \frac{1}{\sqrt{1-\theta^2}} exp\left(\frac{2\theta xy - x^2 - y^2}{2(1-\theta^2)} + \frac{x^2 + y^2}{2}\right) \tag{3}$$