



A development of trivariate drought frequency analysis approach using copula function

Kim, Jin-Young^a · So, Byung-Jin^a · Kim, Tae-Woong^b · Kwon, Hyun-Han^{a*}

^aDepartment of Civil Engineering, Chonbuk National University

^bDepartment of Civil and Environmental Engineering, Hanyang University

Paper number: 16-069

Received: 25 July 2016; Revised: 25 August 2016 / 26 August 2016; Accepted: 26 August 2016

Abstract

This study developed a trivariate Copula function based drought frequency analysis model to better evaluate the recent 2014~2015 drought event. The bivariate frequency analysis has been routinely used for the drought variables of interest (e.g. drought duration and severity). However, the recent drought patterns showed that the intensity can be regarded as an important factor which is being characterized by short duration and severe intensity. Thus, we used the trivariate Copula function approach to incorporate the trivariate drought characteristics into the drought frequency analysis. It was found that the return periods based on the trivariate frequency analysis are, in general, higher than the existing bivariate frequency analysis. In addition, this study concludes that the increase in drought frequency claimed by the Gumbel copula function has been overestimated compared to the Student t Copula function. In other words, the selection of copula functions is rather sensitive to the estimation of trivariate drought return periods at a given duration, magnitude and intensity.

Keywords: Trivariate Copula Function, Return Period, Drought, Trivariate Drought Frequency Analysis

Copula 함수를 활용한 삼변량 가뭄빈도해석 기법 개발

김진영^a · 소병진^a · 김태웅^b · 권현한^{a*}

^a전북대학교 토목공학과, ^b한양대학교 건설환경플랜트공학과

요 지

본 연구에서는 최근 발생한 2014~2015 가뭄 사상을 보다 정확하게 분석하기 위해 삼변량 Copula 함수를 도입하여 연구를 진행하였다. 기존 연구에서는 일반적으로 가뭄 분석시 이변량(가뭄 지속시간, 심도)을 활용한 연구가 다수 진행되었다. 그러나 최근 강우자료의 패턴을 살펴보면 두 변량 이외의 가뭄 강도가 중요한 인자로 평가되어 이를 함께 고려한 삼변량 Copula 분석을 수행하였으며, 기상청 관측소 중 서울 관측소를 대상으로 연구를 진행하였다. 기본적으로, 이변량 빈도해석 결과에 비해 삼변량 해석 결과는 동일한 가뭄 사상에 대해서 다소 증가된 재현기간을 나타내는 것으로 파악됐다. 이와 더불어, Gumbel Copula 함수의 경우 Student t Copula 함수보다 가뭄 위험도 평가 시 다소 과대 추정하는 것으로 확인되었다. 즉, 삼변량 빈도해석 시 고려되는 Copula 함수의 선택이 가뭄의 재현기간을 추정하는데 있어 매우 민감한 사항으로 평가되었다.

핵심용어: Trivariate Copula function, 빈도해석, 가뭄, 삼변량 가뭄빈도해석

1. 서 론

2014~2015년 우리나라 강수량이 평년에 비해 절반 수준에 미치지 못해 극심한 가뭄을 일으켰으며, 이는 댐 용량 부족, 지하수 고갈 등 다양한 피해를 발생시켰다. 특히 소양강댐의

*Corresponding Author. Tel: +82-63-270-2426
E-mail: hkwon@jbnu.ac.kr (H.-H. Kwon)

경우 1978년에 발생한 최저 수위 151.93 EL.m 이후 두 번째로 낮은 수위인 152.25 EL.m를 기록하였다. 우리나라에서 가뭄은 대략적으로 2~3년 주기를 가지고 발생하고 있으며, 특히, 2015년에 겪었던 가뭄은 물 용수공급 측면에서 막대한 영향을 미친 것으로 평가되었다. 지구온난화로 인한 기후변화의 영향으로 강수량의 증가가 일반적으로 전망되지만, 상대적으로 증가된 강우변동성으로 인해 가뭄 발생 빈도 및 강도도 동시에 증가될 것으로 전망되고 있다(Dai et al., 2004; Sheffield and Wood, 2008).

국내외 연구에서는 가뭄 연구를 위해 가뭄 모니터링, 예측 인자 발견 등 다양한 연구가 수행되어 지고 있다. 이러한 가뭄 특성을 분석하기 위해 가뭄에 영향을 미치는 다양한 인자들을 조사하고, 빈도해석을 통해 현재 가뭄을 평가하는 연구가 다수 진행되어 왔다(Fernández and Salas, 1999; Bonaccorso et al., 2003; Kim et al., 2003; Kim et al., 2006; Cancliere and Salas, 2010; Kwon and Lall, 2016). 그러나 가뭄 현상과 같은 경우 다양한 현상이 복합적으로 발생하므로 단변량 빈도해석은 가뭄의 빈도 분석 시 과소 또는 과대 추정될 수 있다(Shiau et al., 2007). 이러한 점에서 다수의 가뭄연구에서는 다변량(multivariate) 가뭄 해석 절차의 필요성이 강조되고 있다. 즉, 다양한 가뭄변량들의 결합분포(joint distribution)를 활용하는 것이 보다 신뢰성 있는 가뭄해석에 필수 사항으로 인식되고 있다. 일반적으로, 가뭄 분석을 위해서 다양한 가뭄의 특성 인자를 정의하고 있으며, 가뭄 지속시간(drought duration)과 가뭄 심도(drought severity), 가뭄 강도(drought intensity), 가뭄 발생빈도(drought frequency)등으로 구분되어진다.

국외 연구동향을 살펴보면 앞서 정의된 가뭄 인자 중 2개의 변량을 선택하여 이변량 가뭄해석을 수행한 연구가 다수 진행되어왔다. Shiau and Modarres (2009)는 Copula 함수를 기반으로 가뭄 지속시간과 심도 자료를 활용하여 SDF (severity-duration-frequency) 곡선을 개발하였으며, 미계측 구역에서도 가뭄에 대한 판단을 수행할 수 있는 모형으로 확장 개발하였다. Shiau and Shen (2001)은 가뭄 심도에 대한 위험도 분석(risk analysis)을 위해 조건부 Copula 모형을 활용하였으며, 특정 지속시간에 대한 심도를 산정함으로써 대상 구역의 가뭄 위험도를 산정한바 있다. 또한 Shiau (2006)는 가뭄 지속시간과 심도를 대상으로 다양한 2차원 Copula를 적용하였으며, Copula 함수별 결합분포 및 조건분포(condition distribution)를 추정하고 이를 이용하여 빈도해석을 수행하였다. 국내에서는 Yoo et al. (2013)은 빈도해석 시 샘플링오차를 개선하기 위하여 추계학적 모형을 이용하였으며, 모의된 수문변량을 대상으로 이변량 가뭄빈도해석을 수행하였다. 이와 더불어

해석 결과에 대한 불확실성을 정량적으로 제시하였다. 국내외 연구 동향을 살펴본 결과 앞서 언급하였듯이 대다수의 연구에서 가뭄 지속시간과 가뭄 심도를 대상 자료로 이변량 빈도해석 수행되고 있다(Chun et al., 2015; Lee and Son, 2016; Yu et al., 2016).

최근 기후변화의 영향으로 인해서 가뭄은 짧은 지속시간과 강한 심도의 특징을 가지고 있다. 이러한 점을 고려하여 가뭄 강도 역시 가뭄 특성을 분석하는데 있어 중요한 인자로 평가되고 있다(Laird et al., 1998; Kumar et al., 2009). 즉, 이변량 분석을 통해 가뭄을 판단하는 연구에서 나아가 다변량(multivariate) 해석기법을 활용한 연구가 대두되고 있다(Zhang and Singh, 2007). Grimaldi and Serinaldi (2006)은 해석변량으로서 한계깊이(critical depth)-총깊이(total Depth)-최대강수(peak)를 활용하여 삼변량 결합확률분포를 산정하고 조건부 확률을 통해 강우모의 발생을 수행하였다. Favre et al. (2004)은 다변량 모델링을 위해 삼변량 Copula를 활용하였으며, 강우자료와 유량자료를 대상으로 분석을 수행하였으며 최종결과물로서 각 Copula 함수별로 위험도를 제시하였다. 국외에서는 이러한 수문변량의 신뢰성 있는 모의를 위해 다변량 해석이 다수 진행되고 있는 반면, 국내에서는 이변량 이상의 수문해석모형의 개발 및 적용이 매우 제한적으로 이루어지고 있다.

이러한 점에서 본 연구에서는 가뭄변량으로서 지속시간, 심도 그리고 강도 자료를 대상으로 삼변량 Copula 함수를 활용한 가뭄빈도해석 모형을 개발하였다. 모형의 적용은 최근 2014~2015년의 가뭄을 대상으로 전반적으로 수행되었으며, 기상청 관측소 중 서울 관측소를 대상으로 연구를 진행하였다.

본 논문의 구성은 다음과 같다. 1장에서는 연구 내용의 전반적인 내용에 대해 서술 하였으며, 2장에서는 가뭄 정의 및 삼변량 Copula 모형에 대해 서술하였다. 3장에서는 이변량 및 삼변량 Copula 빈도해석 모형을 통해 현재 가뭄 빈도 비교 검토를 수행하였으며, 4장에서는 결론에 대해 서술하였다.

2. 본 론

본 연구에서는 Copula 함수를 활용하여 빈도해석을 수행하였으며, 가뭄분석을 위해 사용된 자료는 서울 관측소(108) 자료를 사용하였다. 가뭄해석 변량으로서 가뭄의 기간, 심도 및 강도를 선택하여 가뭄 빈도를 해석하였으며, 가뭄빈도 해석 결과는 삼변량 Copula 함수를 통해 도출하였다. 본 절에서는 본 연구에서 활용한 삼변량 Copula 모형 개념 및 연구 수행

방법에 대해 수록하였다.

2.1 가뭄 정의

가뭄은 일반적으로 기상학적, 수문학적, 농업적 및 사회경제학적 가뭄 등 크게 4가지로 구분되어 진다. 최근 국내외에서 가뭄 분석 연구는 가뭄 지속시간과 가뭄 심도를 활용한 이변량 해석이 주로 수행되어 지고 있다. 즉, 가뭄의 특성상 복합적인 인자간의 상호관계를 파악함으로써 보다 정확한 현상을 파악할 수 있는 것이다. 본 연구에서는 가뭄 빈도해석시 가뭄 특성 인자는 Yevjevich (1967)의 가뭄 정의를 활용하였으며, 6개월 누적 월 강우자료를 활용하였다. 이는 정상년 기준의 강수량을 기준으로 강수의 Anomaly를 구분하고자 가뭄절단수준을 결정하였다(Chun et al., 2015). 가뭄 지속시간은 가뭄절단수준(x_0)이하로 떨어진 시점에서 해갈되는 시점으로 정의하며, 이때 발생한 강우 누적부족량을 가뭄 심도로 정의하고 있다. 또한 가뭄 강도의 경우 강우 누적부족량을 동일 기간의 지속시간으로 나눔으로써 도출될 수 있는 가뭄 인자이다.

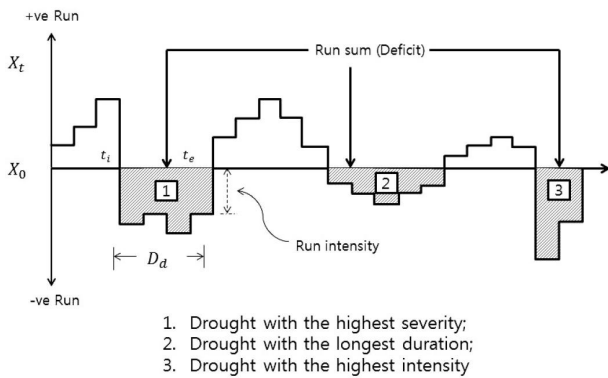


Fig. 1. The identified drought duration, severity and intensity by run theory (Yoo et al., 2013)

2.2 Multivariate return period using Trivariate Copula function

Copula 모형은 널리 알려져 있듯이 Sklar (1959)에 의해 처음 소개되었으며, Joe (1997)과 Ne lsen (2006)에 의해 다양한 기본 이론 및 방법론이 정립되었다. Copula 모형은 현재까지도 통계 및 경제분야에서 빈번하게 사용되어 지고 있으며 (Radice et al., 2015; Salvatierra and Patton, 2015), 최근 수문학 분야에서도 다변량 분석시 활발하게 활용되고 있다. 특히 수문분야에서 상호 변량간의 의존성이 뚜렷한 가뭄현상 및 이에 대한 빈도를 평가하기 위해 활용되고 있다. 이와 더불어 Bayesian Network 등의 조건부 확률계산이 주로 이루어지는 다양한 평가체계에서 사용되어 지고 있다(Kim et al., 2015).

일반적으로 Copula 함수는 누적확률분포를 활용하여 변수들간의 종속성 구조를 평가하는데 이용된다. 수문학적 분석시 사용되는 변량의 경우 꼬리(tail)가 두꺼운 극치값(극대 또는 극소) 값을 사용하므로 이러한 변량의 꼬리를 다루는데 있어 Copula 함수를 사용하여 그 종속성 구조를 파악하는 것이 타당하다고 알려져 있다. Copula 함수의 경우 확률변수 X_1, X_2, \dots, X_n 이 구간 $[0, 1]$ 에서의 균일분포를 따른다고 가정하였을 때 Copula 함수는 다음과 같은 Eq. (1)으로 나타낼 수 있다.

$$C(x_1, x_2, \dots, x_n) = P(X_1 \leq x_1, X_2 \leq x_2, \dots, X_n \leq x_n) \quad (1)$$

이때, Sklar의 이론을 기반으로 이러한 확률변수 X_1, X_2, \dots, X_n 의 주변확률분포함수(marginal distribution) $F_1(x_1), F_2(x_2), \dots, F_n(x_n)$ 이 존재할 때, Copula 함수를 활용하여 결합분포함수를 표현하면 다음 Eq. (2)와 같다.

$$H(x_1, x_2, \dots, x_n) = C(F_1(x_1), F_2(x_2), \dots, F_n(x_n)) \quad (2)$$

Table 1. The Copula functions used in this study

Name of copula	Joint distribution function
Gaussian	$c^G(u_1, u_2, \dots, u_n; \Sigma) = \frac{\partial}{\partial u} C^G(u_1, u_2, \dots, u_n; \Sigma) = \frac{1}{ \Sigma ^{1/2}} \exp(-\frac{1}{2} y^T (\Sigma^{-1} - I) y)$
Clayton	$[\max(u^{-\theta} + v^{-\theta} - 1; 0)]^{-1/\theta}$
Frank	$-\frac{1}{\theta} \ln[1 + \frac{(\exp^{-\theta u} - 1)(\exp^{-\theta v} - 1)}{\exp^{-\theta} - 1}]$
Gumbel	$\exp[-((-\ln(u))^\theta + (-\ln(v))^\theta)^{1/\theta}]$
Student t	$c(u; v; \Sigma) = \frac{\Gamma((v+d)/2)[\Gamma(v/2)]^{d-1}}{[\Gamma((v+1)/2)]^{1/2}} [\prod_{i=1}^d (1 + y_i^2/v)^{(v+1)/2}] (1 + y^T \Sigma^{-1} y/v)^{-(v+d)/2}$

즉, 다양한 확률변수와 그에 부합하는 주변확률분포를 기반으로 Copula 함수를 통해 결합분포확률을 결정할 수 있으며, 각 변량이 강한 상관성을 지니고 있으면 보다 현실적인 모형을 구축할 수 있게 된다. 일반적으로 Copula 함수는 매개변수적과 비매개변수적으로 구분되어진다. Nelssen (2006)은 Copula 함수의 계산시 계산과정이 보다 간편하고, 다양한 확률분포형에 적용이 가능한 5개(Gaussian, Clayton, Frank, Gumbel, Student t)의 Copula 모형을 제안하였으며, 이는 Table 1에 제시하였다.

앞서 언급하였듯이 가뭄빈도해석의 경우 일변량 가뭄빈도 해석을 수행하는 것 보다는 가뭄에 영향을 미치는 복합인자를 함께 고려하는 것이 더 타당하다고 알려져 있다. 이러한 이유로 Shiao and Shen (2001)은 가뭄 지속시간과 가뭄심도를 활용하여 이변량 가뭄빈도해석 식을 제안한 바 있으며, 국내외 다양한 연구에서 이변량 가뭄빈도해석을 활용하여 연구를 진행하였다. 즉, 가뭄빈도해석시 가뭄의 지속시간과 심도는 서로 상관성이 높은 것으로 알려지고 있으며, 이를 활용한 이변량 가뭄빈도해석 모형이 다양하게 적용되었으며, 이는 Eq. (3)과 같이 표현할 수 있다.

$$T_{DDS} = \frac{E(L)}{P(D \geq d, S \geq s)} = \frac{E(L)}{1 - F_D(d) - F_S(s) + C(F_D(d), F_S(s))} \tag{3a}$$

$$T'_{DDS} = \frac{E(L)}{P(D \geq d \text{ or } S \geq s)} = \frac{E(L)}{1 - C(F_D(d), F_S(s))} \tag{3b}$$

이때, $E(L)$ 은 평균가뭄발생 간격이며, T_{DDS} 는 $D \geq d$ and $S \geq s$ 일 때 결합재현기간(joint return period)를 의미하고, T'_{DDS} 는 $D \geq d$ or $S \geq s$ 을 의미한다. Eq. (3)을 살펴보면 가뭄 지속시간과 심도가 동시에 발생할 확률 및 개별로 발생할 확률은 두 변량의 결합확률 $C(F_D(d), F_S(s))$ 을 산정함으로써 가능하다. 그러나 앞서 언급하였듯이 가뭄 지속시간 및 심도 이외 가뭄 강도 또한 중요한 변량 중 하나로서 본 연구에서

는 가뭄 지속시간, 심도 및 강도를 동시에 고려할 수 있는 삼변수 Copula 함수를 기반으로 가뭄빈도분석을 수행하였다. 삼변수의 결합재현기간 산정을 위한 식은 다음 Eq. (4)와 같다. 이때 x_i 는 가뭄 지속시간, y_i 는 심도 그리고 z_i 는 강도를 의미한다.

$$T_{DDSI} = \frac{E(L)}{P(X_i \geq x_i, Y_i \geq y_i, Z_i \geq z_i)} = \frac{E(L)}{(1 - F(x_i) - F(y_i) - F(z_i) + C(F(x_i), F(y_i)) + C(F(x_i), F(z_i)) + C(F(y_i), F(z_i)) - C(F(x_i), F(y_i), F(z_i)))} \tag{4a}$$

$$T'_{DDSI} = \frac{E(L)}{P(X_i \geq x_i \text{ or } Y_i \geq y_i \text{ or } Z_i \geq z_i)} = \frac{E(L)}{1 - C(F(x_i), F(y_i), F(z_i))} \tag{4b}$$

Eq. (4)를 살펴보면 $C(F(x_i), F(y_i))$, $C(F(x_i), F(z_i))$, $C(F(y_i), F(z_i))$ 는 기존 이변량 분석시 정의된 결합확률과 동일하다. 반면, 삼변수의 결합확률은 $C(F(x_i), F(y_i), F(z_i))$ 로 나타낼 수 있으며, 다음 Fig. 2와 같은 절차로 분석되어 질 수 있다. Fig. 2를 살펴보면, 삼변량 Copula 분석시 (a)의 경우는 Archimedean Copula, (b)는 Meta-elliptical Copula (e.g. Gaussian, t Copula)의 분석 방법을 의미한다. 즉, Archimedean Copula의 경우 2 변수간의 조합을 통해 추정된 결합확률을 단계적으로 추정하는 반면, Meta-elliptical Copula의 경우 3개의 변량을 하나의 모형 내에서 추정이 가능하다. 본 연구에서는 Fig. 2에 제시된 두 가지 Copula 방법을 모두 활용하였으며, Archimedean Copula에서는 Gumbel Copula와 Meta-elliptical Copula의 경우 t Copula를 선택하여 분석을 수행하였다.

3. 적용 및 고찰

본 연구에서는 서울 관측소 자료를 기반으로 삼변수 Copula 함수를 활용하여 가뭄 빈도해석을 수행하였으며, 이때 활용되는 가뭄 인자는 가뭄 지속시간, 심도 및 강도이다. 본 연구에서 수행한 방법은 다음과 같이 요약할 수 있다.

- 1) 본 연구에서는 서울 관측소 강우 자료의 월단위 자료를 활용하였으며, 앞서 제시한 Fig. 1과 같은 방법을 기반으로 가뭄 지속시간, 심도 및 강도를 추출하였다. Copula 함수에 적용하기 위해 주변확률분포는 각각 Inverse Gaussian,

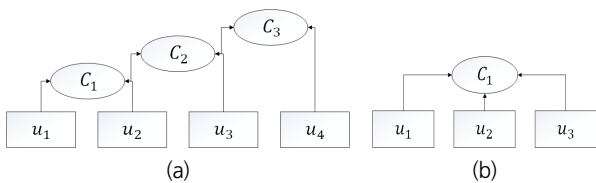


Fig. 2. A schematic representation for the estimation of trivariate joint probability. The left panel indicates the archimedean copula functions based joint probability estimation while the right panel indicates the meta-elliptical copula function based approach

Weibull, Weibull 분포를 활용하였다.

- 2) 삼변량 Copula 함수를 분석하기 위해 각 Copula 함수는 Log-likelihood 비교 분석을 통해 최적의 Copula 함수를 선정하였으며, 본 연구에서는 앞서 언급한 5가지의 Copula 함수를 활용하였다.
- 3) 최종적으로 앞서 제시한 Eq. (4)를 활용하여 삼변수가 고려된 빈도해석을 수행하였으며, 2014~2015년 가뭄을 대상으로 이변량과 삼변량 가뭄빈도해석 결과를 비교하여 제시하였다.

3.1 자료도시 및 Copula 함수 선택

앞서 언급하였듯이 본 연구에서는 구축된 6개월 누적 월강수량 자료를 활용하여 Anomaly를 계산하고, 이를 활용하여 가뭄 특성 인자(가뭄 지속기간, 심도, 강도)를 추출하였다(Fig. 3). 도시 결과 각 가뭄 특성은 1982년과 2014년에 가장 극심하게 발생하고 있는 것을 확인할 수 있었다.

본 연구에서는 삼변수 Copula 함수 적용 전 각 변량에 대한 주변확률분포로서 Log-likelihood와 BIC (Bayesian information criterion, Table 2)를 기준으로 각 변수별로 Inverse Gaussian, Weibull 및 Weibull 분포를 선정하였으며, 누가확률밀도함

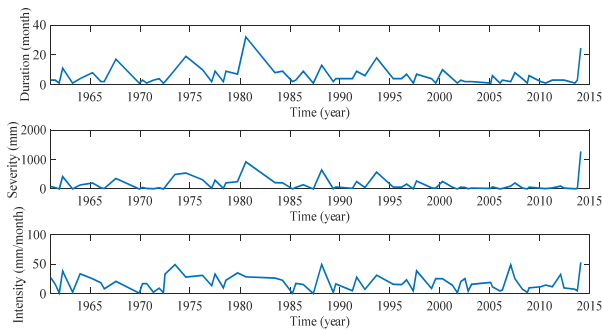


Fig. 3. The drought variables used in this study for Seoul station

수의 형태로 Fig. 4에 도시하였다. 이와 더불어 각 확률분포로 추정된 누가확률밀도 함수와 경험적인 누가확률밀도 함수를 비교한 결과 유사한 거동을 보이고 있는 것을 확인할 수 있었다. 각 주변확률분포에 대한 식은 아래 (5)~(6)과 같다.

$$f(x; \mu, \lambda) = \left[\frac{\lambda}{2\pi x^3} \right]^{1/2} \exp \frac{-\lambda(x - \mu)^2}{2\mu^2 x} \quad (5a)$$

$$F(x; \mu, \lambda) = \Phi \left(\sqrt{\frac{\lambda}{x}} \left(\frac{x}{\mu} - 1 \right) \right) + \exp \left(\frac{2\lambda}{\mu} \right) \Phi \left(- \sqrt{\frac{\lambda}{x}} \left(\frac{x}{\mu} + 1 \right) \right) \quad (5b)$$

$$f(x; k, b) = \frac{k}{b} \left(\frac{x}{b} \right)^{k-1} \exp \left[- \left(\frac{x}{b} \right)^k \right] \quad (6a)$$

Table 2. The table indicated BIC value of each distributions

Drought variable	Distribution	BIC
Duration	Inverse Gaussian	367.357
	Log-normal	370.930
	Gamma	381.899
	Weibull	383.881
	Extreme value	505.158
Severity	Weibull	818.086
	Log-normal	818.314
	Gamma	821.708
	Inverse Gaussian	840.227
	Extreme value	1023.141
Intensity	Weibull	547.765
	Gamma	549.786
	Log-normal	563.086
	Inverse Gaussian	574.050
	Extreme value	587.919

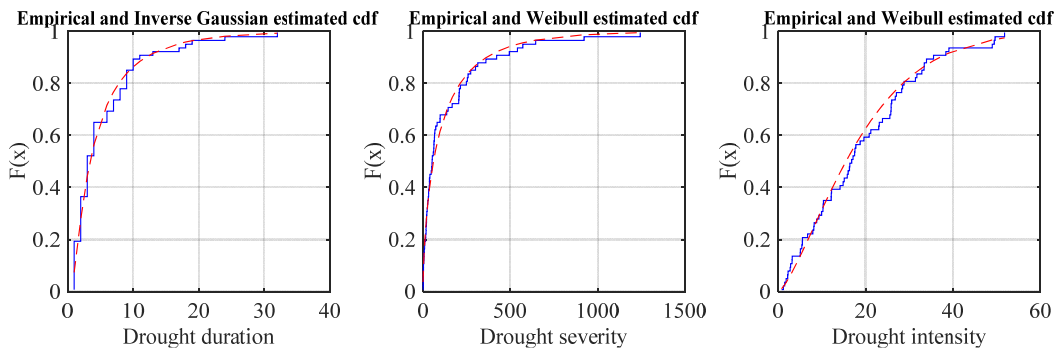


Fig. 4. The theoretical CDFs along with empirical CDFs for the drought variables (i.e. duration, severity and intensity)

$$F(x;k,b) = 1 - \exp\left[-\left(\frac{x}{b}\right)^k\right] \tag{6b}$$

이때, μ 와 λ 는 Inverse Gaussian 분포의 위치(location) 및 형상 매개변수(shape parameter), b 와 k 는 Weibull 분포의 크기 매개변수(scale parameter) 및 형상 매개변수(shape parameter)를 의미한다.

앞서 언급하였듯이 Copula 함수는 각 변량간의 의존성을 파악하고 분석하는데 용이한 기법 중 하나이다. 이에 본 연구에서는 Copula 모형에 적용하기에 앞서 각 변량간 의존성을 상관관계를 통해 분석해 보았다. 산정결과 가뭄 지속시간과 심도는 0.93으로 높은 상관성을 보였으며, 심도와 강도는 0.70 그리고 지속시간과 강도는 0.56을 나타내고 있다. 즉, 가뭄 강도 역시 타 가뭄 변량들과 상관성이 상당히 큰 것으로 파악되었으며, 본 연구에서 제 3의 변량으로 사용하는데 타당성을 가지는 것으로 판단되었다. 상관계수를 정량적으로 산정하여 Table 3에 제시하였으며, 산점도 또한 Fig. 5에 도시하였다. 본 연구에서는 앞서 제시한 Copula 함수 중 각 변량간 최적

의 Copula 함수를 선택하기 위해 Log-likelihood 값을 활용하였다. 잘 알려져 있듯이 Log-likelihood는 모형의 적합성을 판단하기 위해서 사용되어지며, 큰 값을 가지는 Copula 함수가 최적 모형으로 선택되어진다. 다음 Table 4는 각 변량별 Copula 함수의 Log-likelihood 및 매개변수를 나타내었다. 그 결과 Table 4에서 나타냈듯이 1) 관계에서는 Gumbel Copula가 가장 적합한 것으로 도출되었으며, 2)와 3)의 관계에서는 Student t가 가장 적합한 Copula 함수인 것으로 도출되었다. 그러나 국외 논문을 살펴보면 Multivariate Copula 분석시 하나의 Copula 함수를 채택하여 연구를 진행하는 것이 서로 다른 Copula 함수로 사용하는 것보다 통계학적으로 안정적인 결과를 도출한다고 알려져 있다(Ma *et al.*, 2013). 이러한 이유로 Table 4에서 1)은 Gumbel Copula가 가장 적합하지만, Student t Copula의 Log-likelihood가 2번째로 적합한 Copula가 도출되어 본 연구에서는 보다 신뢰성 있는 결과를 도출하기 위해 1)~3)관계 모두 Student t Copula를 채택하여 연구를 진행하였다.

Table 3. The correlation matrix for all drought variables

Correlation coefficient	Duration	Severity	Intensity
Duration	1.000	0.926	0.560
Severity	0.926	1.000	0.704
Intensity	0.560	0.704	1.000

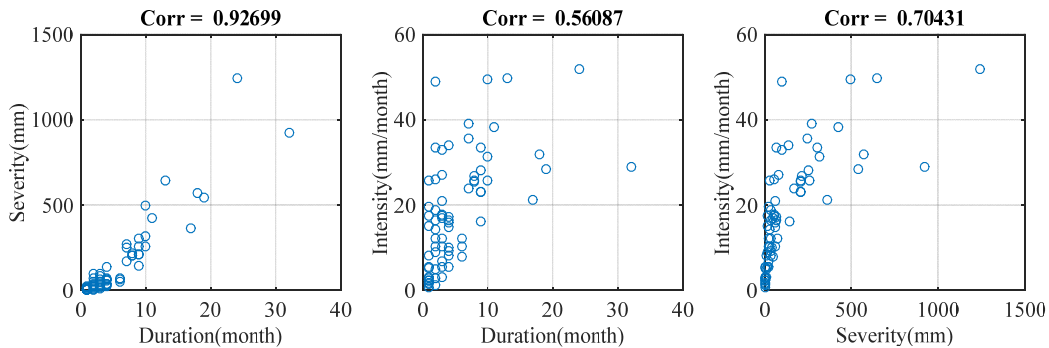


Fig. 5. The scatter plot of pair wise drought variables

Table 4. The estimated log-likelihood between the drought variables

Variables	Gaussian	Clayton	Frank	Gumbel	Student t	θ
1) Duration & Severity	64.176	44.493	57.602	64.958	64.211	3.721 (Gumbel)
2) Duration & Intensity	18.574	15.361	17.600	17.601	18.725	0.638 (Student t)
3) Severity & Intensity	50.978	48.944	49.484	43.979	51.454	0.878 (Student t)

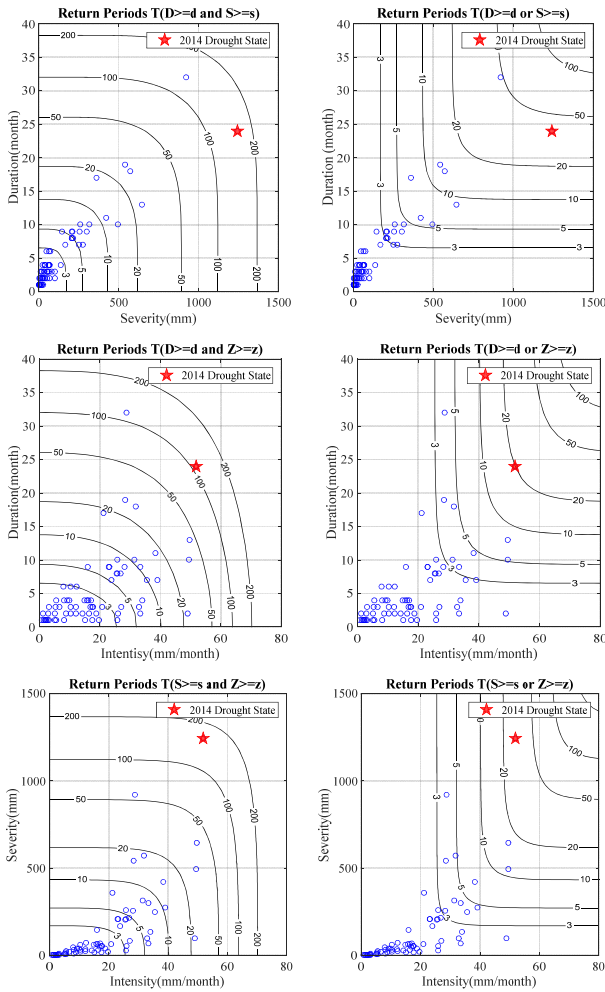


Fig. 6. The joint return periods for the drought variables. The symbol “red stars” indicate the drought state for the 2014~2015 event

3.2 삼변량 Copula 함수를 활용한 빈도해석 결과

앞서 언급하였듯이 삼변량 Student t Copula 함수와 삼변량 Gumbel Copula 함수를 이용하여 빈도해석 결과를 비교 검토하였다. 이에 앞서 본 연구에서는 먼저 이변량 빈도해석을 수행하여 서울 관측소의 현재 가뭄상태를 평가하였으며, 가뭄 빈도해석 결과를 Fig. 6에 제시하였다. 그림을 살펴보면

2014~2015년의 가뭄은 지속시간 24개월, 심도 1,243.60 mm, 강도 51.81 mm/month로 산정되었으며, Table 5의 지속시간과 강도의 관계가 약 110년 빈도(and year), 지속시간 및 심도의 관계에서는 약 160년 빈도(and year)로 산정되었다(Table 5). 반면 삼변량 Student t Copula 함수를 이용한 빈도해석의 경우 약 200년의 재현기간을 갖는 것으로 평가되었다. 이는 이변량 빈도해석 결과와 비교하였을 때 비슷하거나 다소 큰 결과를 나타내고 있다. 삼변량 Copula 모형 결과의 경우 3개의 변량이 모두 고려되어 분석이 이루어지기 때문에 이변량 빈도해석과 비교하는 경우, 비슷하거나 다소 큰 재현기간으로 도출되는 것이 합당하다고 알려져 있다(Chen et al., 2013; Kwon and Lall, 2016; Requena et al., 2013; Salvadori and Michele, 2004). 즉, 이변량 빈도해석 결과와 단변량 해석 결과의 관계는 다음 부등식(inequality)과 같이 나타낼 수 있다.

$$T'_{DDS} \leq \min[T_{DD}, T_{DS}] \leq \max[T_{DD}, T_{DS}] \leq T_{DDS} \quad (7)$$

삼변량 Gumbel Copula 함수를 이용한 빈도해석의 경우 Student t Copula 함수를 이용한 결과에 비해 삼변량 결합재현기간에서 큰 차이를 나타내고 있다. 이는 분포함수의 꼬리에 영향으로 판단되며, 삼변량 빈도해석 시 적합한 Copula 함수의 선정이 매우 중요함을 시사하고 있다.

최종적으로 삼변량 결합재현기간 Eq. (4a)를 활용하여 가뭄상상별 재현기간을 평가하였으며 Fig. 7과 Table 6에 제시하였다. Table 6에서 연도별 첫 번째 행은 Student t Copula 함수로 추정된 결과를 나타내며, 반면 두 번째 행은 Gumbel Copula 함수를 통해 도출된 재현기간을 나타낸다. 앞서 언급하였듯이 이변량 이상 Copula 함수를 이용한 빈도분석시 동일한 Copula 함수를 활용하여 분석하는 것이 더 타당하다고 알려져 있으며, 본 연구에서는 3개의 변량에 대해서 Student t 또는 Gumbel Copula 함수를 모두 적용하는 방식으로 결과를 도출하였다. 그 결과 가뭄 재현기간이 크지 않은 경우에는 추정된 재현기간의 차이가 크지 않았으나, 중·대규모 가뭄 발생 시 추정된 재현기간의 차이가 다소 크게 나는 것을 확인할

Table 5. The comparison of return periods between bivariate and trivariate frequency analysis

Variables	And year		Or year	
	Student t	Gumbel	Student t	Gumbel
1) Duration & Severity	158.292	142.568	37.864	38.890
2) Duration & Intensity	109.625	66.677	19.812	22.423
3) Severity & Intensity	158.364	144.575	28.826	29.335
Trivariate	198.313	288.120	20.019	24.268

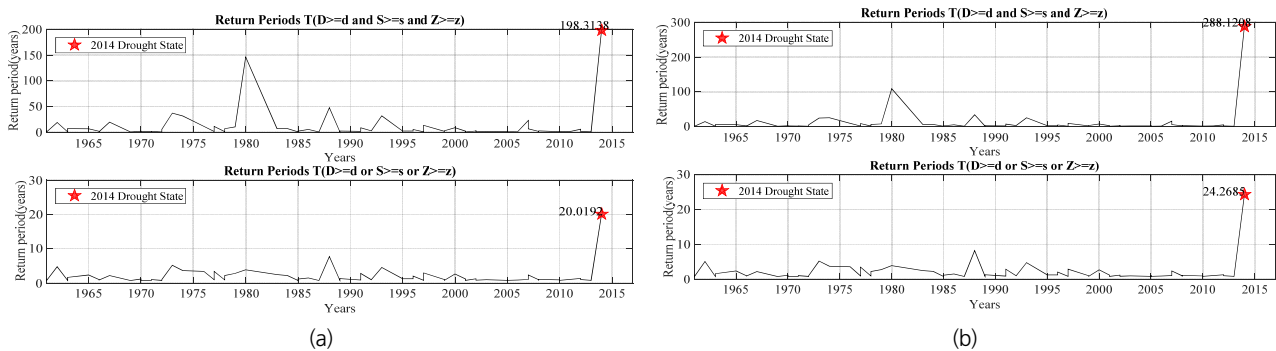


Fig. 7. The result showed return period using trivariate Copula function of Student t (a) and Gumbel (b). The red stars indicate the drought state for the 2014~2015 Drought event

Table 6. The return periods using Student t and Gumbel trivariate copula function

Year	And year				Year	And year			
	Duration & Severity	Duration & Intensity	Severity & Intensity	Trivariate		Duration & Severity	Duration & Intensity	Severity & Intensity	Trivariate
1961	1.841	3.827	3.551	3.971	1978	5.477	6.577	4.187	7.064
	1.851	3.820	3.564	3.308		5.173	6.150	4.138	5.991
1962	11.391	17.154	13.728	19.091	1979	5.093	9.921	8.155	10.679
	10.357	13.824	12.387	13.945		4.866	8.882	7.729	7.526
1963	2.621	6.825	6.196	7.178	1980	132.191	114.632	55.306	146.263
	2.607	6.631	6.123	5.322		104.996	104.040	54.880	108.965
1965	4.853	6.367	4.479	6.843	1983	4.980	6.728	4.835	7.245
	4.585	5.875	4.383	5.573		4.692	6.152	4.703	5.798
1966	1.288	2.025	1.945	2.073	1984	5.489	6.605	4.218	7.098
	1.304	2.060	1.982	1.951		5.182	6.171	4.167	6.010
1967	17.455	18.164	7.625	19.764	1985	1.582	2.181	1.925	2.255
	16.234	17.159	7.578	17.282		1.594	2.210	1.967	2.076
1969	0.828	0.828	0.807	0.834	1986	4.929	5.391	2.754	5.653
	0.835	0.833	0.813	0.845		4.826	5.302	2.782	5.228
1970	1.576	2.151	1.895	2.224	1987	0.821	0.820	0.787	0.823
	1.588	2.182	1.938	2.054		0.826	0.823	0.790	0.830
1971	1.789	1.937	1.358	1.978	1988	24.065	42.342	37.877	47.865
	1.801	1.973	1.396	1.946		22.462	31.699	31.807	33.918
1972	1.597	5.763	5.654	5.869	1989	1.946	2.473	1.978	2.569
	1.609	5.781	5.659	4.918		1.944	2.486	2.020	2.352
1973	13.642	34.337	29.672	37.323	1991	6.124	8.143	5.798	8.844
	12.994	28.161	26.270	24.322		5.683	7.280	5.582	6.993
1974	27.080	26.949	15.799	31.806	1992	2.692	2.789	1.393	2.834
	22.903	23.625	15.456	25.293		2.702	2.821	1.421	2.804
1976	7.844	10.622	7.790	11.672	1993	26.317	26.385	18.102	31.974
	7.159	9.143	7.354	8.978		21.877	22.036	17.464	25.176
1977	6.977	10.606	8.158	11.568	1995	1.923	2.388	1.889	2.478
	6.459	9.137	7.636	8.556		1.923	2.407	1.933	2.291

Table 6. The return periods using Student t and Gumbel trivariate copula function (Continue)

Year	And year				Year	And year			
	Duration & Severity	Duration & Intensity	Severity & Intensity	Trivariate		Duration & Severity	Duration & Intensity	Severity & Intensity	Trivariate
1996	3.921	5.173	3.702	5.515	2006	1.386	1.452	1.095	1.470
	3.761	4.882	3.663	4.581		1.401	1.482	1.123	1.478
1997	5.563	12.637	10.647	13.608	2007	4.874	6.425	4.537	6.908
	5.340	11.266	10.037	9.217		1.979	22.692	22.471	15.621
1999	1.780	1.909	1.318	1.947	2008	2.726	2.894	1.593	2.964
	1.793	1.945	1.354	1.923		2.729	2.923	1.629	2.891
2000	6.846	8.332	5.461	9.064	2010	1.170	1.422	1.332	1.454
	6.331	7.554	5.321	7.446		1.187	1.457	1.370	1.436
2001	1.886	2.256	1.746	2.334	2011	1.462	1.698	1.415	1.742
	1.889	2.282	1.791	2.194		1.477	1.736	1.457	1.695
2002	1.660	2.578	2.324	2.672	2012	2.035	5.854	5.530	6.072
	1.671	2.599	2.360	2.380		2.044	5.809	5.513	4.749
2003	1.244	1.765	1.684	1.807	2013	1.375	1.417	1.033	1.430
	1.260	1.800	1.722	1.732		1.388	1.442	1.056	1.442
2005	2.763	3.001	1.760	3.093	2014	158.293	109.625	158.365	198.314
	2.757	3.022	1.800	2.973		142.568	66.677	144.575	288.121

수 있었다. 따라서 앞서 언급한 바와 같이 삼변량 Copula 분석 시 자료 특성에 적합한 Copula 함수를 활용하는 것이 연구결과의 신뢰성을 확보하는데 유리할 것으로 판단된다.

4. 결론 및 토의

최근 우리나라에서 발생한 가뭄은 댐 용수부족, 지하수 고갈 등 다양한 형태의 가뭄 피해를 발생시켰다. 이러한 가뭄은 피해 발생 시 단순히 용수공급 차원에서의 문제뿐만 아니라, 사회 경제적으로 매우 복합적인 문제를 야기시키고 있다. 이러한 점에서 다양한 관점의 가뭄 위험도를 신뢰성 있게 평가하는 방법론의 개발 및 적용은 가뭄의 인지 및 대응 능력의 효율성을 개선시키는데 매우 중요한 사항이라 할 수 있다. 일반적으로 이러한 가뭄 분석을 위해 가뭄 특성 인자 중 가뭄 지속시간과 심도를 활용하여 연구가 진행되어져 왔다. 그러나 최근 가뭄의 특성을 평가해보면 일반적으로 가뭄 지속시간은 짧아지고, 심도는 커지고 있는 등 가뭄 강도가 증가 경향이 뚜렷해지고 있다. 이러한 점에서 본 연구에서는 가뭄 지속시간, 심도 및 강도를 포괄적으로 고려할 수 있는 삼변량 Copula 함수를 활용하여 2014~2015년의 가뭄을 평가하였으며, 본 연구에서 도출된 결과는 아래와 같다.

첫째, 본 연구에서는 기존 이변량 Copula 모형과 더불어 삼변량 Copula 모형을 활용하여 가뭄의 특성을 평가하였다. 즉, 다양한 가뭄 특성인자를 포괄적으로 고려할 수 있는 분석방안을 도입하고자 하였다. 2014~2015년 가뭄사상을 대상으로 빈도해석을 수행한 결과 가뭄 지속시간과 심도의 경우 약 160년 빈도, 가뭄 지속시간과 강도의 관계는 약 110년 빈도로 평가된 반면, 삼변량 빈도해석의 경우 대략 200년의 재현기간을 평가되었다. 결론적으로 이변량 빈도해석의 경우 삼변량 빈도해석 보다 가뭄위험도를 다소 과소추정 하는 것으로 나타났으며, 가뭄의 대한 위험도를 정량적으로 평가하는데 있어서 이변수 보다는 삼변수 해석이 보다 현실적인 접근방법이 될 수 있다고 판단되었다.

둘째, 삼변량 빈도해석 시 고려되는 Copula 함수의 선택이 가뭄의 재현기간을 추정하는데 있어 매우 민감한 사항으로 평가되었다. 즉, Meta-elliptical Copula 형태의 삼변수 Student t Copula와 Archimedean Copula에 속하는 삼변수 Gumbel Copula를 비교한 결과 Gumbel Copula가 가뭄의 위험도를 다소 과대 추정하는 것으로 평가되었다. 이러한 점에서 최적 Copula 선정에 대한 지표의 도입과 모의를 통한 검증 작업등의 과정이 필수적으로 요구된다 하겠다.

본 연구결과를 통해 다양한 가뭄 특성인자를 고려할 수 있는 삼변량 Copula 함수의 도입으로 복합적인 특성을 가진 가

뮴 분석 시 보다 현실적인 가뭄 위험도 분석이 가능할 것으로 판단된다. 향후 연구로서 기존 이변량 및 삼변량 Copula 모형에 Bayesian 기법을 도입하여 자료, 매개변수 추정, 모형에서 기인하는 불확실성을 정량적으로 추정하고 해석할 수 있는 모형으로 확장하고자 한다.

감사의 글

본 연구는 국토교통부 국토교통기술지역특성화사업 연구개발사업의연구비지원(16RDRP-B076564-03)에 의해 수행되었습니다.

References

- Bonaccorse, B., Cancelliere, A., and Rossi, G. (2003). "An analytical formulation of return period of drought severity." *Journal of Hydrologic Engineering*, Vol. 8, No. 5, pp. 237-246.
- Cancelliere, A., and Salas, J. D. (2010). "Drought probabilities and return period for annual streamflows series." *Journal of Hydrology*, Vol. 391, No. 1-2, pp. 77-89.
- Chen, L., Singh, V. P., Guo, S., Mishra, A. K., and Gou, J. (2013). "Drought analysis based on copulas." *Journal of Hydrologic Engineering*, Vol. 18, No. 7, pp. 797-808.
- Chun, S.-Y., Kim, Y.-T., and Kwon, H.-H. (2015). "Drought frequency analysis using hidden markov chain model and bivariate copula function." *Journal of Korea Water Resource Associate*, Vol. 48, No. 12, pp. 969-979.
- Dai, A., Trenberth, K. E., and Qian, T. (2004). "A global dataset of palmer drought severity index: Relationship with soil moisture and effects of surface warming." *Journal of Hydro-meteorology*, Vol. 5, pp. 1117-1129.
- Favre, A.-C., Adluni, S. E., Perreault, L., Thiémonge, N., and Bobee, B. (2004). "Multivariate hydrological frequency analysis using copulas." *Water Resources Research*, Vol. 40, No. 1, doi: 10.1029/2003WR002456.
- Fernández, B., and Salas, J. D. (1999). "Return period and risk of hydrologic events. I : Mathematical Formulation." *Journal of Hydrologic Engineering*, Vol. 4, No. 4, pp. 297-307.
- Grimaldi, S., and Serinaldi, F. (2006). "Design hyetograph analysis with 3-copula function." *Hydrological Sciences Journal*, Vol. 51, No. 2, pp. 223-238.
- Joe, H. (1997). "Multivariate models and dependence concept." Chapman & Hall, London.
- Kim, J.-Y., Kim, J.-G., Choi, B.-H., and Kwon, H.-H. (2015). "A development of hydrologic dam risk analysis model using bayesian network(BN)." *Journal of Korea Water Resource Associate*, Vol. 48, No. 10, pp. 781-791.
- Kim, T.-W., Valdés, J. B., and Yoo, C. S. (2003). "Nonparametric approach for estimating return periods of droughts in arid regions." *Journal of Hydrologic Engineering*, Vol. 8, No. 5, pp. 237-246.
- Kim, T.-W., Valdés, J. B., and Yoo, C. S. (2006). "Nonparametric approach for bivariate drought characterization using palmer drought index." *Journal of Hydrologic Engineering*, Vol. 11, No. 2, pp. 134-143.
- Kumar, M. N., Murthy, C. S., Sai, M. V. R. S., and Roy, P. S. (2009). "On the use of standardized precipitation index (SPI) for drought intensity assessment." *Meteorological Applications*, Vol. 16, No. 3, pp. 381-389.
- Kwon, H.-H., and Lall, U. (2016). "A copula-based nonstationary frequency analysis for the 2012-2015 drought in California." *Water Resources Research*, doi:10.1002/2016WR018959.
- Laird, K. R., Fritz, S. C., and Cumming, B. F. (1998). "A diatom-based reconstruction of drought intensity, duration, and frequency from moon lake, north dakota: a sub-decadal record of the last 2300 years." *Journal of Paleolimnology*, Vol. 19, pp. 161-179.
- Lee, T. S., and Son, C. Y. (2016). "Analyzing the drought event in 2015 through statistical drought frequency analysis." *Journal of Korea Water Resource Associate*, Vol. 49, No. 3, pp. 177-186.
- Ma, M. W., Song, S. B., Ren, L., Jaing, S. H., and Song, J. L. (2013). "Multivariate drought characteristics using trivariate Gaussian and Student t copulas." *Hydrological processes*, Vol. 27, pp. 1175-1190.
- Nelssen, R. B. (2006). "An Introduction to Copula." Springer, New York, pp. 109-115.
- Radice, R., Marra, G., and Wojtyś, M. (2015). "Copula regression spline models for binary outcomes." *Statistics and Computing*, doi:10.1007/s11222-015-9581-6.
- Requena, A. L., Mediero O., and Marcos, L. G. (2013). "A bivariate return period based on copulas for hydrologic dam design: accounting for reservoir routing in risk estimation", *Hydrology and Earth System Sciences*, Vol. 17, No. 8, pp. 3023-3038.
- Salvadori, G., and Michele, C. (2004). "Frequency analysis via copulas: Theoretical aspects and applications to hydrological events." *Water Resources Research*, Vol. 40, No. 12.
- Salvatierra, I. D. L., and Patton, A. J. (2015). "Dynamic copula models and high frequency data." *Journal of Empirical Finance*, Vol. 30, pp. 120-135.
- Sheffield, J., and Wood, E. F. (2008). "Projected changes in drought occurrence under future global warming from multi-model, multi-scenario, IPCC AR4 simulations." *Climate Dynamics*, Vol. 31, pp. 79-105.
- Shiau, J. T. (2006). "Fitting drought duration and severity with two-dimensional copulas." *Water Resources Management*, Vol. 20, pp. 795-815.
- Shiau, J. T., and Modarres, R. (2009). "Copula-based drought severity-duration-frequency analysis in Iran." *Meteorological Applications*, Vol. 16, pp. 481-489.

- Shiau, J.-T., and Shen, H. W. (2001). "Recurrence analysis of hydrologic droughts of differing severity." *Journal of water resources planning and management*, Vol. 127, No. 1, pp. 30-40.
- Shiau, J.-T., Feng, S., and Nadarajah, S. (2007). "Assessment of hydrological droughts for the Yellow River, China, using copulas." *Hydrological processes*, Vol. 21, pp. 2157-2163.
- Sklar, K. (1959). "Fontions de repartition a n dimensions et leurs marge." *Publ. Inst. Statis. Univ. Paris 8*, p. 11.
- Yevjevich, V. (1967). "On objective approach to derinitions and investigations of continental hydrologic droughts." *Hydrology Paper*, No. 23, Colorado State University, Fort Collins, pp. 4-18.
- Yoo, J. Y., Shin, J. Y., Kim, D. K., and Kim, T.-W. (2013). "Drought risk analysis using stochastic rainfall generation model and copula functions." *Journal of Korea water resources association*, Vol. 46, No. 4, pp. 425-437.
- Yu, J. S., Yoo, J. Y., Lee, J.-H., and Kim, T.-W. (2016). "Estimation of drought risk through the bivariate drought frequency analysis using copula functions." *Journal of Korea Water Resource Associate*, Vol. 49, No. 3, pp. 217-225.
- Zhang, L., and Singh, V. P. (2007). "Gumbel-hougaard copula for trivariate rainfall frequency analysis." *Journal of Hydrologic engineering*, Vol. 12, No. 4, pp. 409-419.