

Copula 모형을 이용한 이변량 강우빈도해석

Bivariate Frequency Analysis of Rainfall using Copula Model

주 경 원* / 신 주 영** / 허 준 행***

Joo, Kyung-Won / Shin, Ju-Young / Heo, Jun-Haeng

Abstract

The estimation of the rainfall quantile is of great importance in designing hydrologic structures. Conventionally, the rainfall quantile is estimated by univariate frequency analysis with an appropriate probability distribution. There is a limitation in which duration of rainfall is restrictive. To overcome this limitation, bivariate frequency analysis by using 3 copula models is performed in this study. Annual maximum rainfall events in 5 stations are used for frequency analysis and rainfall depth and duration are used as random variables. Gumbel (GUM), generalized logistic (GLO) distributions are applied for rainfall depth and generalized extreme value (GEV), GUM, GLO distributions are applied for rainfall duration. Copula models used in this study are Frank, Joe, and Gumbel-Hougaard models. Maximum pseudo-likelihood estimation method is used to estimate the parameter of copula, and the method of probability weighted moments is used to estimate the parameters of marginal distributions. Rainfall quantile from this procedure is compared with various marginal distributions and copula models. As a result, in change of marginal distribution, distribution of duration does not significantly affect on rainfall quantile. There are slight differences depending on the distribution of rainfall depth. In the case which the marginal distribution of rainfall depth is GUM, there is more significantly increasing along the return period than GLO. Comparing with rainfall quantiles from each copula model, Joe and Gumbel-Hougaard models show similar trend while Frank model shows rapidly increasing trend with increment of return period.

Keywords : copula, bivariate frequency analysis, rainfall quantile

요 지

확률강우량은 수공구조물의 설계에 있어 중요한 역할을 하며 이러한 확률강우량의 산정은 일반적으로 일변량 빈도해석을 수행하고 최적의 확률분포형을 찾아냄으로써 계산된다. 하지만 일변량 빈도해석은 수행 시 지속기간이 제한적이라는 단점이 있으며 이를 보완하기 위해 본 연구에서는 이변량 빈도해석을 수행하였다. 다변량 모형인 copula 모형 중 3가지의 분포형을 이용하여 5개 지점의 연최대강우사상에 대해 이변량 빈도해석을 수행하였으며 확률변수로 강우량과 지속기간을 사용하였다. 주변분포형은 강우량에는 Gumbel (GUM), generalized logistic (GLO) 분포형, 지속기간에는 generalized extreme value (GEV), GUM, GLO 분포형이 사용됐으며 copula 모형은 Frank, Joe, Gumbel-Hougaard 모형을 이용하였다. 주변분포형의 매개변수는 확률가중모멘트법을 이용하여 추정하였으며, copula 모형의 매개변수는 준모수방법인 의사최우도법을 사용하여 구하였다. 이를 통해 얻어진 확률강우량을 주변분포형과 copula 모형을 바꾸어가며 비교하였다. 그 결과,

*연세대학교 토목환경공학과, 석사과정 (e-mail: kwjy1@yonsei.ac.kr)

Master Student, School of Civil and Environmental Engineering, Yonsei University, Seoul 120-749, Korea

**연세대학교 산업기술연구소, 연구원 (e-mail: hyjyshin@gmail.com)

Researcher, Engineering Institute, Yonsei University, Seoul 120-749, Korea

***교신저자, 연세대학교 사회환경시스템공학부 토목환경공학과 교수 (e-mail: jhheo@yonsei.ac.kr)

Corresponding Author, Professor, School of Civil and Environmental Engineering, Yonsei University, Seoul 120-749, Korea

주변분포형의 종류에 따른 변화에서는 지속기간의 분포형에는 크게 영향을 받지 않는 것으로 나타났다. 강우량의 분포형에 따라서는 조금씩 차이가 났으며 강우량의 분포형이 GUM일 경우, GLO일 때에 비해 재현기간이 증가할수록 확률강우량이 증가하는 경향이 두드러졌다. Copula 모형별로 비교해보았을 때, Joe, Gumbel-Hougaard 모형은 비슷한 경향을 나타내었으며 Frank 모형은 재현기간의 증가에 따른 확률강우량의 증가가 강하게 나타났다.

핵심용어 : copula, 이변량 빈도해석, 확률강우량

1. 서 론

확률강우량의 산정은 불분명한 미래 사상에 대해 기준을 제시하며 구조물의 설계에 있어 중요한 역할을 한다. 이를 위해서 일반적으로 이변량 빈도해석을 통한 확률강우량의 산정이 이루어지고 있으며 이는 관측된 지속기간별 연최대강우량의 최적의 분포형을 찾아 각 지속기간, 재현기간별 확률강우량을 찾는 것을 기본으로 한다. 하지만 강우의 물리적 특성상, 강우사상은 강우량 외에 지속기간, 강우강도 등의 인자들도 중요하며 이러한 요소들을 동시에 고려할수록 실제 강우사상을 잘 표현할 수 있다. 이러한 점들을 반영하기 위하여 본 연구에서는 다변량 모형을 이용하여 강우사상에 대해 이변량 빈도해석을 수행하고 그 결과를 분석하는 것을 목표로 한다. 다변량 모형은 한 자연현상에 미치는 다양한 요인들에 대하여 종합적으로 고려할 수 있기 때문에 그 현상을 구체적으로 분석할 수 있으며, 수문통계 분야에 있어서도 앞으로 충분히 활용할 수 있는 여지가 많은 연구주제이다.

국내에서는 빈도해석에 대한 많은 연구가 이루어지고 있지만 수문분야에 다변량 모형을 적용한 연구는 아직 미비한 것으로 조사되었다. 이정호 등(2009)은 Gumbel 혼합모형을 이용하여 홍수심도를 평가하였으며, Lee et al. (2010)은 Gumbel 혼합모형을 이용하여 한국의 강우-빈도곡선을 유도하고 확률강우량을 제시하였다. 최근 빈도해석 연구 분야에서는 비정상성을 고려한 빈도해석이나 지역빈도해석을 이용하여 확률강우량을 산정하는 연구가 주를 이루고 있다. 권영문 등(2009)은 강우자료가 Gumbel 분포의 매개변수 변화에 미치는 영향을 고려하였으며, 이길성 등(2009)은 서울지점을 대상으로 기존의 GEV 분포에서 위치매개변수에 시간이 고려된 개선된 GEV 분포를 이용하여 상대평균제곱근오차를 계산하여 비교하였으며, 구여주(2011)는 GEV 분포에서 위치, 규모매개변수 각각에 시간요인을 고려하여 모의실험을 하였으며 평균제곱근오차를 비교, 분석하였다. 지역빈도해석은 관측자료 부

족에 따른 문제를 해결하기 위해 고안되었으며, 이동진과 허준행(2001)은 한강유역의 강우량에 대해 지점빈도해석에 비해 지역빈도해석이 유리함을 밝혔으며, 허준행 등(2007a, 2007b)은 우리나라에 지점빈도해석보다 지역빈도해석이 적합함을 밝혔다.

국외에서는 다변량 빈도해석에 관한 연구가 2000년대 전후로 수문통계 분야에 적용되기 시작했으며 현재까지도 활발히 이루어지고 있는 추세이다. Yue et al. (1999)은 Gumbel 혼합모형을 이용하여 홍수량 빈도해석 기법을 제시하였으며, Zhang and Singh (2006)은 copula 모형을 이용하여 캐나다 Quebec 지역의 홍수량 빈도해석을 수행하였다. Genest and Favre (2007)는 수문자료를 이용하여 copula 모형을 구축하는 흐름을 간단한 예제와 함께 잘 정리하였다.

이변량 빈도해석을 통해 얻을 수 있는 확률강우량은 강우의 지속기간을 정해놓고 표본자료로 사용하기 때문에 다양한 지속기간을 갖는 실제 강우사상과 물리적으로 맞지 않는 단점이 있으며, 산정되는 확률강우량 또한 표본자료에서 사용한 지속기간에 대해서만 얻을 수 있다. 이러한 단점은 이변량 빈도해석을 수행함으로써 해결할 수 있다. 여러 이변량 분포형 중 copula 모형은 전통적인 이변량 분포에 비해 주변분포형을 자유롭게 정할 수 있는 이점이 있어 주변분포의 변수 그대로의 특성을 유지하면서 다변량 모형을 구축할 수 있다(Genest and Favre (2007)). 본 연구에서는 3가지의 이변량 copula 모형을 적용하였으며 매개변수의 추정은 의사최우도법을 사용하였고 확률강우량의 산정은 조건부 copula의 개념을 이용하여 구하였다. 이를 통하여 임의의 지속기간에 대하여 이변량 빈도해석을 통하여 도출된 확률강우량을 산정할 수 있다. 본 연구의 목적은 이러한 일련의 과정을 통해 주변분포형과 copula 모형의 변화에 따른 확률강우량을 도출하고 비교, 분석하였다. 이를 통하여 copula 모형을 우리나라 강우사상에 대하여 적용 시 각 모형과 주변분포형에 의한 특징을 파악하고 기존 이변량 빈도해석을 통해 도출되는 확률강우량과의 차이를 분석하고자 하였다.

2. 이론

2.1 최소무강우시간의 정의 (Inter-Event Time Definition (IETD))

장기간 강우 기록에서 강우사상을 잘 분리해 내기 위해서는 강우사상의 시작과 끝을 구분하는 정확한 기준이 있어야 한다(권재호 등, 2004). 이를 위해서는 지점별로 최소무강우시간(t_{b0})을 결정하여야 하며 이를 통하여 시우량 자료로부터 강우사상들을 분리해낼 수 있다.

2.2 Copula 모형

Copula 모형은 일변량의 주변분포형(marginal distribution)들을 결합확률분포형으로 연결해주는 모형으로 Sklar (1959)가 처음 제시하였으며 2000년을 전후로 다변량 모형의 필요성에 의해 여러 분야에 걸쳐 적용되기 시작했다. 기본적으로 copula 함수는 $C: [0,1]^d \rightarrow [0,1]$ 와 같이 표현된다. 즉, d 개의 서로 다른 확률변수의 결합확률분포형이라 할 수 있으며 Eq. (1)과 같이 정의할 수 있다.

$$H(x_1, x_2, \dots, x_d) = C[F_1(x_1), F_2(x_2), \dots, F_d(x_d)] \quad (1)$$

여기서, H 는 결합확률이며 C 는 copula 함수, $F_i(x_i)$ 는 i 확률변수의 주변분포확률을 의미한다. 본 연구에서 사용된 copula 모형은 Frank (Frank, 1979), Joe (Joe, 1993), Gumbel-Hougaard (Nelson, 2006)이 사용되었으며 이들의 개형은 Table 1과 같다.

여기서, u, v 는 주변분포의 누적확률 값으로 0에서 1 사이의 값을 가지며, θ 는 Copula의 매개변수이며 각 모형별로 범위가 제한되어 있다.

2.3 의사최우도법(Maximum Pseudo-likelihood Method)

매개변수 추정방법에서 최우도법은 널리 사용되는 방법 중 하나이다. 의사최우도법은 순위를 기반으로 하는 준모수 추정방법으로 Oakes (1994)가 제안하고 후에 Genest et al. (1995)과 Shih and Louis (1995)를 통하여 정리되었다.

확률밀도함수 형태의 copula c_θ 가 연속일 때, 의사최우도법을 위한 대수우도(log-likelihood)식은 Eq. (2)와 같이 쓸 수 있다.

$$l(\theta) = \sum_{i=1}^n \log \left\{ c_\theta \left(\frac{R_i}{n+1}, \frac{S_i}{n+1} \right) \right\} \quad (2)$$

여기서, R_i, S_i 는 각 표본집단에서 i 번째 자료의 순위(rank)를 의미하며 위 Eq. (2)를 최대화 하는 θ 가 해당 copula의 매개변수 추정량이 되며, 본 연구에서는 Newton-Raphson 방법을 이용하여 계산하였다.

2.4 재현기간에 따른 조건부 Copula

본 연구에서는 이변량분포형으로 나타나는 확률강우량을 계산하기 위해 조건부 copula의 개념을 이용하였다. 조건부 copula는 하나의 확률변수의 값이 주어졌을 때 다른 쪽의 누가분포확률을 말한다. Copula 모형은 기본적으로 누가분포함수이기 때문에 한 쪽의 확률변수가 일정한 값일 때의 분포를 알기 위해서는 고정되는 변수에 대하여 미분을 한 후에 임의의 값을 입력하면 일변량 확률분포형과 같은 형태를 띠며 quantile를 구할 수 있다. Eqs. (3)~(5)는 조건부 copula를 유도하는 과정을 나타낸다.

$$P = H(X \leq x | Y = y) = C_\theta(u | V = v) \quad (3)$$

$$= \frac{\partial}{\partial v} C_\theta(u, v) | V = v \quad (4)$$

$$= \dot{C}_\theta(u, v) | V = v \quad (5)$$

이와 같은 조건부 copula에서 비초과확률 q 에 따른 각 u 를 얻기 위해서는 copula 역함수가 필요하며 Eq. (6)과 같이 유도할 수 있다.

$$u = \dot{C}_\theta^{-1}(v, q) | V = v, Q = q \quad (6)$$

여기서, 얻게 되는 u 는 $V = v$ 일 때, 비초과확률 q 에 해당하는 확률변수 X 의 값이며, 본 연구에서는 확률강우량에 해당한다.

Table 1. Copula Models

Copula	Formula	Parameter(θ)
Frank	$C_\theta(u, v) = \frac{1}{\ln \theta} \ln \left[\frac{(\theta^u - 1)(\theta^v - 1)}{\theta - 1} \right]$	$\theta \geq 0$
Joe	$C_\theta(u, v) = 1 - [(1-u)^\theta + (1-v)^\theta - (1-u)^\theta(1-v)^\theta]^{1/\theta}$	$\theta \geq 1$
Gumbel-Hougaard	$C_\theta(u, v) = \exp(-[(-\ln u)^\theta + (-\ln v)^\theta]^{1/\theta})$	$\theta \geq 1$

3. 적용

3.1 적용 지점

대상지점으로는 춘천, 강릉, 서울, 원주, 충주 등 총 다섯 곳의 기상청 시우량 자료를 사용하였다. 대상지점들은 우리나라의 동서로 고르게 분포되어 있으며, 강릉의 경우 2002년 24시간 동안 870.5mm라는 기록적인 강우량이 포함되어 있어 빈도해석 시 문제점을 발생시킬 수 있으나, 실제로 발생한 강우사상이고 본 연구의 목적은 copula 모형의 적용성을 검토하는 것이므로 이를 포함하여 분석하였다. 5개 지점 모두 38~50년의 긴 기록자료를 보유하고 있어 빈도해석을 하기에 충분하며 자료에 대한 개요는 Table 2와 같다.

최소무강우시간(t_{10})은 이동률과 정상만(1992)의 연구 결과를 이용하였다. 이를 이용하여 시우량 자료를 강우사상자료로 변환시켰으며, 빈도해석을 수행하기 위하여 연도별 최대강우량을 갖는 사상을 선택하여 사용하였다. 강우사상의 총 기록별, 연최대강우사상별 평균은 Table 3과 같으며 연최대강우사상의 강우량과 지속기간의 시계열 그래프는 Fig. 1과 같다.

3.2 주변분포

주변분포의 분포형 결정을 위해 지점별로 강우량, 지속기간에 대해 몇 가지 분포형을 이용하여 매개변수 추정

및 적합도 검정을 수행하였다. 분포형은 정규분포(NOR), GEV, GUM, 3변수 gamma (GAM), GLO, 총 5가지를 사용하였으며 분포형 결정은 적합도 검정을 토대로 수행하였으며 결과는 Table 4와 같다. 수행된 적합도 검정은 Chi-square (C), Kolmogorov-Smirnov (K), Cramer-von-mises (M), Probability plot correlation coefficient (P) 검정이며 기각되었을 경우 해당 분포형에 약자로 표시하였으며, 전부 통과하지 못하였을 경우에는 'ALL'로 표시하였다.

강우량의 검정결과를 살펴보면 NOR 분포형은 모든 지점에서 기각되며 GEV 분포형은 서울, 원주, 충주에서 PPCC 검정에 대하여, GAM 분포형은 강릉, 서울, 원주에서 모든 검정에 대하여 기각되었다. GUM 분포형은 강릉에서 PPCC 검정만 기각되고 다른 지점에서는 기각되지 않았으며, GLO 분포형은 모든 지점, 모든 검정방법에 대하여 통과하는 것으로 나타났다. 지속기간의 검정결과를 살펴보면 NOR 분포형이 춘천을 제외한 모든 지점에서 1가지 이상의 검정방법에 대하여 기각되며, GAM 분포형은 서울에서 모든 검정방법에 대하여 기각되었다. GEV, GUM, GLO 분포형은 모든 지점, 모든 검정방법에 대하여 통과하는 것으로 나타났다. 따라서 강우량의 분포형은 GUM, GLO 분포형, 지속기간의 분포형은 GEV, GUM, GLO 분포형을 사용하였다.

Table 2. Properties of Target Site

Site	Code	Sample (N)	$t_{10}(hr)$
Chuncheon	000101	1966 ~ 2010 (45)	8
Gangneung	000105	1961 ~ 2010 (50)	10
Seoul	000108	1961 ~ 2010 (50)	10
Wonju	000114	1973 ~ 2010 (38)	10
Chungju	000127	1973 ~ 2010 (38)	11

Table 3. Mean Properties of Rainfall Events

Site	Mean number of events (events/year)	Total Event		Annual Maximum Event	
		Mean Depth (mm)	Mean Duration (hr)	Mean Depth (mm)	Mean Duration (hr)
Chuncheon	70.6	16.3	10.0	196.0	49.0
Gangneung	65.3	18.5	13.4	229.3	68.8
Seoul	68.5	18.3	11.0	208.8	49.8
Wonju	84.1	14.6	10.1	193.0	48.3
Chungju	70.3	15.5	10.6	158.7	47.8

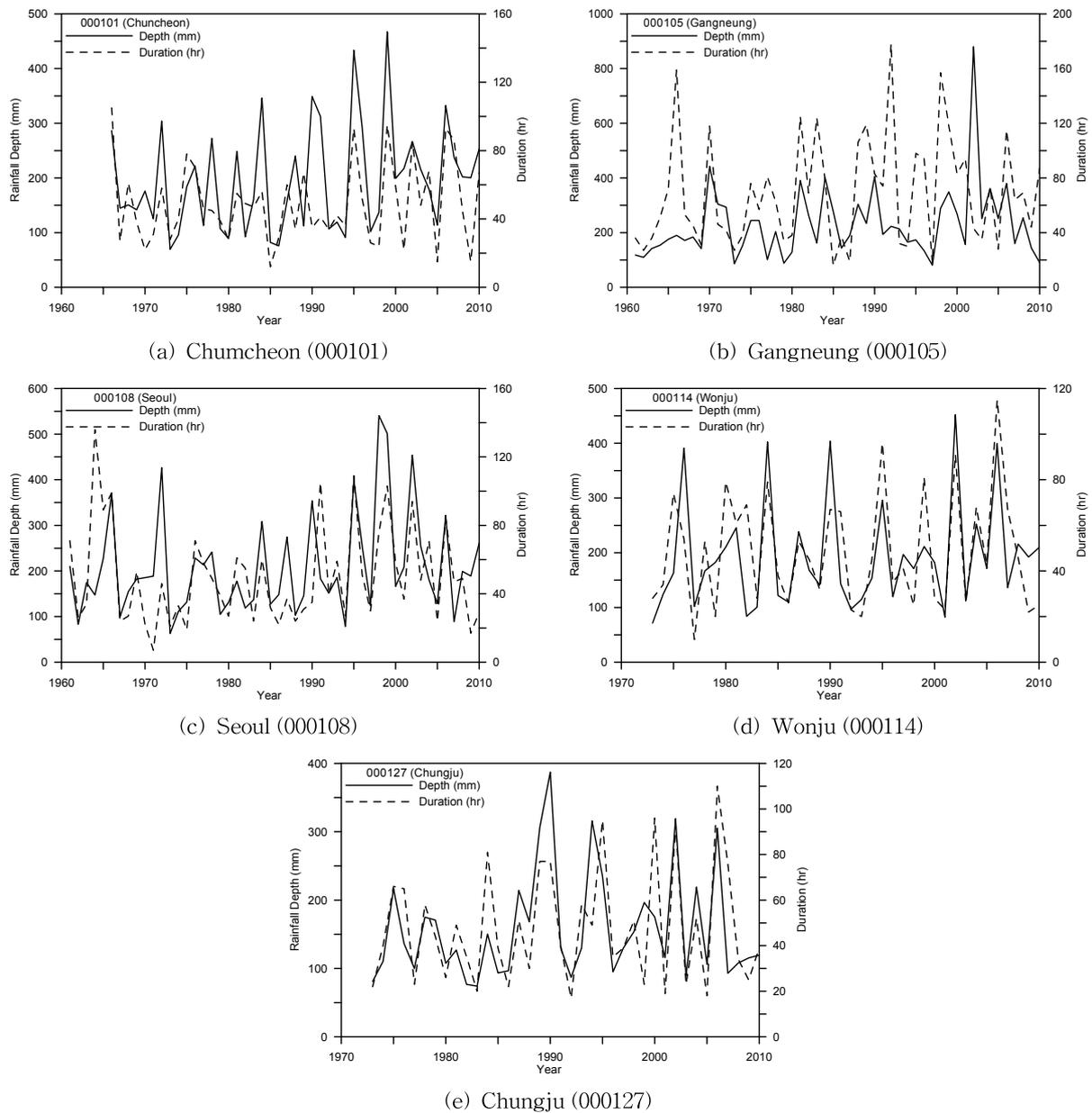


Fig. 1. Time Series of Rainfall Depth and Duration from Annual Maximum Events

Table 4. Results of Goodness-of-Fit Test

Site	Rainfall Depth					Duration				
	NOR	GEV	GUM	GAM	GLO	NOR	GEV	GUM	GAM	GLO
Chuncheon	P	-	-	-	-	-	-	-	-	-
Gangneung	CP	-	P	ALL	-	CP	-	-	-	-
Seoul	KP	P	-	ALL	-	CP	-	-	ALL	-
Wonju	P	P	-	ALL	-	P	-	-	-	-
Chungju	CP	P	-	-	-	P	-	-	-	-

3.3 Copula 모형과 주변분포의 매개변수

Copula 모형의 매개변수 추정은 2.3절에서 소개된 의사 최우도법을 이용하여 추정하였으며 결과는 Table 5와 같다. 주변분포의 매개변수는 확률가중모멘트(probability weighted moment)법을 이용하여 추정하였다.

4. 결과분석

2.4절에서 언급된 조건부 copula의 개념과 3.3절에서 계산된 매개변수들을 이용하여 확률강우량을 계산할 수 있으며, 조건부 copula에 사용되는 지속기간은 24시간, 48시간으로 계산하였다. 4.1절에서는 copula 모형이 같을 때 주변분포별로 확률강우량을 도시하였고 4.2절에서는 주변분포가 같을 때 copula 모형별로 확률강우량을 도시하였다. 각 그래프에서 X축은 재현기간(T)이며 비초과확률 q 에 대해 $T=1/(1-q)$ 와 같이 계산할 수 있다. 또한 각 그림마다 기존의 일변량 지점빈도해석(실선), 지역빈도해석(점선)의 결과도 같이 도시하였으며 지역빈도해석에서의 지역구분은 남우성 등(2008)의 연구를 이용하였다. 분포형은 우리나라에 가장 적합하다고 알려진 GUM 분포형을 적용하였으며 매개변수는 확률가중모멘트법을 이용하여 추정하였다.

4.1 Copula 모형 별 확률강우량

적용된 5개 지점별로 각 3가지 copula 모형의 확률강우량을 도시할 수 있으며 각 copula 모형에는 주변분포의 조합별로 6종류의 확률강우량을 도시할 수 있다. Fig. 2는 서울지점의 각 copula 모형의 주변분포의 조합별 확률강우량을 나타낸 그래프이다.

범례에 쓰여 있는 두 개의 분포형은 주변분포의 분포형(강우량-지속기간)을 나타내며 24시간, 48시간의 지속기간에 대하여 로그 스케일의 재현기간에 따라 선형에 가까운 확률강우량의 증가를 확인할 수 있다. 주변분포의 조

합에 따라 총 6개의 확률강우량이 나타나며 각각의 값이 큰 차이는 나지 않지만 크게 2가지의 경향으로 나뉘어 나타내는데 이는 강우량의 분포형이 GUM인지 GLO인지에 의해 나뉘고 있다. 이는 주변분포형에서 강우량이 지속기간보다 큰 규모 매개변수를 갖기 때문에 기인하는 것으로 보인다. 또한 copula 모형을 통해 산정된 확률강우량을 일변량 빈도해석과의 값과 비교해보면, 지속기간 24시간일 때는 3가지의 copula 모형 모두 비슷한 값을 보이며, 48시간일 때는 Frank 모형은 비슷한 값을 나타내고 나머지 모형에서는 적게 산정됨을 알 수 있다.

4.2. 주변분포의 조합별 확률강우량

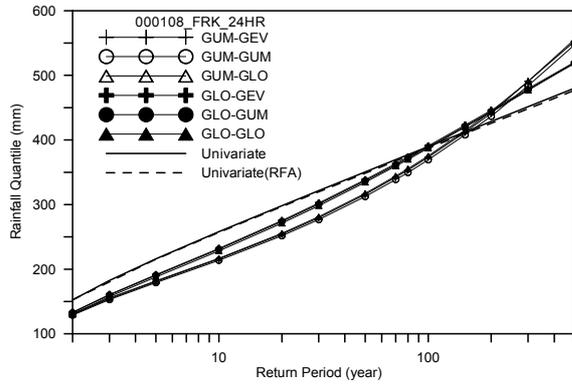
Copula 모형에 따른 확률강우량 변화를 알아보기 위해 주변분포의 조합이 같을 경우 각 copula 모형별 확률강우량과 일변량 빈도해석의 결과를 도시하였으며 Fig. 3과 같다.

전체적으로 Joe와 Gumbel-Hougaard 모형은 비슷한 양상을 나타내며, Frank 모형은 재현기간의 증가에 따른 확률강우량의 상승이 다른 copula 모형에 비해 두드러지며, 지속기간 24시간일 때는 100년 이상, 48시간에는 전 재현기간에서 3개의 모형 중 가장 큰 확률강우량을 갖는다. 일변량 빈도해석 결과와 비교해보면, 강우량이 분포형이 GLO이고 지속기간이 24시간일 때에는 비슷한 값을 나타내며 48시간일 경우에는 Frank 모형은 일변량 빈도해석 결과와 비슷한 확률강우량을 나타내며 다른 모형들은 적게 추정됨을 확인할 수 있다. 강우량의 분포형이 GUM인 경우에도 비슷한 양상을 나타내지만 재현기간의 증가에 따른 상승세가 더욱 두드러지는 것을 확인할 수 있으며 이를 통해 분포형에서 오른쪽 꼬리부분이 더 길 것이라 유추할 수 있다.

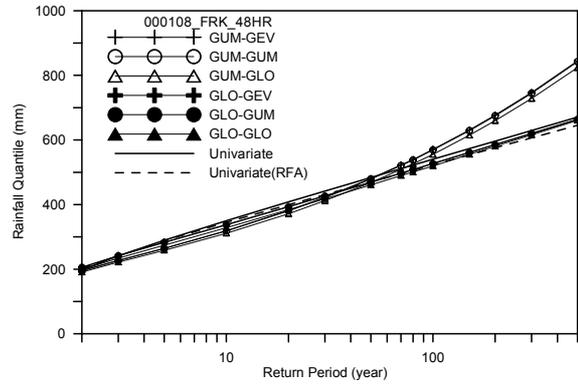
4.1절의 주변분포 분포형에 의한 차이와 4.2절의 copula 모형에 의한 차이를 비교했을 때 copula 모형에 의한 차이가 더 두드러지게 나타나는 것을 확인할 수 있다. 또한 Frank 모형은 지속기간 24시간일 때 짧은 재현기간에서

Table 5. Parameters of Copula(θ)

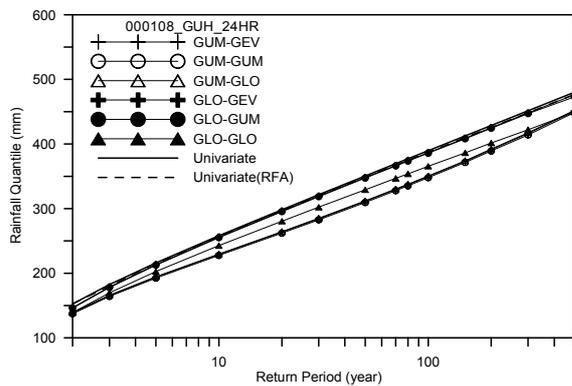
Site	Frank	Gumbel-Hougaard	Joe
Chuncheon	3.90968	1.54404	1.65703
Gangneung	2.70939	1.42618	1.66089
Seoul	4.06160	1.53886	1.66753
Wonju	1.52992	1.58801	4.01999
Chungju	6.60975	1.96108	2.21227



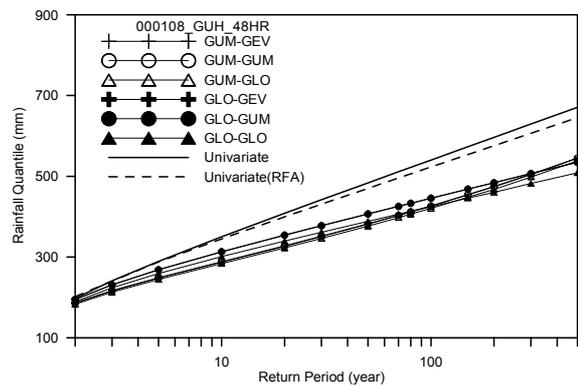
(a) Frank (24 hours)



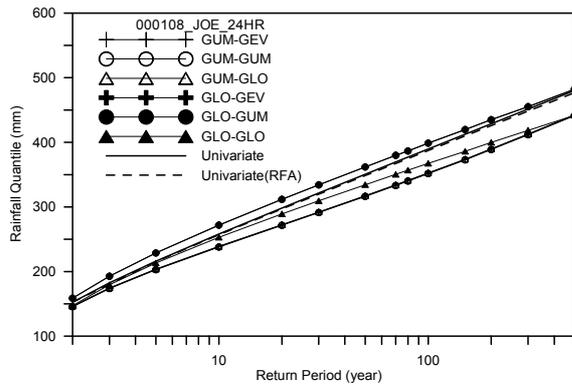
(b) Frank (48 hours)



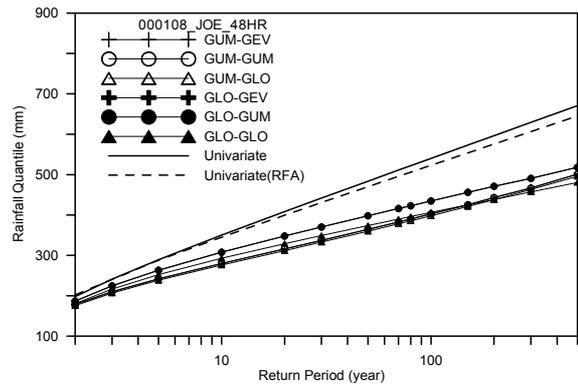
(c) Gumbel-Hougaard (24 hours)



(d) Gumbel-Hougaard (48 hours)



(e) Joe (24 hours)



(f) Joe (48 hours)

Fig. 2. Rainfall Quantile of Seoul from Each Copula Model and Univariate Model

는 다른 copula 모형보다 확률강우량이 작은 부분이 있지만 48시간의 경우에는 계산된 전 재현기간에서 가장 큰 확률강우량을 가지는 것을 확인할 수 있으며, Frank 모형이 지속기간의 증가에 대해 확률강우량의 상승이 가장 큰 것을 알 수 있다.

4.3. 일변량 빈도해석 확률강우량과의 비교·분석

기준에 이루어지던 일변량 빈도해석과의 비교를 위해

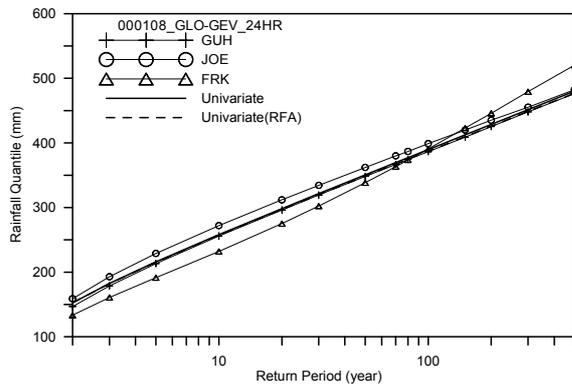
그 결과를 도시하였으며, 일변량 빈도해석의 결과와 비교하여 이변량 빈도해석을 수행했을 경우의 값을 평균절대 상대오차(%)로 Table 6에 나타내었다. 종합적으로 검토할 수 있도록 대표적으로 많이 쓰이는 지속기간 1, 6, 12, 24, 48시간과 재현기간 2, 3, 5, 10, 20, 30, 50, 70, 80, 100, 150, 200, 300, 500년에 대한 절대상대오차를 평균하여 나타낸 값들이며 괄호 안의 값은 지역빈도해석의 평균절대 상대오차이며 그 식은 Eq. (7)과 같다.

$$\text{평균절대상대오차(\%)} = \frac{\sum_{i=1}^n \sum_{j=1}^m \left| \frac{Q_{CPL,i,j} - Q_{UNI,i,j}}{Q_{UNI,i,j}} \right|}{n \times m} \quad (7)$$

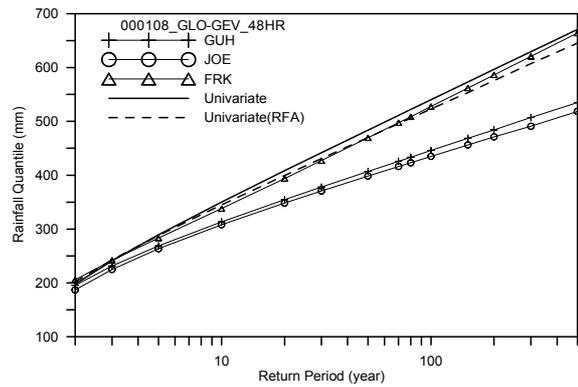
여기서, $Q_{UNI,i,j}$ 와 $Q_{CPL,i,j}$ 는 각각 일변량 빈도분석 결과와 copula 모형의 이변량 빈도분석 결과의 강우지속기간 i , 재현기간이 j 일때의 확률강우량이며 $n=5$ (지속기간), $m=14$ (재현기간)이다.

전체적으로 일변량 빈도분석에 비해 copula 모형의 확률

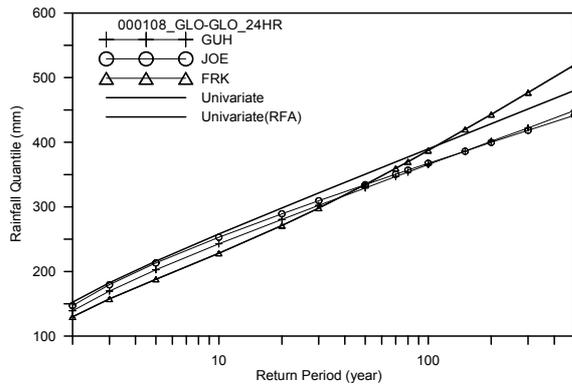
강우량이 크게 계산된다는 것을 알 수 있으며 지점별로 그 차이가 크게 나타나는 것을 확인할 수 있다. Copula 모형별 평균을 살펴보면 세 모형 중 Joe 모형이 41.3% (40.9%)로 가장 크게 계산되며 Frank와 Gumbel-Hougaard 모형은 34.1% (34.0%), 33.8% (33.3%) 크게 산정되었다. 또한 Frank 모형은 지점별로 편차가 큰데 비해 Gumbel-Hougaard 모형은 적은 것으로 보아 각 지점에 따라 Frank 모형이 민감하게 반응하는 것을 알 수 있다. 따라서 일변량 빈도분석과 비교하면 Joe 모형의 경우 다른 모형에 비



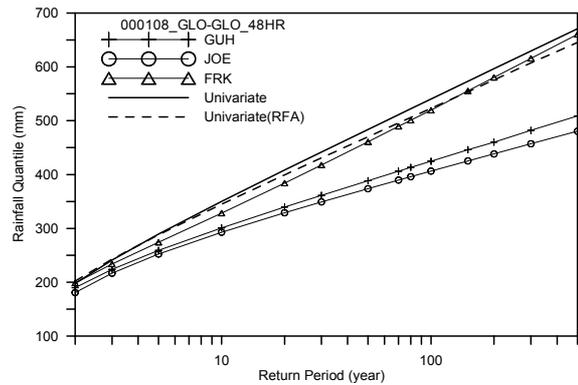
(a) GLO-GEV (24 hours)



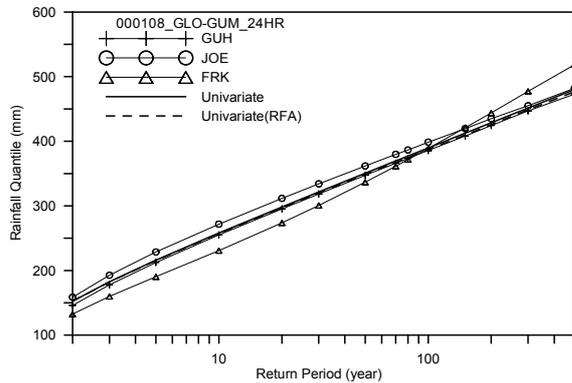
(b) GLO-GEV (48 hours)



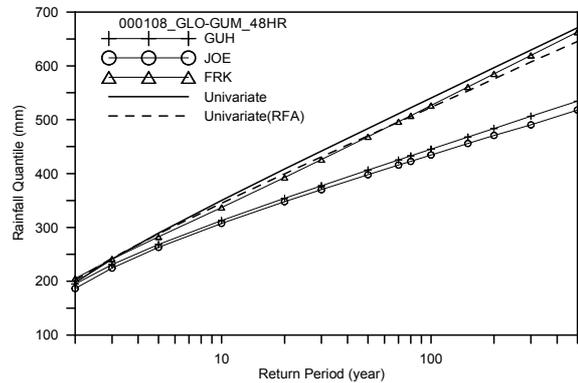
(c) GLO-GLO (24 hours)



(d) GLO-GLO (48 hours)

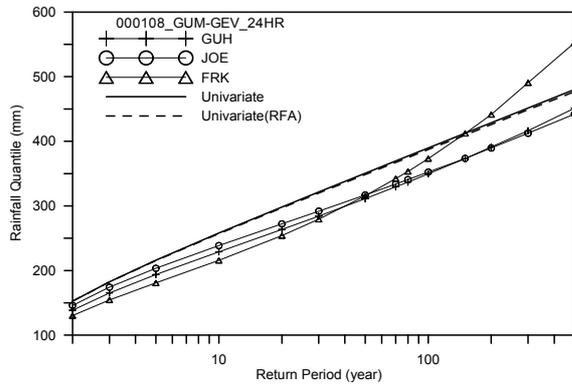


(e) GLO-GUM (24 hours)

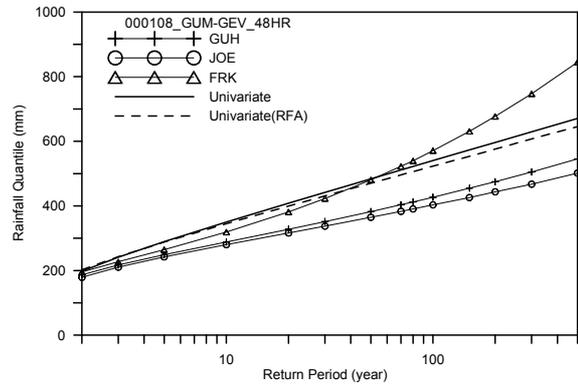


(f) GLO-GUM (48 hours)

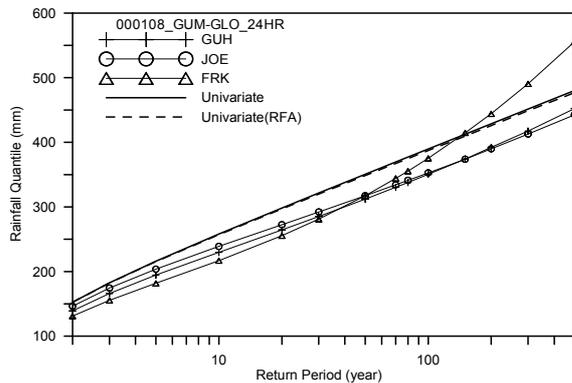
Fig. 3. Rainfall Quantile of Seoul from Each Combination of Marginal Distribution



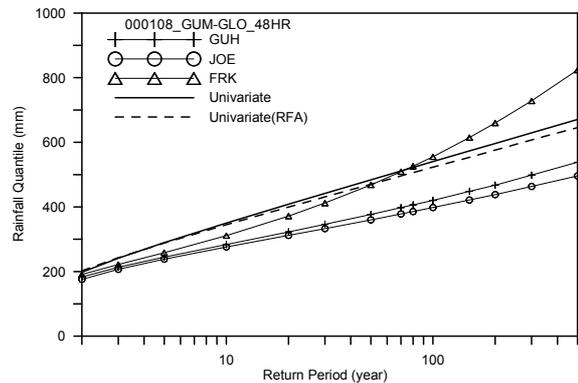
(g) GUM-GEV (24 hours)



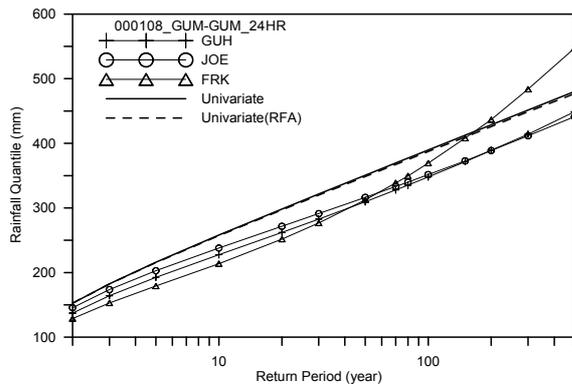
(h) GUM-GEV (48 hours)



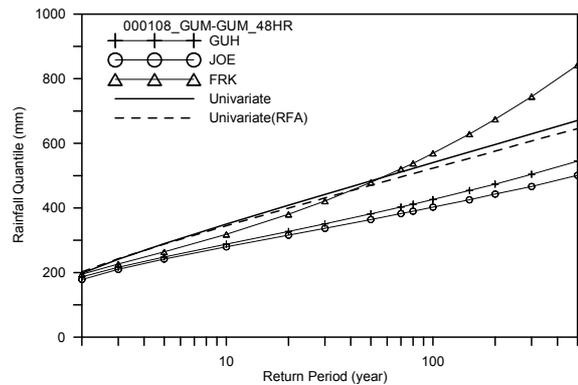
(i) GUM-GLO (24 hours)



(j) GUM-GLO (48 hours)



(k) GUM-GUM (24 hours)



(l) GUM-GUM (48 hours)

Fig. 3. Rainfall Quantile of Seoul from Each Combination of Marginal Distribution (continued)

해 증가폭이 너무 크므로 실제 적용에 어려움이 있을 것으로 판단되며, Frank와 Gumbel-Hougaard 모형은 비슷한 양상을 나타내지만 Frank 모형을 각 지점별로 살펴보면 변동폭이 커서 표본자료에 민감한 특징이 있어 Gumbel-Hougaard 모형이 적용지점에 적절한 것으로 판단된다.

5. 결론

본 연구의 목적은 copula 모형을 이용한 강우사상의 이

변량 빈도해석과 주변분포의 분포형 및 copula 모형에 따른 확률강우량의 산정 및 비교, 분석에 있다. 강우사상의 빈도해석을 위하여 이변량 확률모형인 copula 모형 중 세가지의 모형을 이용하였으며 연최대강우량 사상을 입력자료로 사용하고, 확률변수로는 강우량과 지속기간을 사용하였다. Copula 모형 매개변수의 추정은 의사최우도법을 이용했으며, 주변분포는 확률가중모멘트법을 이용하여 매개변수를 추정하고 적합도 검정을 실시한 결과 강우량은 GEV와 GUM분포형, 지속기간은 GEV, GUM, GLO

Table 6. Mean Absolute Relative Error Between of Bivariate and Univariate (%)

Site		Frank		Gumbel-Hougaard		Joe	
		GLO	GUM	GLO	GUM	GLO	GUM
Chuncheon	GEV	44.7(35.6)	44.1(35.3)	54.0(46.1)	45.0(36.0)	45.8(36.9)	54.0(46.2)
	GLO	44.7(35.6)	43.7(34.9)	54.0(46.1)	45.4(36.4)	44.7(36.1)	54.7(46.9)
	GUM	45.7(36.7)	46.4(37.7)	54.7(46.9)	45.4(36.3)	44.3(35.7)	54.7(46.9)
Gangneung	GEV	51.2(58.4)	44.7(50.2)	50.6(56.5)	51.6(58.9)	46.0(51.7)	50.7(56.6)
	GLO	51.5(58.8)	45.7(51.4)	50.7(56.5)	50.0(56.9)	45.8(51.4)	51.7(57.5)
	GUM	50.4(57.3)	47.1(52.8)	51.8(57.6)	50.3(57.2)	46.8(52.5)	51.7(57.6)
Seoul	GEV	25.0(26.0)	24.7(25.2)	33.2(33.7)	25.2(26.2)	25.8(26.3)	33.2(33.8)
	GLO	25.2(26.2)	26.5(27.0)	33.4(33.9)	25.4(26.0)	26.0(26.4)	34.1(34.5)
	GUM	25.6(26.2)	27.1(27.5)	34.1(34.6)	25.8(26.4)	27.7(28.2)	34.2(34.7)
Wonju	GEV	29.2(28.5)	30.6(29.7)	39.9(38.7)	29.7(28.9)	31.6(30.7)	40.0(38.8)
	GLO	29.2(28.4)	30.4(29.6)	39.9(38.7)	29.6(28.7)	31.8(30.9)	40.3(39.1)
	GUM	30.0(29.1)	32.8(31.9)	40.4(39.2)	29.5(28.6)	31.6(30.8)	40.3(39.1)
Chungju	GEV	19.5(20.9)	19.4(20.6)	26.2(27.1)	19.5(20.9)	20.4(21.5)	26.3(27.1)
	GLO	19.5(21.0)	20.2(21.3)	26.3(27.1)	19.8(21.1)	19.8(20.8)	27.8(28.1)
	GUM	19.9(21.2)	21.0(22.0)	27.8(28.2)	19.9(21.1)	20.7(21.8)	27.9(28.2)
Average		34.1(34.0)		33.8(33.3)		41.3(40.9)	

분포형을 적용하였다. 조건부 copula의 개념을 이용하여 지속기간 24시간, 48시간 강우사상에 대하여 확률강우량을 계산하였다. 확률강우량이 특정 copula에 대하여 주변 분포 조합의 변화에 따라 어떤 특성을 나타내는지, 그리고 주변분포가 같을 때 copula 모형에 따라 어떻게 달라지는지 연구를 수행하였다.

본 연구에서 수행된 copula 모형을 이용한 이변량 빈도 해석은 일변량 빈도해석에 비해 하나의 자료계열을 더 사용함으로써 복잡한 자연현상을 실제 모형으로 구축하는데 있어서 물리적으로 더 적절하며, 이변량 모형 중에서도 copula 모형을 사용함으로써 주변분포형을 자유롭게 선정하여 자료계열의 특성을 유지하면서 모형을 구성할 수 있었다. 강우사상에 대해 copula 모형을 통한 이변량 빈도해석의 결과로 copula 모형 중 Frank 모형은 재현기간의 증가에 따른 확률강우량의 증가가 두드러졌고 Joe 모형에 의한 확률강우량이 가장 크게 추정되었으며, 강우량의 주변분포형이 GUM일 경우에 재현기간에 따른 확률강우량의 증가세가 두드러지는 것을 확인하였다. 분석 결과에 따르면 각 3가지의 copula 모형별, 강우량의 주변분포형, 지속기간의 분포형별로 확률강우량의 차이가 나타났으

며 나열한 순서대로 확률강우량의 차이가 크게 두드러졌다. 따라서 copula 모형을 이용하여 확률강우량을 제시할 경우, copula 모형의 선정이 가장 중요하며 그 다음으로는 강우량의 주변분포형 선정에 주의하여야 한다. 또한 기존의 일변량 빈도해석의 결과와 비교해 보았을 때 Joe 모형은 크게 추정되는 경향이 있으며 Frank 모형은 다른 모형에 비해 표본자료에 민감하여 Gumbel-Hougaard 모형이 적절한 것으로 판단된다. 또한 지점별로 편차의 변동이 크므로 copula 모형을 적용하여 빈도해석을 실시할 경우, 일괄적인 모형을 적용하기보다 각 지점별로 적합한 copula 모형이 무엇인지에 대하여 고려할 필요가 있다.

Copula 모형은 수학적으로는 많은 발전이 이루어지고 있지만 매개변수의 추정이나 수문자료에 적합한 모형의 정립이 이루어지지 않아 아직 실제 실무에 쓰이기에는 다소 어려운 부분이 있다. 또한 copula 모형이 국내의 수문자료에 대하여 적용된 사례가 전무하여 아직 적용하는 단계라 할 수 있지만 copula 모형을 사용함으로써 이변량, 더 나아가 다변량 빈도해석을 수행할 수 있다는 점과 국외에서도 빠른 추세로 연구 및 적용되고 있다는 점에서 복잡한 자연현상을 이해하고 예측하는 데 있어 사용되는

빈도해석의 새로운 방법을 제시할 수 있을 것이다.

감사의 글

본 연구는 국토해양부 한국건설교통기술평가원의 2009 건설기술혁신사업인 ‘기후변화에 의한 수문 영향분석과 전망’ 과제에 의해 지원되었습니다.

참고문헌

구여주 (2011). 비정상성을 고려한 GEV 분포의 매개변수 추정. 석사학위논문, 연세대학교.

권영문, 박진원, 김태웅 (2009). “강우량의 증가 경향성을 고려한 목표연도 확률강우량 산정.” **대한토목학회 논문집**, 대한토목학회, 제29권, 제2B호, pp. 131-139.

권재호, 박무중, 김중훈 (2004). “비점오염원 산정을 위한 강우 분석.” **한국수자원학회 2004년 학술발표회 논문집**, 한국수자원학회, pp. 206-211.

남우성, 김태순, 신주영, 허준행 (2008). “다변량 분석 기법을 활용한 강우 지역빈도해석.” **한국수자원학회 논문집**, 한국수자원학회, 제41권, 제5호, pp. 517-525.

이길성, 강원규, 박경신, 성진영 (2009). “개선된 GEV 분포를 이용한 강우량 빈도분석.” **한국수자원학회 2009년 학술발표회 논문집**, 한국수자원학회, pp. 1321-1326.

이동률, 정상만 (1992). “한강유역 강우의 시·공간적 특성.” **한국수자원학회지**, 한국수자원학회, 제25권, 제4호, pp. 75-85.

이동진, 허준행 (2001). “L-모멘트법을 이용한 한강유역 일강우량 자료의 지역빈도해석.” **한국수자원학회 논문집**, 한국수자원학회, 제34권, 제2호, pp. 119-130.

이정호, 정건희, 김태웅 (2009). “이변량 Gumbel 혼합모형을 이용한 홍수심도 평가.” **한국수자원학회 논문집**, 한국수자원학회, 제42권, 제9호, pp. 725-736.

허준행, 이영석, 신흥준, 김경덕 (2007a). “우리나라 강우자료의 지역빈도해석 적용성 연구(I): 확률강우량산정.” **대한토목학회 논문집**, 대한토목학회, 제27권, 제2B호, pp. 101-111.

허준행, 이영석, 신흥준, 김경덕 (2007b). “우리나라 강우자료의 지역빈도해석 적용성 연구(II): 모의실험 및 적합기법 선정.” **대한토목학회 논문집**, 대한토목학회, 제27권, 제2B호, pp. 113-123.

Frank, M.J. (1979). “On the simultaneous associativity

of $F(x,y)$ and $x+y-F(x,y)$.” *Aequationes Math*, Vol. 10, No. 1, pp. 194-226.

Genest, C., and Favre, A.C. (2007). “Everything you always wanted to know about copula but afraid to ask.” *Journal of Hydrology*, No. 12, issue. 4, pp. 347-368.

Genest, C., Ghoudi, K., and Rivest, L.P. (1995). “A semiparametric estimation procedure of dependence parameters in multivariate families of distribution.” *Biometrika*, Vol. 82, No. 3, pp. 543-552.

Joe, H. (1993). “Parametric families of multivariate distributions with given margins.” *Journal of Multivariate Analysis*, Vol. 46, No. 2, pp. 262-282.

Lee, C.H., Kim, T.W., Chung, G., Choi, M., and Yoo, C. (2010). “Application of bivariate frequency analysis to the derivation of rainfall-frequency curves.” *Stochastic Environmental Research and Risk Assessment*, Vol. 24, No. 3, pp. 389-397.

Nelson, R.B. (2006). *An Introduction of Copulas*, Springer-Verlag, New York.

Oakes, D. (1994). “Multivariate survival distribution.” *Journal of Nonparametric Statistics*, Vol. 3, No. 3, pp. 343-354.

Shih, J.H., and Louis, T.A. (1995). “Inferences on the association parameter in copula models for bivariate survival data.” *Biometrics*, Vol. 51, No. 4, pp. 1384-1399.

Sklar, A. (1959). “Fonctions de repartition a n dimension et leurs marges.” *Publications de Institut de Statistique Universite de Paris*, Vol. 8, pp. 229-231.

Yue, S., Ouarda, T.B.M.J., Bobee, B., Legendre, P., and Bruneau, P. (1999). “The Gumbel mixed model for flood frequency analysis.” *Journal of Hydrology*, Vol. 226, pp. 88-100.

Zhang, L., and Singh, V.P. (2006). “Bivariate flood frequency analysis using the copula method.” *Journal of Hydrologic Engineering*, Vol. 11, No. 2, pp. 150-164.

논문번호: 11-122	접수: 2011.11.11
수정일자: 2012.05.15/05.29	심사완료: 2012.05.29