

Multivariate frequency analysis of urban rainfall characteristics using three-dimensional copulas

Chenglin Liu, Yuwen Zhou, Jun Sui and Chuanhao Wu

ABSTRACT

Urban runoff is a major cause of urban flooding and is difficult to monitor in the long term. In contrast, long term continuous rainfall data are generally available for any given region. As a result, it has become customary to use design rainfall depth as a proxy for runoff in urban hydrological analyses, with an assumption of the same frequency for runoff and rainfall. However, this approach has lack of overall coordination and cannot fully reflect the variability of rainfall characteristics. To address this issue, this study presents a three-dimensional copula-based multivariate frequency analysis of rainfall characteristics based on a long term (1961–2012) rainfall data from Guangzhou, China. Firstly, continuous rainfall data were divided into individual rainfall events using the rainfall intensity method. Then the characteristic variables of rainfall (design rainfall depth, DRD; total rainfall depth, TRD; peak rainfall depth, PRD) were sampled using the annual maximum method. Finally, a copula method was used to develop the multivariate joint probability distribution and the conditional probability distribution of rainfall characteristics. The results showed that the copula-based method is easy to implement and can better reflect urban rainstorm characteristics. It can serve a scientific reference for urban flood control and drainage planning.

Key words | copula function, design rainfall depth, multivariate analysis, peak rainfall depth, total rainfall depth, urban rainfall

Chenglin Liu
Yuwen Zhou

College of Architecture and Civil Engineering,
Beijing University of Technology,
Beijing 100124,
China

and

Key Laboratory of Beijing for Water Quality Science
and Water Environment Recovery Engineering,
Beijing 100124,
China

Chenglin Liu
Jun Sui

Guangzhou Municipal Engineering Design &
Research Institute,
Guangzhou 510060,
China

Chuanhao Wu (corresponding author)

Institute of Groundwater and Earth Sciences,
Jinan University,
Guangzhou 510632,
China

E-mail: wuch0907@jnu.edu.cn

NOTATION

DRD	design rainfall depth
TRD	total rainfall depth
PRD	peak rainfall depth
P3	P-III distribution
LN2	lognormal distribution
GEV	generalized extreme value distribution
PPCC	probability point correlation coefficient
RMSE	root-mean-squared error
MAE	mean absolute error
MSE	mean-square error
AIC	Akaike information criterion
OLS	ordinary least squares

INTRODUCTION

The accuracy of the design flood estimates primarily depends on the runoff data series used. In urban

catchments, runoff data are usually limited and cannot be directly used for the design flood calculations. In contrast, long term rainfall monitoring is widespread and relatively systematic, which provides an indirect method for the design flood estimates (Zhang *et al.* 2008). Thus, urban hydrological analyses (e.g. design flood estimates and determinations of facility scales for urban flood control system) are generally conducted based on the rainfall data, with the assumption of the same frequency for rainfall and runoff.

Rainfall events are generally described using characteristic variables, such as rainfall depth, duration, intensity, and rainfall variation (Andreassian *et al.* 1996; Aronica *et al.* 2005; Notaro *et al.* 2013). Design rainfall is an important parameter in surface runoff calculations, which includes the following key elements: frequency (return period), depth (design rainfall depth, DRD; total rainfall depth, TRD; and peak rainfall depth, PRD), and duration (design duration, total duration, and peak position). As known, rainstorm intensity formula is the most common univariate method

for the calculation of design rainfall (Ren *et al.* 2004). However, it fails to calculate the characteristic variables of rainfall (e.g. TRD and PRD) that may influence the storage capacities of the flood facilities. The limitation of univariate methods for the frequency analysis of rainstorm events has attracted attention in recent decades. For this reason, numerous studies have attempted to describe the rainstorm events in a more comprehensive manner by using bivariate distributions to get a detailed relationship between the different eigenvalues of the rainstorm (Hashino 1985; Singh & Singh 1991; Bacchi *et al.* 1994; Yue *et al.* 2001; Zhang *et al.* 2008; Lee *et al.* (2010); Fontanazza *et al.* 2011). Unfortunately, there is very little research on the joint distributions of three or more variables to describe rainstorm events except for a few studies, such as Grimaldi & Serinaldi (2006), which performed a multivariate analysis, using copula approach, to define the trivariate cumulative distribution functions (CDF) of critical depth–total depth–peak in Umbria. It is therefore necessary to perform additional comparative analyses to better understand the adaptability of multivariate analysis, especially for waterlogging prone cities.

Copulas (Sklar 1959) are considered as an excellent mathematical tool for solving multivariate probability problems, since they can easily connect univariate marginal distribution functions with the multivariate probability

distribution (Salvadori & De Michele 2004; Nelsen 2006). During the past several decades, copulas have been used extensively in multivariate frequency analysis of rainfall (Zhang & Singh 2007a, 2007b; Fontanazza *et al.* 2011; Zhang *et al.* 2013; Wu *et al.* 2014), floods (Zhang & Singh 2006; Hou *et al.* 2010), droughts (Shiau 2006), and low water situations (Yu *et al.* 2014). On the basis of long-term (1961–2012), short-duration rainfall data, the aim of this study is to develop a multivariate probability distribution model of the characteristic variables of rainfall (DRD, PRD, and TRD) in Guangzhou City, China, using the three-dimensional (3-D) copulas. The model is employed to determine the conditional probabilities for various combinations of the rainfall characteristic variables.

STUDY AREA

Guangzhou (from 22° 26' to 23° 56' N and 112° 57' to 114° 03' E), located in the northern edge of the Pearl River Delta, is the capital of Guangdong province (Figure 1). It borders the South China sea and is adjacent to Hong Kong and Macau. Guangzhou has a maritime, subtropical, monsoon climate, with an annual mean temperature of 21.9 °C, an annual mean relative humidity of 77%, and an annual

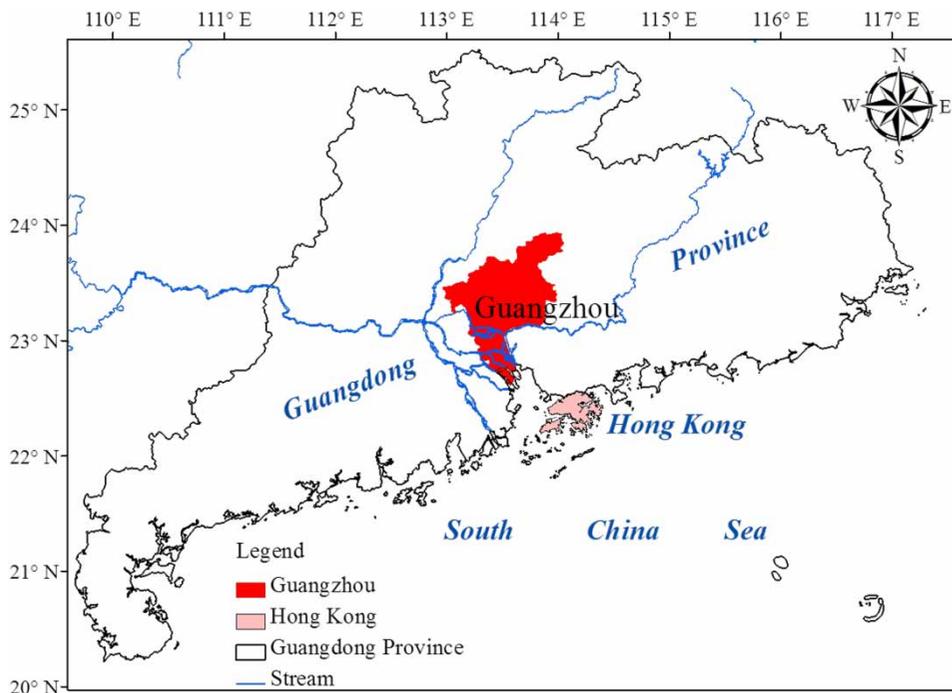


Figure 1 | Location of Guangzhou City.

mean rainfall depth more than 1,700 mm. The rainy season occurs from April to June, which shows the largest number of heavy rain days in China. In recent years, urban waterlogging has become more frequent in Guangzhou, and short-duration rainstorms are one of the main factors leading to waterlogging (Wu et al. 2014).

METHODOLOGY

Rainfall dataset

The long-term (1961–2012) rainfall data with a temporal resolution of 1 min, measured by the tipping bucket rain gauge, are driven from the Wushan station in Tianhe District of Guangzhou. The continuous rainfall time series was first divided into individual rainfall events using the rainfall intensity method (Powell et al. 2003). Independent rainfall events are identified when they can be separated by a dry period, in which the rainfall intensity less than 0.1 mm/h lasts for at least 10 h. As a result, a total of 7,833 individual rainfall events were collected during the period 1961–2012.

Table 1 | Kendall's τ correlation matrix of DRD, TRD, and PRD

	DRD	TRD	PRD
DRD	1.00	0.89	0.17
TRD	0.89	1.00	0.12
PRD	0.17	0.12	1.00

Table 2 | Archimedean copulas used in this study

	Copulas	Formula	Empirical relationship between θ and τ
2-D	Clayton copula	$C(u_1, u_2) = (u_1^{-\theta} + u_2^{-\theta} - 1)^{-1/\theta}, \theta \in (0, \infty)$	$\tau = \frac{\theta}{2 + \theta}$
	GH copula	$C(u_1, u_2) = \exp \{- [(-\ln u_1)^\theta + (-\ln u_2)^\theta]^{1/\theta}\}, \theta \in [1, \infty)$	$\tau = 1 - \frac{1}{\theta}$
	AMH copula	$C(u_1, u_2) = u_1 u_2 / [1 - \theta(1 - u_1)(1 - u_2)], \theta \in [-1, 1)$	$\tau = \left(1 - \frac{2}{3\theta}\right) - \frac{2}{3} \left(1 - \frac{1}{\theta}\right)^2 \ln(1 - \theta)$
	Frank copula	$C(u_1, u_2) = -\frac{1}{\theta} \ln \left[1 + \frac{(e^{-\theta u_1} - 1)(e^{-\theta u_2} - 1)}{(e^{-\theta} - 1)} \right], \theta \neq 0$	$\tau = 1 + \frac{4}{\theta} \left[\frac{1}{\theta} \int_0^\theta \frac{t}{e^t - 1} dt - 1 \right]$
3-D	Clayton copula	$C(u_1, u_2, u_3) = (u_1^{-\theta} + u_2^{-\theta} + u_3^{-\theta} - 2)^{-1/\theta}, \theta \in (0, \infty)$	-
	GH copula	$C(u_1, u_2, u_3) = \exp \{- [(-\ln u_1)^\theta + (-\ln u_2)^\theta + (-\ln u_3)^\theta]^{1/\theta}\}, \theta \in [1, \infty)$	-
	AMH copula	$C(u_1, u_2, u_3) = u_1 u_2 u_3 / [1 - \theta(1 - u_1)(1 - u_2)(1 - u_3)], \theta \in [-1, 1)$	-
	Frank copula	$C(u_1, u_2, u_3) = -\frac{1}{\theta} \ln \left[1 + \frac{(e^{-\theta u_1} - 1)(e^{-\theta u_2} - 1)(e^{-\theta u_3} - 1)}{(e^{-\theta} - 1)^2} \right], \theta \neq 0$	-

θ denotes the parameter of copula; μ_1, μ_2 , and μ_3 are marginal distribution functions; τ denotes Kendall rank correlation coefficient.

The annual maximum method (Jenkinson 1955) was then employed for sampling according to 12-h design duration, which yields a 52-year time series of annual maximum rainfall events. Finally, the characteristic variables of rainfall, i.e. DRD, TRD and PRD, are obtained from each of the annual maximum rainfall events. Table 1 shows the Kendall rank correlation matrix for DRD, TRD and PRD. As can be seen, all the variables are positively correlated. Moreover, the correlation between DRD and TRD (PRD) is stronger than that between PRD and DRD (TRD).

Copula function

Copulas are functions that describe the correlation of variables without limiting their marginal distribution. On the basis of copulas, a number (k) of marginal distributions (in any form) can be combined to generate a multivariate joint probability distribution. A more detailed introduction to the principles of copulas can be found in Salvadori & De Michele (2004), Nelsen (2006), and Zhang & Singh (2006). There are three main families of copulas: elliptical, Archimedean, and quadratic. The Archimedean copula family, which allows for greater flexibility and simplicity of use, is more suitable for hydrologic analyses (Zhang & Singh 2007a, 2007b). The most common 2-D or 3-D Archimedean copulas include the Clayton copula, the Gumbel-Hougaard (GH) copula, the Ali-Mikhail-Haq (AMH) copula, and the Frank copula (as shown in Table 2). The four forms of Archimedean copulas were all considered in this study.

Parameter estimation and goodness-of-fit testing

Determining a suitable marginal distribution for the study variables is the first step when using copulas. To fit DRD, PRD, and TRD, this study utilised three types of distribution functions that are commonly applied in hydrological frequency analyses: the P-III distribution (P3), the lognormal distribution (LN2), and the generalised extreme value distribution (GEV). Probability point correlation coefficient (PPCC), root-mean-squared error (RMSE), and mean absolute error (MAE) goodness-of-fit tests were used to determine the appropriate type of marginal distribution curve for each data series (Li & Chen 2010).

There are usually two methods for estimating the copula's parameters: nonparametric and parametric. The nonparametric method requires clear expressions of Kendall's τ and copula's θ , which is suitable for the estimation of a single parameter in 2-D copulas. In this study, the parameters of 2-D copulas were estimated using the empirical expressions between Kendall's τ and copula's θ (as shown in Table 2), while the parameters of 3-D copulas were determined by the inference of functions for margins (IFM; Favre et al. 2004; Xie et al. 2012). The IFM method was performed following the two-step process: (1) the estimation of parameters φ_1 , φ_2 , and φ_3 in marginal distribution by using the maximum likelihood (ML) method (Myung 2003), as shown in Equation (1), and (2) estimation of the 3-D copula's θ using the ML method as shown in Equation (2).

$$\begin{aligned}\hat{\varphi}_1 &= \arg \max \sum_{i=1}^n \ln f_{X_1}(x_{1i}; \varphi_1) \\ \hat{\varphi}_2 &= \arg \max \sum_{i=1}^n \ln f_{X_2}(x_{2i}; \varphi_2) \\ \hat{\varphi}_3 &= \arg \max \sum_{i=1}^n \ln f_{X_3}(x_{3i}; \varphi_3) \\ \hat{\theta} &= \arg \max \sum_{i=1}^n \ln c(F_1(x_{1i}, \hat{\varphi}_1), F_2(x_{2i}, \hat{\varphi}_2), F_3(x_{3i}, \hat{\varphi}_3); \theta)\end{aligned}\quad (1)$$

where $F_1(x_1)$, $F_2(x_2)$ and $F_3(x_3)$ represent the marginal distribution functions of X_1 , X_2 and X_3 , respectively; $f_{X_1}(x_1)$, $f_{X_2}(x_2)$ and $f_{X_3}(x_3)$ represent the probability density functions of X_1 , X_2 and X_3 , respectively.

After fitting the copula to rainfall data, it is necessary to identify the copula form with the highest goodness-of-fit.

Many proposals have been made recently for the goodness-of-fit testing of copulas (Zhang & Singh 2007b; Hou et al. 2010; Jiang et al. 2014). In this study, the Kolmogorov–Smirnov (K-S) test (Frank & Masse 1951) was first applied to check if the empirical and theoretical values could have uniform distribution. The agreement between the empirical and theoretical cumulative probabilities was then assessed. Finally, the evaluations of mean-square error (MSE; as shown in Equation (3)), Akaike information criterion (AIC; Akaike 1974; as shown in Equation (4)), and the ordinary least squares (OLS; Hou et al. 2010; as shown in Equation (5)) were performed to assess the goodness-of-fit of a copula function to the dataset.

$$MSE = \frac{1}{n} \sum_{i=1}^n (F_{emp}(x_{i1}, x_{i2}, \dots, x_{im}) - C(u_{i1}, u_{i2}, \dots, u_{im}))^2 \quad (3)$$

$$AIC = n \ln(MSE) + 2K \quad (4)$$

$$OLS = \sqrt{\frac{1}{n} \sum_{i=1}^n (F_{emp}(x_{i1}, x_{i2}, \dots, x_{im}) - C(u_{i1}, u_{i2}, \dots, u_{im}))^2} \quad (5)$$

where $F_{emp}(x_{i1}, x_{i2}, \dots, x_{im})$ and $C(u_{i1}, u_{i2}, \dots, u_{im})$ are empirical and theoretical cumulative probabilities, respectively; M is the function dimension and K is the number of model parameters.

Conditional risk probability

Given different variable conditions, the corresponding conditional probability distributions can be derived. This study considered four main types of conditional risk probability: (1) given $X_3 = x_3$, the conditional distribution function of X_1 and X_2 was expressed using Equation (6), (2) given $X_3 \leq x_3$, the conditional distribution function of X_1 and X_2 was expressed using Equation (7), (3) given $X_2 = x_2$ and $X_3 = x_3$, the conditional distribution function of X_1 was expressed using Equation (8), and (4) given $X_2 \leq x_2$ and $X_3 \leq x_3$, the conditional distribution function of X_1 was expressed using Equation (9).

$$\begin{aligned}F(x_1, x_2 | X_3 = x_3) &= \frac{\partial F(x_1, x_2, x_3) / \partial x_3}{f_{X_3}(x_3)} = C(U_1 \leq u_1, \\ U_2 \leq u_2 | U_3 = u_3) &= \frac{\partial C(u_1, u_2, u_3)}{\partial u_3}\end{aligned}\quad (6)$$

$$F(x_1, x_2 | X_3 \leq x_3) = \frac{F(x_1, x_2, x_3)}{F(x_3)} = \frac{C(u_1, u_2, u_3)}{u_3} \quad (7)$$

$$\begin{aligned} F(x_1 | X_2 = x_2, X_3 = x_3) &= \frac{\partial^2 F(x_1, x_2, x_3) / \partial x_2 \partial x_3}{f_{X_2 X_3}(x_2, x_3)} \\ &= C(U_1 \leq u_1 | U_2 = u_2, U_3 = u_3) \\ &= \frac{\partial^2 C(u_1, u_2, u_3) / \partial u_2 \partial u_3}{\partial^2 C(u_2, u_3) / \partial u_2 \partial u_3} \end{aligned} \quad (8)$$

$$F(x_1 | X_2 \leq x_2, X_3 \leq x_3) = \frac{F(x_1, x_2, x_3)}{F(x_2, x_3)} = \frac{C(u_1, u_2, u_3)}{C(u_2, u_3)} \quad (9)$$

where $u_1 = F_1(x_1), u_2 = F_2(x_2), u_3 = F_3(x_3)$; $F(x_1, x_2, x_3)$ represents the joint probability distribution function of X_1, X_2 and X_3 ; $F(x_2, x_3)$ represents the joint probability distribution function of X_2 and X_3 ; $f_{X_3}(x_3)$ represents

the probability density function of X_3 ; $f_{X_2 X_3}(x_2, x_3)$ represents the joint probability density function of X_2 and X_3 .

RESULTS

Determination of marginal distribution functions

Figure 2 shows the fitted curves of marginal distribution functions for DRD, PRD, and TRD. Generally, the chosen distribution curves show good agreement with the observations, despite slight differences among the fitted curves. Goodness-of-fit tests were conducted to further identify the most appropriate distribution function, and the results are shown in Table 3. As can be seen, for DRD, the LN2 distribution shows the smallest

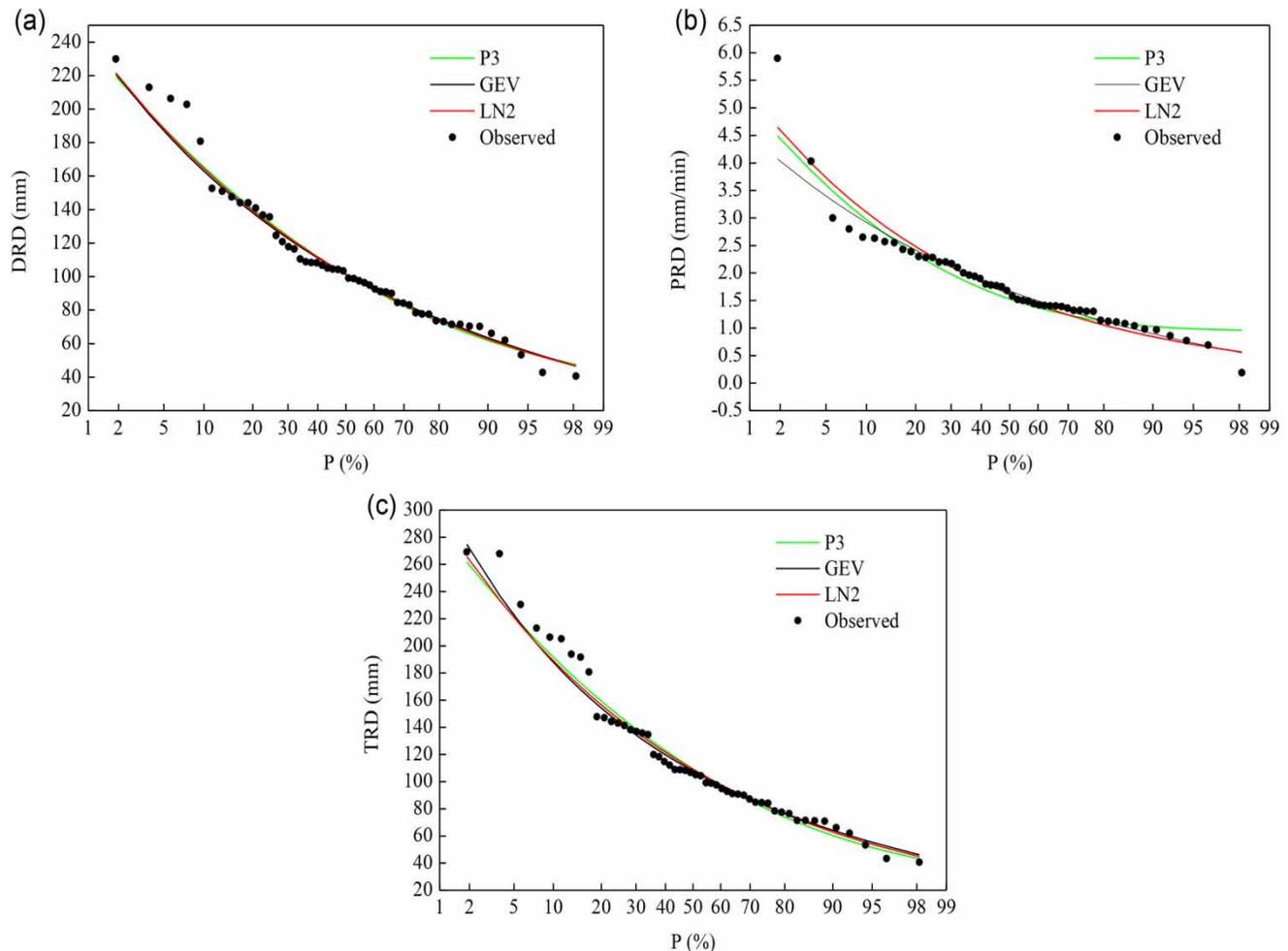


Figure 2 | Fit of univariate frequency curves for (a) DRD, (b) PRD, and (c) TRD.

Table 3 | Results of goodness-of-fit test of marginal distribution function

Index	Frequency distribution	RMSE	MAE	PPCC
DRD	P3	6.996	4.621	0.988
	GEV	7.095	4.208	0.990
	LN2	6.876	4.306	0.989
PRD	P3	0.273	0.162	0.955
	GEV	0.282	0.114	0.952
	LN2	0.250	0.140	0.960
TRD	P3	8.873	6.431	0.988
	GEV	8.786	5.582	0.990
	LN2	8.930	5.954	0.990

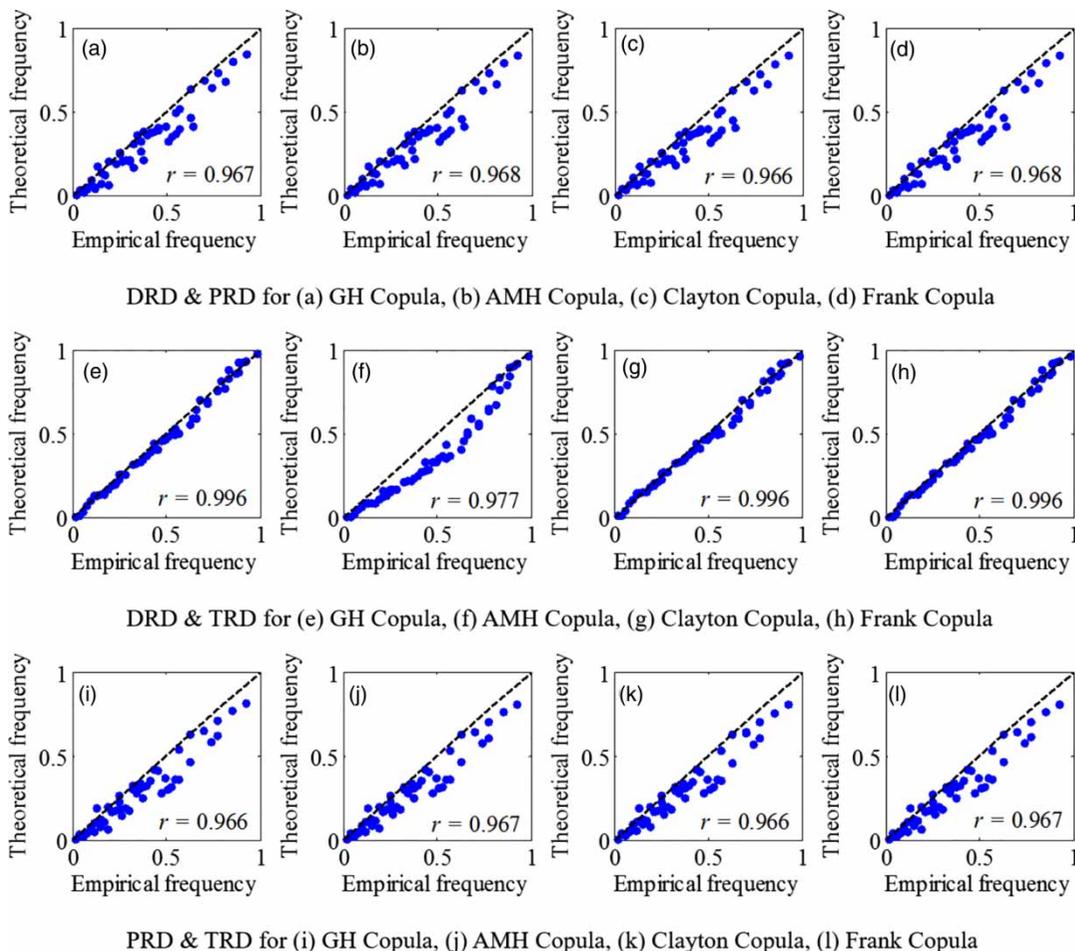
RMSE value, whereas the GEV distribution shows the smallest MAE value and the largest PPCC value. For PRD, the GEV distribution shows the lowest MAE value, whereas the LN2 distribution shows the smallest RMSE value and the largest PPCC value. For TRD, the

GEV distribution exhibits the lowest MAE and RMSE values and the largest PPCC value. On the basis of these results, the GEV, LN2, and GEV are chosen as the marginal distribution functions of DRD, PRD, and TRD, respectively.

Copula parameter estimation and goodness-of-fit testing

2-D copulas

The K-S tests were first used to evaluate the goodness-of-fit for the fitted distributions. The results show that all the tests for the 2-D copulas are significant at the 95% confidence level, indicating that the four Archimedean copulas can be applied for the bivariate rainfall data of the study region. Figure 3 shows the fit of empirical and theoretical frequencies of 2-D copulas for DRD & PRD, DRD &

**Figure 3** | Fit of empirical and theoretical frequencies of 2-D copulas.

TRD, and PRD & TRD. As can be seen, most copulas fit well the bivariate joint distributions of the rainfall characteristic variables, especially for DRD & TRD, with the correlation coefficients more than 0.97. For DRD & PRD and PRD & TRD, there are similar scatter patterns among the four Archimedean copulas, while for DRD & TRD, the simulation results of AMH copula are slightly worse than those of the other copulas. The results of goodness-of-fit test of 2-D copulas are shown in Table 4. It can be seen that the best-fitted joint distributions for DRD & PRD, DRD & TRD, and PRD & TRD all point to Frank copula, with the smallest values of MSE, AIC, and OLS. Therefore, the Frank copula is selected for further analyses of the bivariate joint distributions.

3-D copulas

The results of K-S test for 3-D copulas show that the tests for GH copula, Clayton copula and Frank copula are significant at the 95% confidence level, but AMH copula are not significant at the 95% confidence level. That is to say, AMH copula is not suitable for the trivariate rainfall data. Figure 4 displays the fit of empirical and theoretical frequencies of 3-D copulas for DRD, PRD and TRD. As shown, the theoretical frequencies of GH copula, Clayton copula and Frank copula are in good agreement with the empirical frequencies, with the correlation coefficients more than 0.98. In contrast, for AMH copula, the theoretical frequency tended to be lower than that of the empirical frequency. The goodness-of-fit test of 3-D copulas were further evaluated using MAE, OLS and AIC (Table 4). The results indicate the smallest MAE for GH copula and Frank copula, the smallest OLS for GH copula, Clayton copula and Frank copula, and the smallest AIC for GH copula. On the basis of these assessments, the GH copula is chosen as the trivariate joint distribution function of DRD, PRD, and TRD.

Conditional probability analysis

Univariate condition

Equations (6) and (7) were conducted to calculate the combinatorial risk probability of rainfall variables under the univariate condition that one of the variables (DRD, PRD, or TRD) is equal or not more than a given

Table 4 | Results of goodness-of-fit test of 2-D copulas and 3-D copulas

	MSE			AIC			OLS					
	GH	AMH	Clayton	Frank	GH	AMH	Clayton	Frank	GH	AMH	Clayton	Frank
DRD, PRD	0.00741	0.00739	0.00769	0.00728	-253.071	-253.173	-251.098	-254.006	0.086	0.086	0.088	0.085
DRD, TRD	0.00084	0.01319	0.00077	0.00075	-366.382	-223.083	-371.054	-371.944	0.029	0.115	0.028	0.027
PRD, TRD	0.00837	0.00838	0.00866	0.00831	-246.707	-246.667	-244.955	-247.073	0.092	0.092	0.093	0.091
DRD, PRD, TRD	0.00145	0.03509	0.00147	0.00145	-338.0423	-172.1847	-337.2247	-338.0384	0.038	0.187	0.038	0.038

DRD, design rainfall depth; PRD, peak rainfall depth; TRD, total rainfall depth.

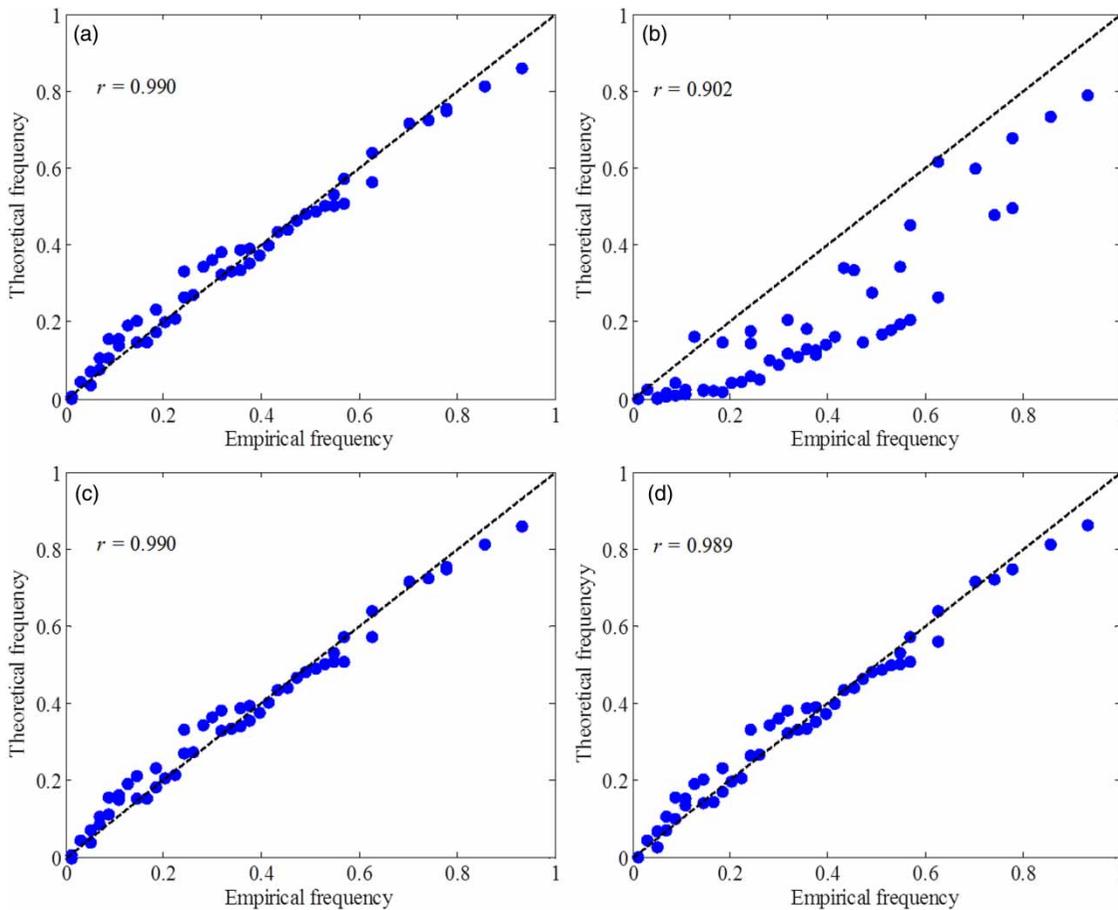


Figure 4 | Fit of empirical and theoretical frequencies of 3-D copulas for (a) GH copula, (b) AMH copula, (c) Clayton copula, and (d) Frank copula.

design value of a certain return period. The results under the condition of the 100-year return period are shown in Figure 5. Compared with a single variable equal to a given design value, the probability distribution under the condition with a single variable less than or equal to a given design value tends to be 'higher' or 'fatter'. This indicates that the risk probability is larger when the single variable is less than or equal to the design value.

Table 5 shows the joint distribution probability of PRD and TRD under the condition that the DRD is equal to a given design value (\sim less than or equal to a given design value). As can be seen, when the return period of DRD is fixed, the joint distribution probability of PRD and TRD shows a decreasing trend as their return periods decrease. In contrast, for the given design values of the PRD and TRD, their joint distribution probability shows an increasing trend when the return period of DRD decreases. Specifically, when DRD is equal to the 100-

year design value, the probability that PRD and TRD are simultaneously less than or equal to the 100-year design values is 69.26%. However, when the PRD and TRD are simultaneously less than or equal to the 10-year design values, the corresponding probability is reduced to 21.49%. Generally, the variation of the risk probability, under the condition of DRD equal to or less than a given design value, is similar to, but larger than that under the condition of DRD equal to a given design value. For example, under the condition with the DRD equal to or less than the 100-year design value, the probabilities that PRD and TRD are simultaneously equal to or less than the 100- and 10-year design values are 98.92% and 85.22%, respectively.

Bivariate condition

Equations (8) and (9) were conducted to calculate the risk probability for one of the variables under the

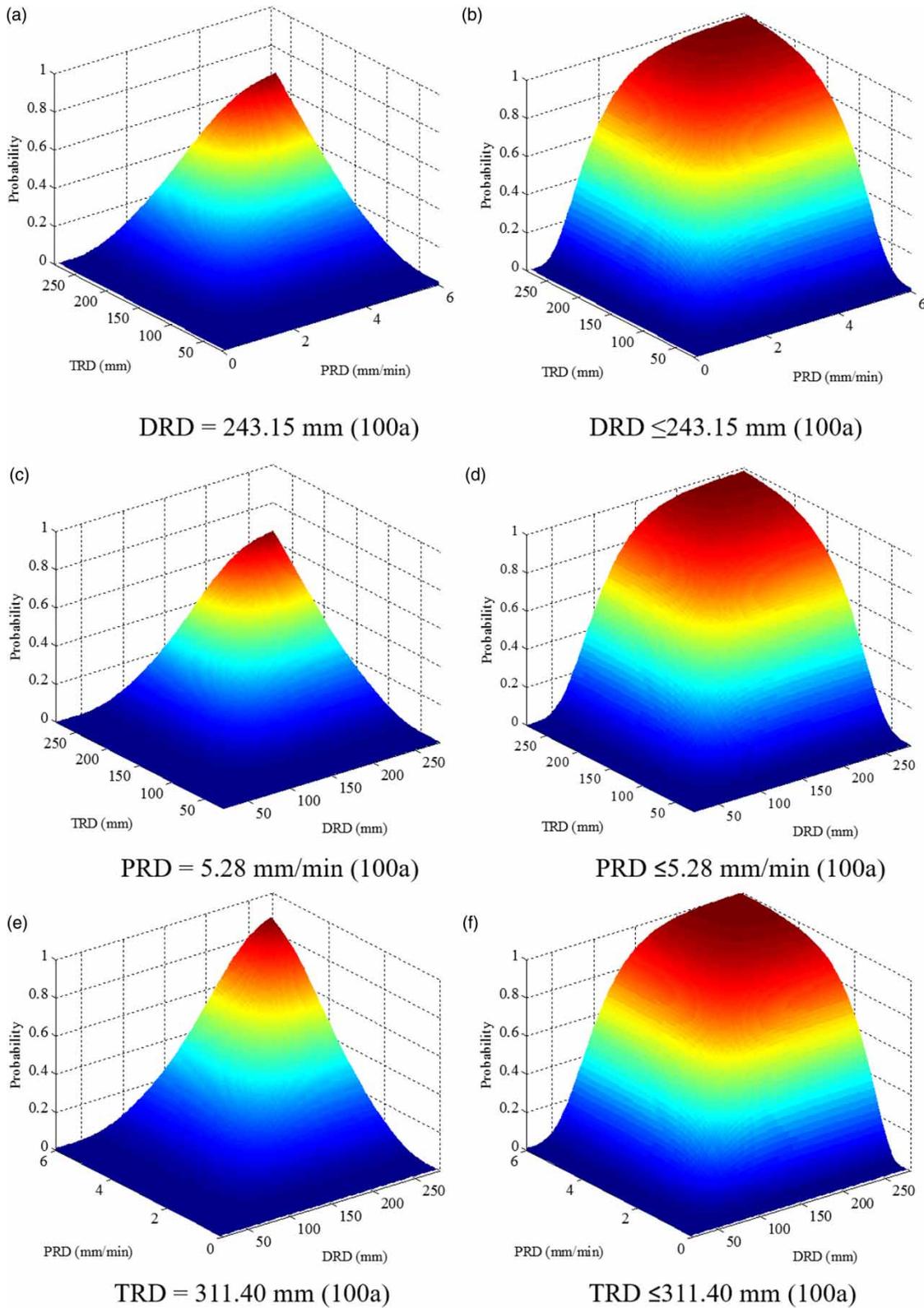


Figure 5 | Risk probability under the univariate condition.

Table 5 | The joint distribution probability of PRD and TRD under the condition that DRD is equal to a given design value (also less than or equal to a given design value)

DRD (mm)	Return period (a)	PRD (mm/min)	Return period (a)	TRD (mm)	Return period (a)	Conditional risk probability (%)
243.2	100	5.3	100	311.4	100	69.26 (98.92)
243.2	100	4.6	50	271.6	50	52.45 (97.45)
243.2	100	3.7	20	222.6	20	32.64 (92.72)
243.2	100	3.1	10	187.8	10	21.49 (85.22)
218.9	50	5.3	100	311.4	100	83.37 (99.14)
218.9	50	4.6	50	271.6	50	68.31 (97.82)
218.9	50	3.7	20	222.6	20	45.05 (93.26)
218.9	50	3.1	10	187.8	10	30.15 (85.82)
187.0	20	5.3	100	311.4	100	94.33 (99.42)
187.0	20	4.6	50	271.6	50	86.22 (98.41)
187.0	20	3.7	20	222.6	20	65.50 (94.42)
187.0	20	3.1	10	187.8	10	46.50 (87.30)
162.8	10	5.3	100	311.4	100	97.71 (99.58)
162.8	10	4.6	50	271.6	50	93.86 (98.83)
162.8	10	3.7	20	222.6	20	80.26 (95.56)
162.8	10	3.1	10	187.8	10	62.16 (89.09)

bivariate condition that the other two rainfall variables are equal to or not more than a given design value. Figure 6 shows the risk probability under the bivariate conditions. As shown in Figure 6(a) and 6(b), the probability of the DRD less than or equal to a certain design value, under the condition of PRD & TRD less than or equal to the given design values, is larger than that under the condition of PRD & TRD equal to the given design values. Furthermore, this probability tends to increase as the design value of TRD decreases. A similar phenomenon is found for the cases of the probability for PRD (Figure 6(c) and 6(d)) and TRD (Figure 6(e) and 6(f)).

Table 6 shows the probability for the TRD less than or equal to the 100-, 50-, 20-, and 10-years design values, by conditioning on the PRD and TRD equal to a given design value (also less than or equal to a given design value). As shown, when PRD = 5.3 mm/min and DRD = 243.2 mm, the probability for the TRD \leq 311.4 mm (100-year) is 58.20%, but for the TRD \leq 187.8 mm (10-year) drops to 2.37%. However, when PRD \leq 5.3 mm/min and DRD \leq 243.2 mm, the probability that TRD \leq 311.4 and 187.8 mm (100-, 10-years) can be up to 99.5% and 91.05%, respectively.

CONCLUSIONS

In this paper, a 3-D Archimedean copula-based multivariate method was used to develop the joint probability distribution model for the characteristic variables of rainfall. On the basis of the long-term (1961–2012) rainfall data, the model was applied to the conditional probability analyses of DRD, PRD and TRD in Guangzhou, China. The main findings of this study are as follows:

- (1) The DRD, PRD and TRD are positively correlated, and the correlation between DRD and TRD (PRD) is stronger than that between PRD and DRD (TRD). The GEV distribution is found to be the most appropriate for the marginal distribution of DRD and TRD, while the LN2 distribution is the most suitable for that of PRD.
- (2) Not all the copulas are appropriate for representation of multivariate probability distributions. Based on the goodness-of-fit tests, the Frank copula and GH copula are found to best fit the bivariate and trivariate rainfall probability distributions of rainfall data, respectively.
- (3) The probability under the single condition with a variable less than or equal to a given design value is larger than that under the condition of a variable equal to a given design value. For double conditioning, the probability under the condition of the two variables less than or equal to the given design values is larger than that under the condition of the two variables equal to the given design values.

The copula-based multivariate probability analysis presented in this study provides a new scientific reference for the urban flood control design. However, it should be noted that the parameter estimation and function selection for the 3-D and higher-dimensional copulas remain a challenge, and need to be further improved to reduce parameter uncertainty of copulas. In addition, when the estimation of the urban design flood is conducted using rainfall data, there is a hypothesis that floods and rainfall events are characterized by the same return period. However, in fact, the hydrological processes that cause floods are very complex in urban areas, where the floods can be produced by many phenomena. The wrong selection of the synthetic rainfall shape may show significant impacts on the urban flooding frequency (Fontanazza et al. 2011). Therefore, uncertainty evaluation of design

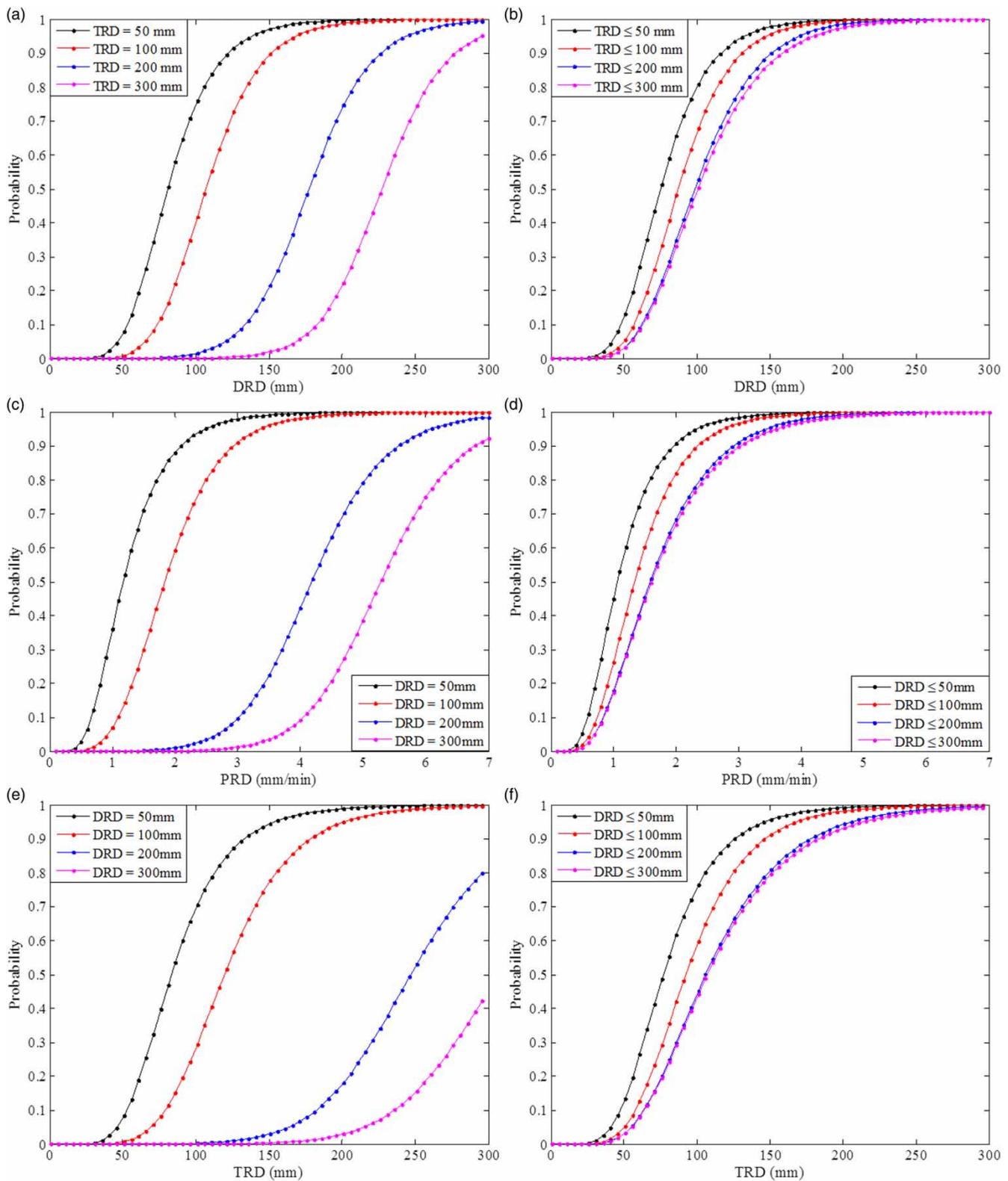


Figure 6 Risk probability under the bivariate conditions: (a) PRD = 5 mm/min and TRD = 50, 100, 200, 300 mm; (b) PRD ≤ 5 mm/min and TRD ≤ 50, 100, 200, 300 mm; (c) TRD = 300 mm and DRD = 50, 100, 200, 300 mm; (d) TRD ≤ 300 mm and DRD ≤ 50, 100, 200, 300 mm; (e) PRD = 5 mm/min and DRD = 50, 100, 200, 300 mm; (f) PRD ≤ 5 mm/min and DRD ≤ 50, 100, 200, 300 mm.

Table 6 | The probability for the TRD less than or equal to a certain design value under the condition that PRD and TRD are equal to a given design value (also less than or equal to a given design value)

PRD (mm/min)	Return period (a)	DRD (mm)	Return period (a)	TRD (mm)	Return period (a)	Conditional risk probability (%)
5.3	100	243.2	100	311.4	100	58.20 (99.50)
5.3	100	243.2	100	271.6	50	31.05 (98.71)
5.3	100	243.2	100	222.6	20	8.24 (95.95)
5.3	100	243.2	100	187.8	10	2.37 (91.05)
5.3	100	218.9	50	311.4	100	73.55 (99.58)
5.3	100	218.9	50	271.6	50	48.24 (98.89)
5.3	100	218.9	50	222.6	20	16.52 (96.30)
5.3	100	218.9	50	187.8	10	5.27 (91.51)
5.3	100	187.0	20	311.4	100	90.34 (99.71)
5.3	100	187.0	20	271.6	50	76.47 (99.19)
5.3	100	187.0	20	222.6	20	43.06 (97.06)
5.3	100	187.0	20	187.8	10	18.71 (92.67)
5.3	100	162.8	10	311.4	100	96.36 (99.79)
5.3	100	162.8	10	271.6	50	90.30 (99.41)
5.3	100	162.8	10	222.6	20	69.34 (97.74)
5.3	100	162.8	10	187.8	10	42.21 (93.98)

rainfall for urban flood risk analysis should be considered in future work.

ACKNOWLEDGEMENTS

This work was supported by the Fundamental Research Funds for the Central Universities (Grant No. 21617301) and partly supported by the Collaborative Research on Flood Resilience in Urban areas (CORFU) project of the European Union's Seventh Framework Programme.

REFERENCES

- Akaike, H. 1974 [A new look at the statistical model identification](#). *IEEE Transactions on Automatic Control* **19** (6), 716–722.
- Andreassian, V., Perrin, C., Michel, C., Usart-Sanchez, I. & Lavabre, J. 1996 [Impact of imperfect rainfall knowledge on the efficiency and the parameters of watershed models](#). *Journal of Hydrology* **250**, 206–223.
- Aronica, G. T., Freni, G. & Oliveri, E. 2005 [Uncertainty analysis of the influence of rainfall time resolution in the modelling of urban drainage systems](#). *Hydrological Processes* **19**, 1055–1071.

- Bacchi, B., Becciu, G. & Kottegoda, N. T. 1994 [Bivariate exponential model applied to intensities and durations of extreme rainfall](#). *Journal of Hydrology* **155**, 225–236.
- Favre, A. C., El Adlouni, S., Perreault, L., Thiémmonge, N. & Bobée, B. 2004 [Multivariate hydrological frequency analysis using copulas](#). *Water Resources Research* **40**, W01101.
- Fontanazza, C. M., Freni, G., La Loggia, G. & Notaro, V. 2011 [Uncertainty evaluation of design rainfall for urban flood risk analysis](#). *Water Science & Technology* **63** (11), 2641–2650.
- Frank, J. & Masse, J. 1951 [The Kolmogorov-Smirnov test for goodness of fit](#). *Journal of the American Statistical Association* **46** (253), 68–78.
- Grimaldi, S. & Serinaldi, F. 2006 [Design hyetograph analysis with 3-copula function](#). *Hydrological Sciences Journal* **51** (2), 223–238.
- Hashino, M. 1985 [Formulation of the joint return period of two hydrologic variates associated with a Poisson process](#). *Journal of Hydrosience and Hydraulic Engineering* **3** (2), 73–84.
- Hou, Y., Song, S., Zhao, L. & Wang, J. F. 2010 [Research of trivariate flood frequency based on Copula function](#). *Journal of Northwest A&F University (Nat. Sci. Ed.)* **38** (2), 219–228 (in Chinese).
- Jenkinson, A. F. 1955 [The frequency distribution of the annual maximum \(or minimum\) values of meteorological elements](#). *Quarterly Journal of the Royal Meteorological Society* **81** (348), 158–171.
- Jiang, C., Xiong, L., Xu, C. Y. & Guo, S. 2014 [Bivariate frequency analysis of nonstationary low-flow series based on the time-varying copula](#). *Hydrological Processes* **29** (6), 1521–1534.
- Lee, C. H., Kim, T. W., Chung, G., Choi, M. & Yoo, C. 2010 [Application of bivariate frequency analysis to the derivation of rainfall-frequency curves](#). *Stochastic Environmental Research and Risk Assessment* **24** (3), 389–397.
- Li, X. K. & Chen, Y. F. 2010 [Study on selection methods for the best frequency distribution of rainstorm](#). *Journal of China Hydrology* **30** (2), 50–53 (in Chinese).
- Myung, I. J. 2003 [Tutorial on maximum likelihood estimation](#). *Journal of Mathematical Psychology* **47**, 90–100.
- Nelsen, R. 2006 *An Introduction to Copulas*. Springer Verlag, New York, NY, USA.
- Notaro, V., Fontanazza, C. M., Freni, G. & Puleo, V. 2013 [Impact of rainfall data resolution in time and space on the urban flooding evaluation](#). *Water Science & Technology* **68** (9), 1984–1993.
- Powell, D. N., Aziz, N. M. & Khan, A. A. 2006 [Impact of new rainfall patterns on the design of hydraulic structures](#). In: *World Environmental and Water Resource Congress 2006: Examining the Confluence of Environmental and Water Concerns*, Omaha, NE, USA.
- Ren, B. Z., Long, T. R. & Shi, L. 2004 [Method to determine parameters in the rainstorm intensity formula](#). *Journal of Chongqing University (English Edition)* **3** (2), 30–33.
- Salvadori, G. & De Michele, C. 2004 [Frequency analysis via copulas: theoretical aspects and applications to hydrological events](#). *Water Resources Research* **40** (12), W12511.
- Shiau, J. T. 2006 [Fitting drought duration and severity with two-dimensional copulas](#). *Water Resources Management* **20** (5), 795–815.

- Singh, K. & Singh, V. P. 1991 Derivation of bivariate probability density functions with exponential marginals. *Stochastic Hydrology and Hydraulics* **5** (1), 55–68.
- Sklar, A. 1959 *Fonction de répartition à n dimensions et leurs marges*, Vol. 8. Publications de L'Institut de Statistique, Université de Paris, Paris, France, pp. 229–231.
- Wu, C. H., Huang, G. R. & Wu, S. Y. 2014 Risk analysis of combinations of short duration rainstorm and tidal level in Guangzhou based on Copula function. *Journal of Hydroelectric Engineering* **33** (2), 33–41 (in Chinese).
- Xie, H., Liu, Q. & Huang, J. S. 2012 Synchronous asynchronous encounter analysis of multiple hydrologic regions based on 3 D copula function. *Advances in Water Science* **23** (2), 186–193 (in Chinese).
- Yu, K. X., Xiong, L. & Gottschalk, L. 2014 Derivation of low flow distribution functions using copulas. *Journal of Hydrology* **508**, 273–288.
- Yue, S., Ouarda, T. B. M. J. & Bobee, B. 2001 A review of bivariate gamma distributions for hydrological application. *Journal of Hydrology* **246** (1–4), 1–18.
- Zhang, L. & Singh, V. P. 2006 Bivariate flood frequency analysis using the copula method. *Journal of Hydrologic Engineering* **11** (2), 150–164.
- Zhang, L. & Singh, V. P. 2007a Gumbel–Hougaard copula for trivariate rainfall frequency analysis. *Journal of Hydrologic Engineering* **12** (4), 409–419.
- Zhang, L. & Singh, V. P. 2007b Bivariate rainfall frequency distributions using Archimedean copulas. *Journal of Hydrology* **332** (1), 93–109.
- Zhang, N., Guo, S. L., Xiao, Y., Fang, B. & Yan, B. W. 2008 Design storm method based on bivariate joint distribution. *Water Power* **34** (1), 18–21.
- Zhang, Q., Li, J., Singh, V. P. & Xu, C. Y. 2013 Copula-based spatio-temporal patterns of precipitation extremes in China. *International Journal of Climatology* **33** (5), 1140–1152.

First received 29 March 2017; accepted in revised form 22 February 2018. Available online 7 March 2018