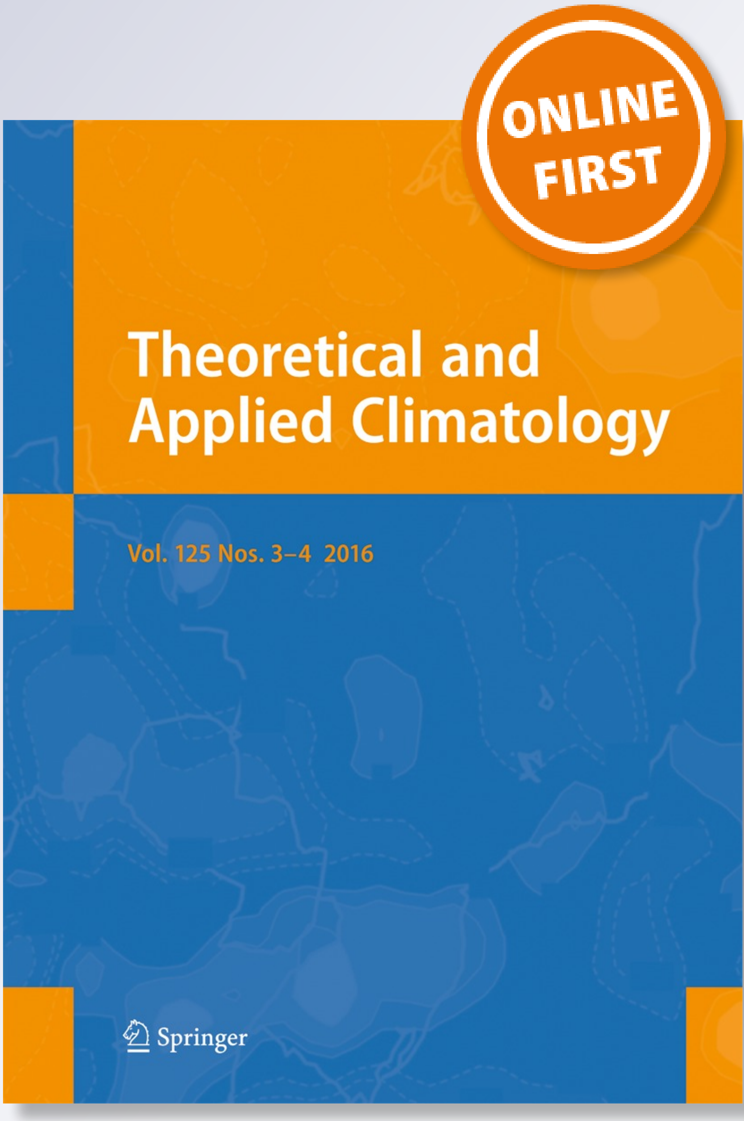
Modeling flood event characteristics using D-vine structures

Maryam Shafaei, Ahmad Fakheri-Fard, Yagob Dinpashoh, Rasoul Mirabbasi & Carlo De Michele

Theoretical and Applied Climatology

ISSN 0177-798X

Theor Appl Climatol
DOI 10.1007/s00704-016-1911-x



Springer

Theor Appl Climatol
DOI 10.1007/s00704-016-1911-x

ORIGINAL PAPER

Modeling flood event characteristics using D-vine structures

Maryam Shafaei[1] · Ahmad Fakheri-Fard[1] · Yagob Dinpashoh[1] · Rasoul Mirabbasi[2] · Carlo De Michele[3]

Received: 16 March 2016 / Accepted: 12 August 2016

Abstract The authors investigate the use of drawable (D-)vine structures to model the dependences existing among the main characteristics of a flood event, i.e., flood volume, flood peak, duration, and peak time. Firstly, different three- and four-dimensional probability distributions were built considering all the permutations of the conditioning variables. The Frank copula was used to model the dependence of each pair of variables. Then, the appropriate D-vine structures were selected using information criteria and a goodness-of-fit test. The influence of varying the data length on the selected D-vine structure was also investigated. Finally, flood event characteristics were simulated using the four-dimensional D-vine structure.

1 Introduction

Copula is a multivariate probability distribution with uniform marginal distributions (Salvadori et al. 2007). The copula allows investigating the dependence structure among the variables, independently by the probabilistic behavior of each variable. According to the Sklar theorem, any multivariate joint distribution can be written in terms of univariate marginal distributions and a copula (Sklar 1959). Early studies involving copulas in hydrology were performed by De Michele and Salvadori (2003), Salvadori and De Michele (2004a, b), and Favre et al. (2004).

✉ Maryam Shafaei
m.shafaei65@gmail.com

1 Water Engineering Department, Tabriz University, Tabriz, Iran

2 Water Engineering Department, Shahrekord University, Shahrekord, Iran

3 Department of Civil and Environmental Engineering, Politecnico di Milano, Piazza L. da Vinci 32, 20133 Milan, Italy

Since then, copulas have been widely applied in several hydrologic applications, including flood frequency analysis (De Michele et al. 2005; Zhang and Singh 2006; Karmakar and Simonovic 2009; Janga and Ganguli 2012a; Sraj et al. 2015), drought-related studies (Kao and Govindaraju 2008; Janga and Ganguli 2012b; Mirabbasi et al. 2012; De Michele et al. 2013; Mirabbasi et al. 2013; Salvadori and De Michele 2015), rainfall characteristic analysis (Salvadori and De Michele 2006; Grimaldi and Serinaldi 2006; Zhang and Singh 2006; Singh and Zhang 2007) and groundwater-related studies (Janga and Ganguli 2012c), the return period of multivariate events (Salvadori and De Michele 2004b, 2007, 2010; Salvadori et al. 2011 2013), and multivariate hazard scenarios and failure probabilities (Salvadori et al. 2016).

However, it must be noted that the use of copula function in multidimensional hydrologic problems (with a dimension $d \geq 3$) is quite limited. In particular, Serinaldi and Grimaldi (2007), Genest et al. (2007), and Ganguli and Reddy (2013) considered three variables in the analysis of flood events, Kao and Govindaraju (2008) and Salvadori and De Michele (2006) in rainfall investigations, and Song and Singh (2010) and Ma et al. (2013) in drought studies. Applications of four-dimensional copulas in hydrology are rare, and we can refer only to De Michele et al. (2007), where a four-dimensional copula has been for the analysis of sea storm events. Serinaldi and Grimaldi (2007) applied a fully nested asymmetric Archimedean copula for a three-dimensional frequency analysis. However, the flexibility of this method is progressively reduced with the increasing of the dimension. Genest et al. (2007) presented meta-elliptic copulas for modeling three-dimensional hydrologic data. The use of copulas, in higher dimensions than two, gives a heavy computational burden. The vine construction method does not include any of the mentioned drawbacks.

Published online: 29 August 2016

A vine tree structure is a method to build high-dimensional probability distributions, based on the mixing of conditional two-dimensional distributions, which represent the building blocks of multivariate distributions. This method has been introduced by Joe (1997). Bedford and Cooke (2001, 2002) developed the technique of multivariate density decomposition and represented the vine method graphically. If the distributions are copulas, then the vine is referred as vine copula. This last is becoming a popular model in finance (see, e.g., Nikololoupoulos et al. (2012); Zhang (2014)), as well as in geophysics and hydrology (see, e.g., Gyasi-Agyei and Melching (2012); Gräler et al. (2013); Xiong et al. (2014)). Here, three- and four-dimensional multivariate distributions of flood variables (namely, flood volume V, flood peak P, duration D, and peak time T_P) are constructed using D-vines, considering data collected at Ripetta station (Rome) on Tiber river basin (central Italy). In this study, D-vine structures are constructed for different lengths of data sample (namely, first 30, 60, 120, 200, and 300, i.e., all the sample), to investigate if data length can affect the selection of the D-vine structure. Three-dimensional D-vine structures are considered to represent the dependence structure among the flood volume, flood peak, and the duration and among flood volume, flood peak, and time peak while the four-dimensional D-vine structures to model the link among flood volume, flood peak, duration, and also peak time. In the "Materials and methods" section, the authors illustrate the methodological issues, including vine and D-vine structures (vine structure), how to estimate the parameters, and how to simulate samples from dependent variables. In the "Application" section, the authors present the case study, where the most appropriate D-vines are selected using information criteria, and in addition, a goodness-of-fit test is applied. A simulation of four-dimensional D-vine structures of flood characteristics is illustrated. In the last section, the conclusions are presented.

2 Materials and methods

2.1 Copulas and vine structures

Firstly, we recall here the Sklar theorem (Sklar 1959), because it allows formulating the vine structure in terms of copulas. In d dimensions, the joint cumulative distribution function (CDF), $F_{12\ldots d}(.)$, of a random vector (X_1, X_2, … X_d), having univariate continuous marginals $F_i(x_i)$ can be written as:

$$F_{12\ldots d}(x_1, x_2, \ldots, x_d) = C_{12\ldots d}\{F_1(x_1), F_2(x_2), \ldots, F_d(x_d), \theta\} \tag{1}$$

where $C_{12\ldots d}(.)$ is a d-dimensional copula that maps the marginal uniform distributions into their multivariate CDF, defined in $[0, 1]^d$ with values in [0, 1]. θ is the set of parameters. The inverse of the Sklar theorem has the expression:

$$C_{12\ldots d}(u_1, u_2, \ldots, u_d) = F_{12\ldots d}\{F_1^{-1}(u_1), F_2^{-1}(u_2), .., F_d^{-1}(u_d)\} \tag{2}$$

where $F_i^{-1}(u_i)$'s are the inverse of marginal distributions of X_i's (Salvadori et al. 2007).

For a dimension $d \geq 3$, the multivariate probability distribution has generally been built using symmetric Archimedean copulas, asymmetric Archimedean copulas (Serinaldi and Grimaldi 2007), or multivariate elliptical copulas. Nevertheless, the mentioned methods contain some restrictions:

1. In multivariate symmetric Archimedean copulas, the dependence among all pairs of variables is identical.
2. In multivariate asymmetric Archimedean copulas, narrow range of negative dependence is allowed as the dimension increases (Joe 1997).

To overcome these restrictions, the dependence among the variables will be modeled using vine structures (Kurowicka and Cooke 2007; Aas et al. 2009). The vine structure on d variables is defined as a nested set of trees $\{T_j, j = 1,\ldots,d-1\}$, built one on top of another, so that the edges of the tree at a certain level j become the nodes of the tree at the level $j+1$. The nodes at the level 1 (T_1) of the vine correspond to the random variables. The vines can be distinguished in regular and non-regular (Kurowicka and Cooke 2007). The vine is regular if two edges in tree j are joined by an edge in tree $j+1$ only if these edges share a common node. The regular vines include drawable (D-) vines and canonical (C-) vines. A regular vine is called D-vine if each node in T_1 has a degree of at most 2. A regular vine is called C-vine if each tree T_i has a unique node of degree $d-$i. In the next section, we will focus the attention on D-vine structures. The joint density function $f(x_1,\ldots, x_d)$ of a vector $X = (X_1,\ldots, X_d)$ of random variables can be written as:

$$f(x_1, x_2, \ldots, x_d) = f(x_d)\cdot f(x_{d-1}|x_d)\cdot f(x_{d-2}|x_{d-1}, x_d)\cdots f(x_1|x_2, \ldots, x_d) \tag{3}$$

Equation (3) indicates that the joint distribution function implicitly includes both the dependency structure and the probability distributions of variables. According to Aas et al. (2009), each term of the right-hand side of Eq. (3) can be written in terms of copulas as:

$$f(x|\vartheta) = C_{x\vartheta_j|\vartheta_{-j}}\{F(x|\vartheta_{-j}), F(\vartheta_j|\vartheta_{-j})\}\cdot f(x|\vartheta_{-j}) \tag{4}$$

where ϑ is a p-dimensional conditioning vector; ϑ_j corresponds to the arbitrarily chosen component from the vector ϑ; and ϑ_{-j} denotes the $(p-1)$ dimensional vector, where it is not included ϑ_j. In fact, the conditional probability distribution function in Eq. (3) can be decomposed into the suitable pair-copula $\left(c_{x\vartheta_j|\boldsymbol{\vartheta}_{-j}}\right)$ times a conditional marginal density, using Eq. (4). The pair-copula structure contains as argument conditional distributions like $F(x|\boldsymbol{\vartheta})$. Joe (1997) demonstrated that for each j:

$$F(x|\boldsymbol{\vartheta}) = \frac{\partial C_{x\vartheta_j|\boldsymbol{\vartheta}_{-j}}\left\{\left(F\left(x|\boldsymbol{\vartheta}_{-j}\right), \quad F\left(\vartheta_j|\boldsymbol{\vartheta}_{-j}\right)\right)\right\}}{\partial F\left(\vartheta_j|\boldsymbol{\vartheta}_{-j}\right)} \tag{5}$$

where $C_{x\vartheta_j|\boldsymbol{\vartheta}_{-j}}$ is a bivariate copula function. In cases where $\boldsymbol{\vartheta}$ is univariate, Eq. (5) may be written:

$$F(x|\vartheta) = \frac{\partial C_{x\vartheta}\{F(x), F(\vartheta), \theta\}}{\partial F(\vartheta)} \tag{6}$$

Figure 1 shows four-dimensional and three-dimensional D-vine structures. The d-dimensional D-vine structure contains $d-1$ trees: T_j, $j = 1, ..., d-1$. Tree T_j has $(d+1-j)$ nodes (Xs in the first tree and Fs in the other trees) and $(d-j)$ edges (Cs). Each edge denotes a pair-copula (C), and the edge label indicates the index of pair-copula, e.g., edge of 13|2 corresponds to the pair-copula $C_{13|2}$. In d-dimensional D-vine structure, $d(d-1)/2$ pair-copulas, or edges, are considered.

2.2 Parameters' estimation of D-vine copula

Here, we refer to the cases $d = 3$ and $d = 4$, and to fix the ideas, we consider the structures reported in Fig. 1. Estimates of D-vine copula parameters can be obtained using the maximum likelihood (ML) method, in the following way:

1. Calculate using ML the parameters (θ_{12} and θ_{23} in the three-dimensional case and θ_{12}, θ_{23}, and θ_{34} in the four-dimensional case) in T_1 of the two-dimensional copulas (C_{12} and C_{23} for $d = 3$ and C_{12}, C_{23}, and C_{34} for $d = 4$).
2. Determine the conditional distribution functions (v_{12}, v_{32} for $d = 3$ and v_{12}, v_{32}, and v_{43} for $d = 4$) of T_2 using the obtained copula parameters from T_1 and the Eq. (6) as follows:

$$v_{12} = F_{2|1}(x_2|x_1) = \frac{\partial C_{12}(u_1, u_2, \theta_{12})}{\partial u_1} \tag{7}$$

$$v_{32} = F_{3|2}(x_3|x_2) = \frac{\partial C_{32}(u_3, u_2, \theta_{23})}{\partial u_2} \tag{8}$$

$$v_{43} = F_{4|3}(x_4|x_3) = \frac{\partial C_{43}(u_4, u_3, \theta_{34})}{\partial u_3} \tag{9}$$

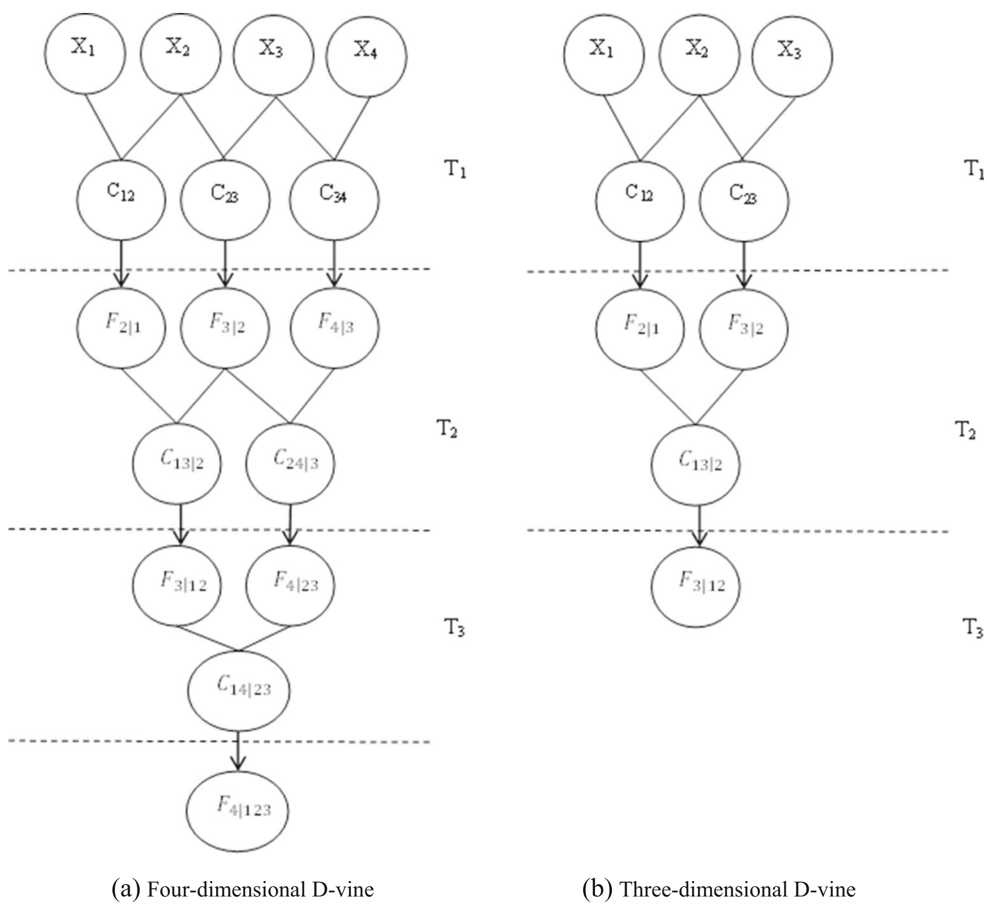


Fig. 1 Scheme of four- and three-dimensional D-vine structures

Table 1 The daily flow statistics of the dataset collected at Ripetta station

Data period	X_{min} (m^3/s)	X_{max} (m^3/s)	X_{mean} (m^3/s)	St. dev. (m^3/s)	Skewness
1921–2000	24.1	2730	217.52	164.05	3.72

3. Compute using the ML parameters ($\theta_{13|2}$ for $d = 3$ and $\theta_{13|2}, \theta_{24|3}$ for $d = 4$) in T_2 of the two-dimensional copulas ($C_{13|2}$ and $C_{13|2}, C_{24|3}$) having conditional distributions v_{12}, v_{32}, or v_{12}, v_{32}, and v_{43} from point 2.
4. Determine the conditional distribution functions $v_{3|12}$ and $v_{4|23}$ for T_3 using the computed copula parameters at T_2 and the Eq. (5) as follows:

$$v_{3|12} = F_{3|12}(x_3|x_1, x_2) = \frac{\partial C_{13|2}\left(F_{3|1}(x_3|x_1), \quad F_{3|2}\left(x_3\middle|x_2\right), \theta_{13|2}\right)}{\partial F_{3|2}\left(x_3\middle|x_2\right)} \tag{10}$$

$$v_{4|23} = F_{4|23}(x_4|x_2, x_3) = \frac{\partial C_{24|3}\left(F_{4|3}(x_4|x_3), F_{4|2}(x_4|x_2), \theta_{24|3}\right)}{\partial F_{4|2}(x_4|x_2)} \tag{11}$$

5. Compute using the ML parameters ($\theta_{14|23}$ for $d = 4$) in T_3 of the two-dimensional copulas $C_{14|32}$,having conditional distributions using $v_{3|12}$ and $v_{4|23}$ from point 4.

2.3 Simulation of random variables from vine structures

The simulation algorithm of samples of dependent variables, considered in the next, is based on the theory of conditional copulas (De Michele et al. 2007). To fix the ideas, we consider the case $d = 4$. Consider the conditional cumulative distribution function:

$$H_{4|123}(u_4|u_1, u_2, u_3) = F(U_4 \le u_4 | U_1 = u_1, U_2 = u_2, U_3 = u_3) = \frac{\frac{\partial^3}{\partial U_1 \partial U_2 \partial U_3} C_{1234}(u_1, u_2, u_3, u_4)}{\frac{\partial^3}{\partial U_1 \partial U_2 \partial U_3} C_{123}(u_1, u_2, u_3)} \tag{12}$$

Table 2 Some statistics of flood characteristics: flood peak (P), volume (V), peak time (T_p), and base time (D)

	P (m^3/s)	D (days)	V (m^3)	T_p (days)
Mean	498.86	24.46	371,518,807	9.04
Max	2512	91	3.00E + 09	72
Min	24.5	24.50	10,152,000	1
St. dev.	401.26	20.48	469,833,735	11.59
Skewness	1.31	1.50	2.49	2.91

where U_1, U_2, U_3, and U_4 are uniform in [0, 1]. In Eq. (12), the numerator of the right side is the mixed partial derivative of the four-dimensional copula with respect to the conditioning variables U_1, U_2, and U_3. The denominator is the density of the three-dimensional copula of the conditioning variables. For simulating a random sample (u_1, u_2, u_3, u_4) from Eq. (12), a random sample was (w_1, w_2, w_3, w_4) from uniform in [0,1] and independent from each other (W_1, W_2, W_3, W_4). Then, we have that $u_1 = w_1$, $u_2 = H^{-1}_{2|1}(w_2|u_{1.})$

where

$$H_{2|1}(w_2|u_1) = \frac{\partial}{\partial U_1} C_{12}(u_1, u_2), \quad u_3 = H^{-1}_{3|12}(w_3|u_1, u_2)$$

where

$$H_{3|12}(w_3|u_1, u_2) = \frac{\frac{\partial^2}{\partial U_1 \partial U_2} C_{123}\left(u_1, u_2, u_3\right)}{\frac{\partial^2}{\partial U_1 \partial U_2} C_{12}\left(u_1, u_2\right)}, \quad u_4$$

$$= H^{-1}_{4|123}(w_4|u_1, u_2, u_3)$$

where

$$H_{4|123}(w_4|u_1, u_2, u_3) = \frac{\frac{\partial^3}{\partial U_1 \partial U_2 \partial U_3} C_{1234}(u_1, u_2, u_3, u_4)}{\frac{\partial^3}{\partial U_1 \partial U_2 \partial U_3} C_{123}(u_1, u_2, u_3)}.$$

Essential for the calculation of $H_{4|123}$ is:

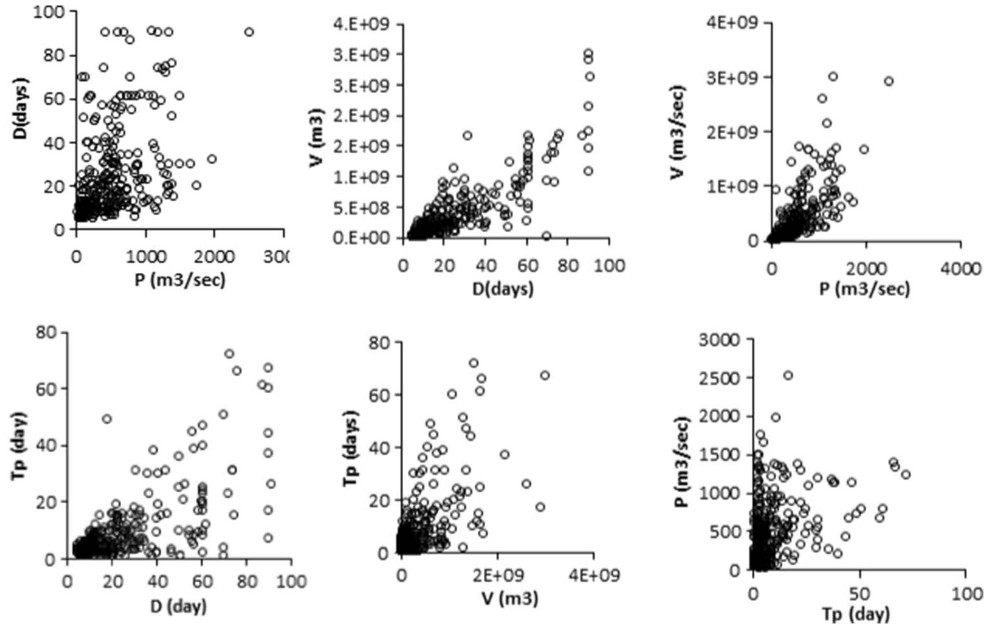
$$\frac{\partial^3}{\partial U_1 \partial U_2 \partial U_3} C_{1234}(u_1, u_2, u_3, u_4) = \frac{\partial}{\partial U_1} C_{14|23}\left(H_{1|23}(u_1|u_2, u_3), H_{4|23}(u_4|u_2, u_3)\right), \tag{13}$$

with

$$H_{1|23}(u_1|u_2, u_3) = \frac{\frac{\partial^2}{\partial U_2 \partial U_3} C_{123}(u_1, u_2, u_3)}{\frac{\partial^2}{\partial U_2 \partial U_3} C_{23}(u_2, u_3)} = \frac{\frac{\partial}{\partial U_3} C_{13|2}\left(\frac{\partial}{\partial U_2} C_{12}\ (u_1, u_2), \frac{\partial}{\partial U_2} C_{23}\ (u_2, u_3)\right)}{\frac{\partial^2}{\partial U_2 \partial U_3} C_{23}\ (u_2, u_3)} \tag{14}$$

and

Fig. 2 Pair observations of flood event characteristics



Once u_1, u_2, u_3, and u_4 are calculated, the corresponding values of T_P, D, V, and P are computed using the inverse of marginal cumulative distribution functions $F_{T_P}^{-1}$, F_D^{-1}, F_V^{-1}, and F_P^{-1}, respectively. In the present study, the different D-vine structures were built permuting the conditioning variables. Then, in order to select the appropriate structures, we used information criteria (AIC, BIC, and log-likelihood) and the probability integral transform (PIT) goodness-of-fit (GOF) test (Aas et al. 2009). For more details about PIT, readers can refer to Aas et al. (2009).

$$H_{4|23}(U_4|U_2,U_3) = \frac{\frac{\partial^2}{\partial U_2 \partial U_3} C_{234}(u_2,u_3,u_4)}{\frac{\partial^2}{\partial U_2 \partial U_3} C_{23}(u_2,u_3)} = \frac{\frac{\partial}{\partial U_3} C_{24|3}\left(\frac{\partial}{\partial U_3} C_{23}\,(u_2,u_3), \frac{\partial}{\partial U_3} C_{34}\,(u_3,u_4)\right)}{\frac{\partial^2}{\partial U_2 \partial U_3} C_{23}\,(u_2,u_3)} \tag{15}$$

Table 3 Pair-wise dependence between the flood characteristics: flood peak (P), peak time (T_p), volume (V), and base time (D)

Dependence measure	Kendall's tau
D, T_p	0.4
D, V	0.67
D, P	0.37
T_p, V	0.38
T_p, P	0.22
V, P	0.66

3 Application

3.1 Case study description

Here we consider daily flow data, collected at Ripetta station (41° 54′ 20″ N, 12° 28′ 32″ E), located in Rome, on Tiber river basin (central Italy). Data are provided by the Regional Agency of Environmental Protection of Lazio region. The data cover a period of 80 years (29,200 days) from January 1921 to March 2000. The dataset has only 2 months missing: August and September 1992. The daily statistics of the dataset are presented in Table 1.

3.2 Selection of flood event characteristics

Flood events were identified with the following characteristics: volume (V), duration (D), flood peak (P), and peak time (T_P). The identification of a flood event requires the determination of its start and end. These are generally determined by the following:

Table 4 Agreement between the non-parametric distribution and the cumulative frequency for each of the four variables via the evaluation criteria and the p value

	R	RMSE	MAE	p value
D	0.996	0.046	0.023	0.97
P	0.997	0.056	0.021	0.96
T	0.998	0.049	0.019	0.97
V	0.994	0.056	0.025	0.98

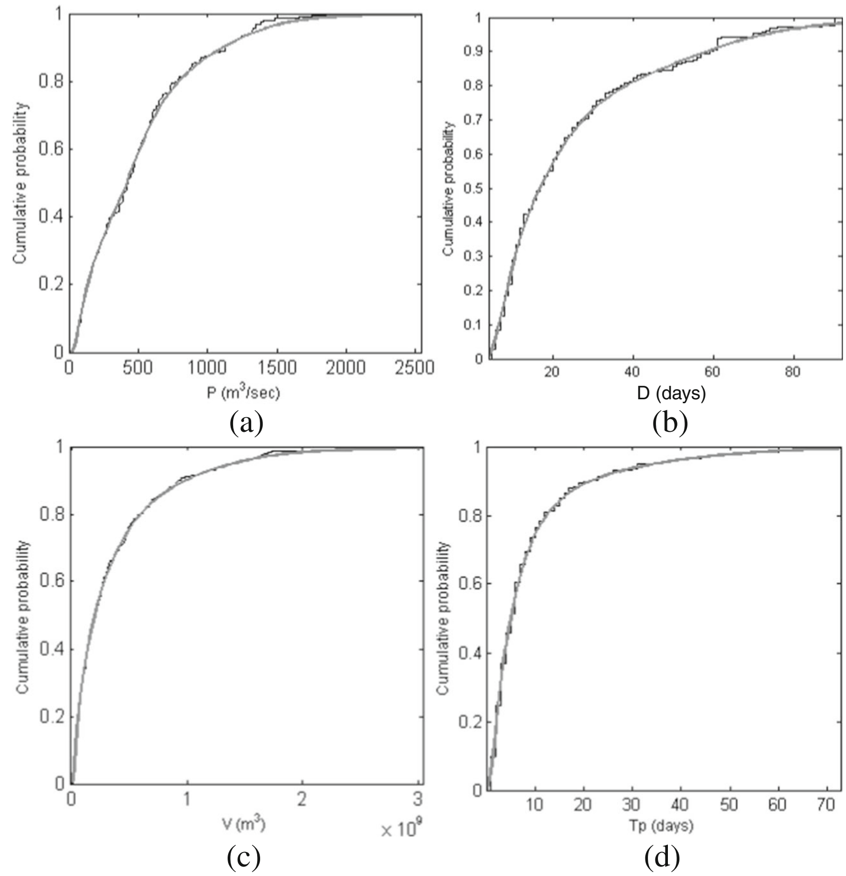


Fig. 3 Comparison between empirical and theoretical univariate distributions, for the peak (P) in **a**, flood duration (D) in **b**, volume (V) in **c**, and peak time (T_p) in **d**. *Continuous lines* represent the theoretical distributions while *broken lines* the empirical distributions

1. Fixing a threshold on the discharge and considering the above-the-threshold part as flood event (see, e.g., Ashkar and Rousselle 1982; Grimaldi and Serinaldi 2006; Todorovic and Zelenhasic 1970)
2. Detecting the time boundaries signed by a rise of discharge from the base flow (the start of flood runoff), and a return to the base flow, indicating the end of the flood runoff (see, e.g., Linsley et al. 1975; Yue et al. 1999); In this study, the flood event is specified using a threshold on the discharge, the mean annual discharge, 217.52 m^3/s. Thus, 300 events during the period 1921–2000 were selected. The time series of the four variables are not

Table 5 D-vine structures for the variables P, V, D, and T_p, using the first 30 values of sample data

$n = 30$									
1-2-3-4	θ_{12}	θ_{23}	θ_{34}	$\theta_{13\|2}$	$\theta_{24\|3}$	$\theta_{14\|23}$	Log-Lik	AIC	BIC
V-P-D-T	8.18	1.89	4.37	9.94	−0.70	−0.64	42.28	−72.56	−64.15
V-D-T-B	**8.18**	**0.27**	**4.37**	**2.77**	**1.77**	**7.92**	**42.67**	**−73.34**	**−65.01**
V-D-T-P	6.13	4.37	0.27	−1.00	1.70	9.94	36.25	−60.50	−52.90
V-D-P-T	6.13	1.89	0.27	11.9	4.06	−0.60	37.63	−63.27	−54.86
V-T-P-D	1.73	0.27	1.89	11.2	4.06	7.92	41.26	−70.53	−62.13
V-T-D-P	1.79	4.37	1.89	4.36	−0.87	9.78	33.30	−58.60	−50.20
P-V-T-D	8.18	1.73	4.37	−1.54	4.36	−4.06	38.12	−64.24	−55.83
P-V-D-T	8.18	6.13	4.37	−3.70	−1.00	0.81	37.13	−62.27	−53.86
P-D-V-T	1.89	6.13	1.73	11.9	4.38	−0.19	37.33	−62.66	−54.25
P-T-V-D	0.27	1.73	6.13	11.25	4.38	−4.00	39.15	−65.47	−57.06
D-P-V-T	1.89	8.18	1.73	9.90	−1.50	3.03	42.14	−72.28	−63.88
D-V-P-T	6.13	8.18	0.27	−3.70	2.77	3.14	37.6	−63.25	−54.84

For each structure, the parameter value (θ) of each pair-wise copula involved, maximum value of log-likelihood, AIC, and BIC is reported. Bold indicated the best structure

Table 6 D-vine structures for the variables P, V, D, and T_p, using the first 60 values of sample data

n = 60									
1-2-3-4	θ_{12}	θ_{23}	θ_{34}	$\theta_{13\|2}$	$\theta_{24\|3}$	$\theta_{14\|23}$	Log-Lik	AIC	BIC
V-P-D-T	9.27	2.92	4.89	6.85	−0.66	0.13	82.24	−152.49	−139.92
V-P-T-D	**9.27**	**1.29**	**4.95**	**3.04**	**2.50**	**5.50**	**84.05**	**−156.11**	**−143.54**
V-D-T-P	6.35	4.89	1.29	−0.23	2.47	9.62	77.70	−143.40	−130.83
V-D-P-T	6.35	2.92	1.29	10.42	3.91	0.12	78.05	−144.11	−131.54
V-T-P-D	2.98	1.29	2.92	10.24	3.91	5.33	81.20	−150.40	−137.83
V-T-D-P	2.98	4.95	2.92	5.07	−0.66	9.30	77.93	−143.87	−131.31
P-V-T-D	9.27	2.98	4.95	−1.96	5.07	−2.70	82.06	−152.12	−139.55
P-V-D-T	9.27	6.35	4.89	−3.56	−0.23	−0.54	79.20	−146.40	−133.84
P-D-V-T	2.92	6.35	1.98	10.42	3.41	−0.86	77.62	−143.37	−130.81
P-T-V-D	1.29	2.98	6.35	10.24	3.41	−3.11	80.46	−148.92	−136.35
D-P-V-T	2.92	9.27	2.98	6.85	−1.96	2.27	81.58	−151.16	−138.60
D-V-P-T	6.35	9.27	1.29	−3.50	3.04	2.42	80.17	−148.34	−135.78

For each structure, the parameter value (θ) of each pair-wise copula involved, maximum value of log-likelihood, AIC, and BIC, is reported. Bold indicated the best structure

autocorrelated with a p value of 5 % T_P, for and 1 % for V, P, and D. Table 2 reports some statistics of the flood event variables. Figure 2. shows the pair observations: (P,V), (P,T_P), (P,D), (D,V), (V,T_P), and (D,T_P).

3.3 Dependence among flood characteristics

Here, the Kendall's tau (τ) is used to investigate the statistical dependence between each pair of variables. Table 3 gives the value of τ between each couple of flood characteristics, with also the corresponding p value. Table 3 suggests that the pair dependences peak-volume, peak time-duration, volume-peak time, flood peak-peak time, flood peak-duration, and volume-duration are all statistically significant at a 5 % level.

3.4 Marginal distribution of flood characteristics

Univariate frequency analysis of flood characteristics is made using non-parametric probability distribution functions. Here, non-parametric distributions with type Normal kernel are used, having a probability density function (PDF):

$$f(x) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x-x_i}{h}\right) \quad (16)$$

Table 7 D-vine structures for the variables P, V, D, and T_p, using the first 120 values of sample data

n = 120									
1-2-3-4	θ_{12}	θ_{23}	θ_{34}	$\theta_{13\|2}$	$\theta_{24\|3}$	$\theta_{14\|23}$	Log-Lik	AIC	BIC
V-P-D-T	8.66	2.775	4.07	10.54	0.52	0.93	179.72	−347.44	−330.72
V-P-T-D	**8.66**	**2.52**	**4.14**	**3.98**	**1.94**	**8.48**	**184.94**	**−357.80**	**−341.66**
V-D-T-P	7.89	4.07	2.52	1.14	1.94	10.12	172.42	−332.84	−316.12
V-D-P-T	7.89	2.75	2.52	10.47	3.61	1.38	178.81	−345.63	−328.91
V-T-P-D	4.20	2.52	2.75	7.70	3.61	7.76	173.93	−335.87	−319.14
V-T-D-P	4.20	4.14	2.75	6.42	0.53	9.67	171.78	−331.75	−315.02
P-V-T-D	8.66	4.20	4.14	−1.69	6.42	−4.29	178.59	−345.19	−328.47
P-V-D-T	8.66	7.89	4.07	−5.07	1.14	−0.34	174.92	−337.84	−321.12
P-D-V-T	2.75	7.89	4.20	10.47	1.99	−0.71	177.84	−343.68	−326.96
P-T-V-D	2.52	4.20	7.89	7.70	1.99	−4.59	174.76	−337.5	−320.80
D-P-V-T	2.75	8.66	4.20	10.54	−1.69	0.73	181.98	−351.95	−335.23
D-V-P-T	7.89	8.66	2.52	−5.07	3.98	0.96	180.49	−348.98	−332.26

For each structure, the parameter value (θ) of each pair-wise copula involved, maximum value of log-likelihood, AIC, and BIC, is reported. Bold indicated the best structure

Table 8 D-vine structures for the variables P, V, B, and T_p, using the first 200 values of sample data

$n = 200$									
1-2-3-4	θ_{12}	θ_{23}	θ_{34}	$\theta_{13\|2}$	$\theta_{24\|3}$	$\theta_{14\|23}$	Log-Lik	AIC	BIC
V-P-D-T	8.92	3.23	4.60	11.37	0.21	0.80	320.32	−608.86	−628.65
V-P-T-D	**8.92**	**2.41**	**4.60**	**4.20**	**2.42**	**9.52**	**334.94**	**−657.88**	**−638.04**
V-D-T-P	9.57	4.60	2.41	0.89	2.42	9.78	313.46	−614.92	−595.14
V-D-P-T	9.57	3.23	2.41	10.46	3.91	1.23	322.41	−632.82	−613.03
V-T-P-D	4.36	2.41	3.23	7.88	3.91	8.53	315.27	−618.54	−598.75
V-T-D-P	4.36	4.73	3.23	7.65	0.21	9.88	322.77	−633.54	−613.75
P-V-T-D	8.92	4.36	4.73	−1.83	7.65	−4.43	329.49	−646.98	−627.19
P-V-D-T	8.92	9.57	4.60	−4.95	0.84	−0.57	314.18	−616.36	−596.57
P-D-V-T	3.23	9.57	4.36	10.46	1.93	−0.85	322.28	−632.57	−612.78
P-T-V-D	2.41	4.36	9.57	7.88	1.93	−4.76	320.43	−628.87	−609.08
D-P-V-T	3.23	8.92	4.36	11.37	−1.83	0.95	325.28	−638.37	−618.78
D-V-P-T	9.57	8.92	2.41	−4.95	4.29	0.91	321.50	−631.08	−611.22

For each structure, the parameter value of each pair-wise copula involved, maximum value of log-likelihood, AIC, and BIC, is reported. Bold indicated the best structure

with

$$h = \left(\frac{4}{3n}\right)^{1/5}\sigma, K(x) = \frac{1}{\sqrt{2\pi}}\exp\left\{\frac{-x^2}{2}\right\}, -\infty < x < \infty \tag{17}$$

where h, σ, and n are, respectively, the bandwidth of kernel, the standard deviation, and the total length of data.

The Weibull plotting position, $P\left(X \leq x_{(i)}\right) = \frac{i}{n+1}$, is used to calculate the empirical cumulative frequency, where i is the rank in ascending order and $x_{(i)}$ is the ith largest observation in a dataset of size n. The values of performance criteria (RMSE, MAE, and R) and the results of a goodness-of-fit test (K-S test) with the p values are given in Table 4. Considering the p-values, it can be found that, in all cases, empirical and non-parametric distributions have no significant differences. The comparison between empirical and non-parametric distributions for each variable is given in Fig. 3. Thus, the non-parametric distribution result is suitable for the modeling of flood volume, peak time, duration, and flood peak.

3.5 Vine copula structures

Here, the authors will (a) consider different D-vine structures permuting the variables, (b) fit the Frank copula for each pair-copula of the D-vine structures, (c) estimate the parameters of the pair-copulas, and (d) select the suitable D-vine copula

Table 9 D-vine structures for the variables P, V, D, and T_p, using all the data

$n = 300$									
1-2-3-4	θ_{12}	θ_{23}	θ_{34}	$\theta_{13\|2}$	$\theta_{24\|3}$	$\theta_{14\|23}$	Log-Lik	AIC	BIC
V-P-D-T	10.13	3.87	4.63	11.42	−0.23	0.75	508.82	−1005.6	−982.60
V-P-T-D	**10.13**	**2.16**	**4.74**	**4.31**	**3.35**	**9.15**	**525.92**	**−1039.8**	**−1017.62**
V-D-T-P	10.14	4.63	2.16	0.39	3.35	9.76	487.37	−962.75	−940.25
V-D-P-T	10.14	3.87	2.16	10.42	4.06	1.09	505.07	−998.15	−975.93
V-T-P-D	3.92	2.16	3.87	9.65	4.06	8.01	495.69	−979.93	−957.16
V-T-D-P	3.92	4.77	3.87	8.25	−0.23	9.70	497.91	−983.82	−962.60
P-V-T-D	10.13	3.92	4.74	−2.07	8.25	−4.40	519.35	−1026.7	−1004.49
P-V-D-T	10.13	10.14	4.63	−4.92	0.39	−0.58	498.40	−984.80	−962.58
P-D-V-T	3.87	10.14	3.92	10.42	2.27	−0.20	503.97	−995.95	−973.73
P-T-V-D	2.16	3.92	10.14	9.65	2.27	−4.48	507.96	−1002.9	−980.770
D-P-V-T	3.87	10.13	3.92	10.42	−2.07	0.98	512.95	−1013.9	−991.69
D-V-P-T	10.14	10.13	2.16	−4.92	4.31	1.08	510.62	−1009.2	−987.024

For each structure, the parameter value (θ) of each pair-wise copula involved, maximum value of log-likelihood, AIC, and BIC is reported. Bold indicated the best structure

Table 10 D-vine structures for the variables P, V, and D, using the different length of data

n = 30			n = 60			n = 120			
Structure	Log-Lik	AIC	BIC	Log-Lik	AIC	BIC	Log-Lik	AIC	BIC
V-P-D	**37.8**	**−69.6**	**−65.4**	**69.1**	**−132.3**	**−129.0**	**155.0**	**−304.1**	**−295.7**
V-D-P	32.1	−58.3	−54.1	64.3	−122.7	−116.4	148.87	−291.7	−283.3
D-V-P	31.8	−57.6	−53.4	66.2	−126.3	−120.1	149.88	−293.7	−285.4
n = 200				n = 300					
Structure	Log-Lik	AIC	BIC	Log-Lik	AIC	BIC			
V-P-D	**274.3**	**−542.6**	**−532.8**	**443.2**	**−880.3**	**−869.1**			
V-D-P	269.3	−532.6	−522.7	429.8	−853.7	−842.6			
D-V-P	267.1	−528.3	−518.4	443.0	−860.1	−849.0			

For each structure, maximum value of log-likelihood, AIC, and BIC is reported. Bold indicated the best structure

having the minimum value of AIC and BIC and maximum value of log-likelihood, both in three- and four-dimensional cases. To simplify the problem, and make the focus on the structure of D-vines, we have decided to consider only one family of bivariate copulas, the Frank family, which has presented the highest percentage of best agreement of pair variables. Initially, we have considered the case $d = 4$, with the variables P, D, T_P, and V. Twenty-four different D-vine structures are possible permuting the four variables, but only 12 give different decompositions. First, the D-vine structures are constructed using the first 30 events of dataset. Then, D-vine structures are calculated also using the first 60, 120, and 200 events and finally all the sample of data, i.e., 300 events. Table 5 provides for the first 30 events of data, the different D-vine structures, the pair-wise copulas involved, the constructed vine copulas, the estimated parameters, the maximum value of the log-likelihood function, and the values of the AIC and BIC. In Table 5, for the T_1 level, the parameters of pair-wise copulas involved are θ_{12}, θ_{23}, and θ_{34}. For the T_2 level, the parameters of pair-wise copulas involved are $\theta_{31|2}$ and $\theta_{24|3}$, and for the T_3 level, the parameter of copula involved is $\theta_{14|23}$. From Table 5, it can be found that *V-P-*T_P*-D* structure has the minimum value of AIC (−73.34) and BIC (−65.01) and the maximum value of log-likelihood (42.67) among the all considered structures. Tables 6, 7, 8, and 9 report the same information given in Table 5, using respectively the first 60, 120, 200, and 300 events of dataset. AIC, BIC, and log-likelihood values indicate that the structure *V-P-*T_P*-D* is ranked as first among all the possible D-vine structures, even if we vary the sample size from 30 to 300. For the first 60 events, *V-P-*T_P*-D* structure has the minimum value of AIC (−156.11) and BIC (−143.54) and the maximum value of log-likelihood (84.05) among the all considered structures, as well as for the first 120 events, it has the minimum value of AIC (−357.80) and BIC (−341.66) and the maximum value of log-likelihood (184.94). Similarly, for 200 and 300 events, *V-P-*T_P*-D* structure gives respectively the minimum values of AIC (−657.88, −1039.8) and BIC (−638.04, −1017.62) and the maximum values (334.94, 525.92) of log-likelihood. In this case, the vine structure among the flood characteristics does not change increasing the sample size from 30 to 300. This result is surprising, because we recognize that the selection of the vine structure could change, according to the uncertainty of sample observations. However, the fact that the vine structure is the same for all the samples (with quite different sample length) could be the indication of a

Table 11 D-vine structures for the variables P, V, and T_p, using the different length of data

n = 30		n = 60				n = 120			
Structure	Log-Lik	AIC	BIC	Log-Lik	AIC	BIC	Log-Lik	AIC	BIC
V-T_p*-P*	19.4	−32.8	−28.6	44.5	−83.7	−76.8	84.45	−162.9	−154.5
V-P-T_p	**20.3**	**−34.7**	**−30.2**	**45.9**	**−85.8**	**−79.5**	**90.2**	**−175.5**	**−167.7**
T_p*-V-P*	19.0	−32.0	−27.3	44.8	−83.6	−77.3	87.62	−169.2	−160.85
n = 200				n = 300					
Structure	Log-Lik	AIC	BIC	Log-Lik	AIC	BIC			
V-T_p*-P*	147.6	−289.2	−279.3	240.7	−475.4	−464.2			
V-P-T_p	**154.2**	**−302.4**	**−292.6**	**248.6**	**−491.2**	**−480.1**			
T_p*-V-P*	150.5	−295.1	−285.0	241.61	−477.2	−466.1			

For each structure, maximum value of log-likelihood, AIC, and BIC is reported. Bold indicated the best structure

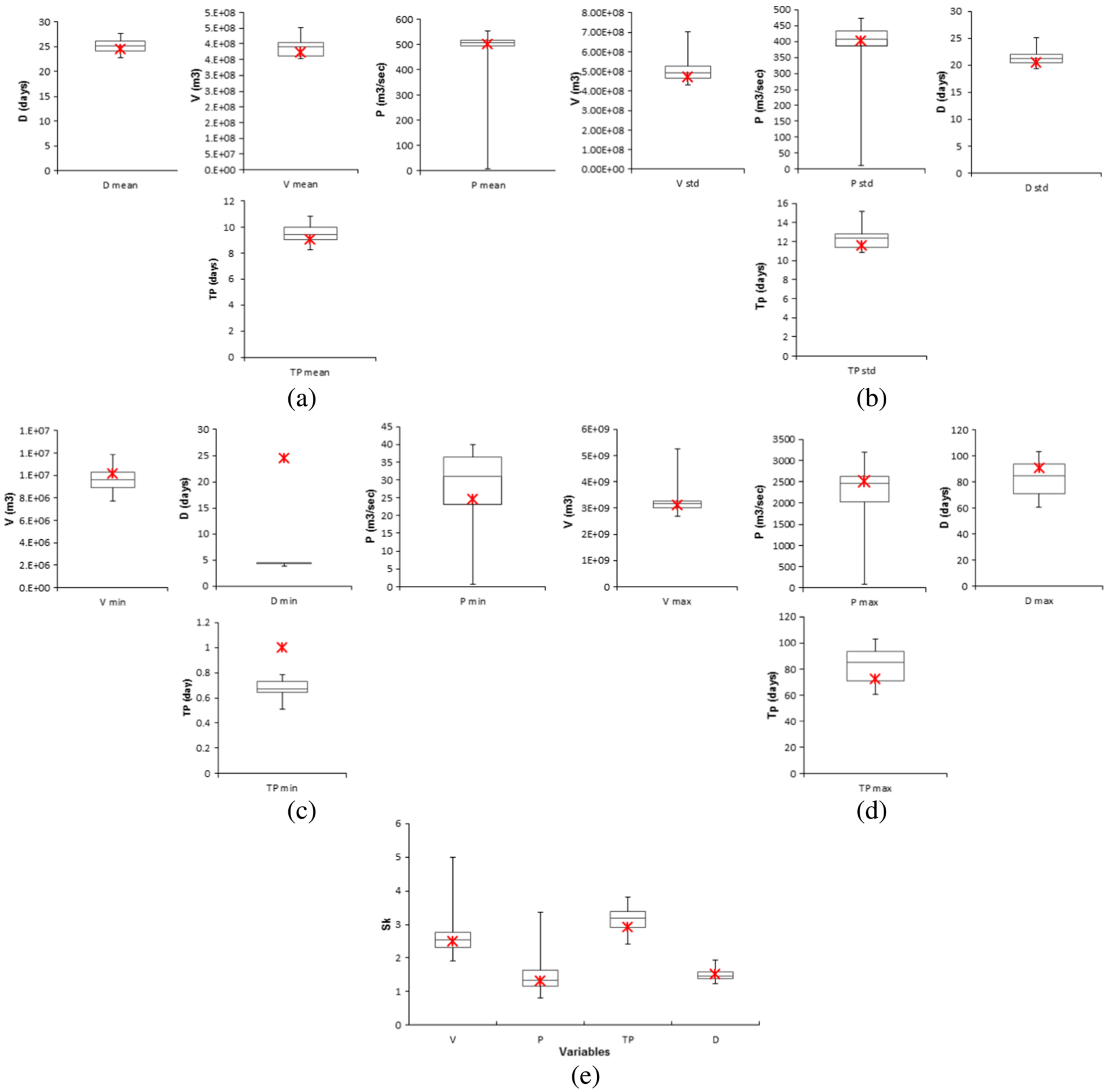


Fig. 4 Box plot of basic statistics **a** mean, **b** standard deviation (std), **c** minimum (min), **d** maximum (max), and **e** skewness (sk) of simulated samples by four-dimensional D-vine structure *V-P-T_P-D* with reported the observed value (×)

characteristic behavior of the variables under exam. Additional investigations are necessary in order to support this conclusion. The PIT test was applied as goodness-of-fit test. The results indicate that the null hypothesis cannot be rejected at 5 % significance level, for all the sample sizes, because *p* values are always higher than 5 %.

In literature, the ordering of the variables in the D-vine structure, i.e., the choice of D-vine structure, was based on the values of Kendall's tau. In particular, the most dependent pairs in the level T_1 were considered to identify the best D-vine structure (Aas et al. 2009; Vernieuwe et al. 2015). In this study, all the possible structures were considered by permuting the variables. From Tables 5, 6, 7, 8, and 9, it can be seen that the use of the most dependent pairs in T_1 does not provide the best model. In fact, *P-V-D-T_P* model is the one with the most dependent pairs: $\tau_{PV} = 0.66$, $\tau_{VD} = 0.67$, and $\tau_{DTp} = 0.40$, but it has not the largest log-likelihood and the minimum values of AIC and BIC in none of the vine copulas constructed using a sample size of 30, 60, 120, 200, and 300. Thus, this structure does not represent the model with the best performances. The *V-P-T_P-D* model with $\tau_{VP} = 0.66$, $\tau_{PTp} = 0.22$, and $\tau_{TpD} = 0.40$ has the minimum values of AIC and BIC and

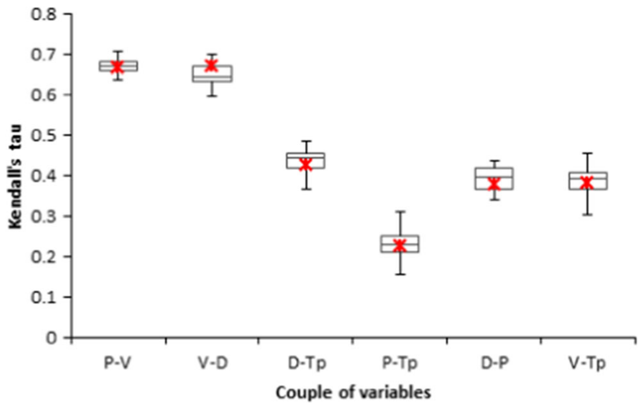


Fig. 5 Box plot of the Kendall's tau for each couple of variables using simulated samples and comparison to the Kendall's tau of the observed sample (×)

the largest log-likelihood in the constructed vine copula for each length of data. Similar considerations can be done in the analysis of three-dimensional D-vine structures. Tables 10 and 11 show three-dimensional structures respectively for the variables (P, V, and D) and (P, V, and T_P). The best vine structures are V-P-T_P and V-P-D. Here, similarly to the quadrivariate case, the increase of data length does not influence the selection of the best structure. The PIT test was applied to the best vine structures as goodness-of-fit test, indicating that the null hypothesis cannot be rejected at 5 % significance level, in both the two cases.

In order to evaluate the reliability of the V-P-T_P-D model in representing the behavior of the variables, a comparison between the statistics of simulated datasets and the statistics of the observed data has been done. In particular, an ensemble of 100 series, of size 300, and of the variables flood volume, flood peak, duration, and peak time was generated. The univariate statistics (namely mean, standard deviation, maximum, minimum, skewness) of each simulated sample are calculated and reported in box plots (see Fig. 4). From this, it is showed that mean, maximum, and standard deviation of all the flood characteristics are well reproduced; in fact, the statistics calculated using the observed data are within the box plots of simulated. Similarly, the minimum of V and P is well reproduced, while the minimum of T_P and D is not represented well. From Fig. 4, it can be found that the skewness of P, V, and D is well reproduced, while that of T_P is slightly overestimated.

In addition, in order to check if the selected vine structure is able to preserve the sample dependences among the flood characteristics, we have calculated the Kendall's tau for each pair of variables of the D-vine structure of each simulated sample. These values are represented in box plots given in Fig. 5 and compared to the values calculated on the observed sample. From Fig. 5, the comparison shows the similarity between observed and simulated values of Kendall's tau for each couple of variables.

4 Conclusions

We have investigated the use of D-vine structures to describe the statistical dependence among the main characteristics of a flood event: flood volume, flood peak, duration, and peak time. Here, we have decided to use vine structures because, with these, d-variate problems (with $d \geq 3$) can be modeled through pairwise distributions; in particular, we have investigated D-vine structures. The three-dimensional vine structures are used to describe the dependence among flood volume, flood peak, and the duration and among flood volume, flood peak, and peak time, while the four-dimensional D-vine structures the dependence among flood volume, flood peak, duration, and peak time. The three- and four-dimensional D-vine structures were fitted to the observed data. Marginals are assumed non-parametric, while the dependence parameters of pair-copulas are determined using maximum likelihood method. Samples of different size are selected from the dataset, namely the first 30, 60, 120, 200, and 300 (i.e., all the sample), to investigate if data length can affect the selection of the D-vine structure. The different D-vine structures were built permuting the variables, and the Frank copula was used for each pair-copula of D-vine structures. The best structure was selected according to AIC, BIC, and log-likelihood values. The PIT test was applied as goodness-of-fit test, indicating that the null hypothesis cannot be rejected at 5 % significance level, in all the models considered. V-P-T_P-D, V-P-T_P, and V-P-D D-vine structures were selected as the best structures for the modeling, four- and three-dimensional problems, respectively. It is found that the increasing data length does not influence the selection of the best D-vine structure. In addition, we found that the best D-vine structures (both for three- and four-dimensional cases) are not the ones having the most dependent pairs in T_1.

Then, the V-P-T_P-D D-vine structure has been used to simulate synthetic samples of flood characteristics. The statistics of the simulated samples are compared to the ones of the observed sample showing good correspondence with each other.

References

Aas K, Czado C, Frigessi A, Bakken H (2009) Pair-copula constructions of multiple dependence. Insurance. Insur Math Econ 44:182–198

Ashkar F, Rousselle J (1982) A multivariate statistical analysis of flood magnitude, duration, and volume. In: Statistical analysis of rainfall and runoff. Water Resource Publication, Fort Collins, pp. 651–669

Bedford T, Cooke R (2001) Probability density decomposition for conditionally dependent random variables modeled by vines. An Mat Arti Intel 32:245–268

Bedford T, Cooke R (2002) Vines—a new graphical model for dependent random variables. Ann Stat 30:1031–1068

Salvadori G, De Michele C (2007) On the use of copulas in hydrology: theory and practice. J Hydrol Eng 12(4):369–380

De Michele C, Salvadori G (2003) A generalized Pareto intensity-duration model of storm rainfall exploiting 2-copulas. J Geophys Res 108(D2):4067. doi:10.1029/2002JD002534

De Michele C, Salvadori G, Canossi M, Petaccia A, Rosso R (2005) Bivariate statistical approach to check adequacy of dam spillway. J Hydrol Eng 10:50–57

De Michele C, Salvadori G, Passoni G, Vezzoli R (2007) A multivariate model of sea storms using copulas. Coast Eng 54:734–751

De Michele C, Salvadori G, Vezzoli R, Pecora S (2013) Multivariate assessment of droughts: frequency analysis and dynamic return period. Water Resour Re 49(10):6985–6994

Favre A, Adlouni SE, Perreault L, Thiemonge N, Bobee B (2004) Multivariate hydrological frequency analysis using copulas. Water Resour Res 40:W01101. doi:10.1029/2003WR002456

Ganguli P, Reddy M (2013) Probabilistic assessment of flood risks using trivariate copulas. Theor Appl Climatol 111:341–360

Genest C, Favre A, Beliveau J, Jacques C (2007) Metaelliptical copulas and their use in frequency analysis of multivariate hydrological data. Water Resour Res 43:W09401. doi:10.1029/2006WR005275

Gräler B, Van den Berg MJ, Vandenberghe S, Petroselli A, Grimaldi S, De Baets B, Verhoest NEC (2013) Multivariate return periods in hydrology: a critical and practical review focusing on synthetic design hydrograph estimation. Hydrol Earth Syst Sci 17:1281–1296. doi:10.5194/hess-17-1281

Grimaldi S, Serinaldi F (2006) Asymmetric copula in multivariate flood frequency analysis. Adv Water Resour 29:1155–1167

Gyasi-Agyei Y, Melching C (2012) Modelling the dependence and internal structure of storm events for continuous rainfall simulation. J Hydro 464:249–261

Janga RM, Ganguli P (2012a) Bivariate flood frequency analysis of upper Godavari River flows using Archimedean copulas. Water Resour Manag 26(14):3995–4018

Janga RM, Ganguli P (2012b) Application of copulas for derivation of drought severity–duration–frequency curves. Hydrol Process 26(11):1672–1685

Janga RM, Ganguli P (2012c) Risk Assessment of Hydroclimatic Variability on Groundwater Levels in the Manjara Basin Aquifer in India Using Archimedean Copulas. J. Hydrol Eng 1345–1357. doi:10.1061/(ASCE)HE.1943-5584.0000564

Joe H (1997) Multivariate models and dependence concepts. Chapman & Hall, London

Kao S, Govindaraju R (2008) Trivariate statistical analysis of extreme rainfall events via the Plackett family of copulas. Water Resour Res 44:W02415. doi:10.1029/2007WR006261

Karmakar S, Simonovic SP (2009) Bivariate flood frequency analysis. Part 2: a copula-based approach with mixed marginal distributions. J Flood Risk Manag 2(1):32–44

Kurowicka D, Cooke R (2007) Sampling algorithms for generating joint uniform distributions using the vine-copula method. Comput Stat Data An 51:2889–2906

Linsley RK, Kohler MA, Paulhus JLH (1975) Applied hydrology. McGraw-Hill, New York

Ma M, Song S, Ren L, Jiand S, Song J (2013) Multivariate drought characteristics using trivariate Gaussian and Student t copulas. Hydrol Process 27:1175–1190

Mirabbasi R, Anagnostou EN, Fakheri-Fard A, Dinpashoh Y, Eslamian S (2013) Analysis of meteorological drought in Northwest Iran using the joint deficit index. J Hydrol 492:35–48

Mirabbasi R, Fakheri-Fard A, Dinpashoh Y (2012) Bivariate drought frequency analysis using the copula method. Theor Appl Climatol 108:191–206

Nikololoupoulos A, Joe H, Li H (2012) Vine copulas with asymmetric tail dependence and applications to financial return data. Comput Stat Data An 56:3659–3673

Salvadori G, De Michele C (2004a) Analytical calculation of storm volume statistics involving Pareto-like intensity-duration marginals. Geophys Res Lett 31:L04502. doi:10.1029/2003GL018767

Salvadori G, De Michele C (2004b) Frequency analysis via copulas: theoretical aspects and applications to hydrological events. Water Resour Res 40 (12)

Salvadori G, De Michele C (2006) Statistical characterization of temporal structure of storms. Adv Water Resour 29:827–842

Salvadori G, De Michele C (2010) Multivariate multiparameter extreme value models and return periods: a copula approach. Water Resour Res 46:W10501. doi:10.1029/2009WR009040

Salvadori G, De Michele C (2015) Multivariate real-time assessment of droughts via copula-based multi-site hazard trajectories and fans. J Hydrol 526:101–115

Salvadori G, De Michele C, Durante F (2011) On the return period and design in a multivariate framework. Hydrol Earth Syst Sci 15(11): 3293–3305

Salvadori G, De Michele C, Kottegoda NT, Rosso R (2007) Extremes in nature: an approach using copulas. Springer, Dordrecht

Salvadori G, Durante F, De Michele C (2013) Multivariate return period calculation via survival functions. Water Resour Res 49(4):2308–2311

Salvadori G, Durante F, De Michele C, Bernardi M, Petrella L (2016) A multivariate copula-based framework for dealing with hazard scenarios and failure probabilities. Water Resour Res, 52, doi: 10.1002/2015WR017225

Serinaldi F, Grimaldi S (2007) Fully nested 3-copula: procedure and application on hydrological data. J Hydrol Eng 12:420–430

Singh VP, Zhang L (2007) IDF curves using the Frank Archimedean copula. J Hydrol Eng 12:651–662. doi:10.1061/(asce)1084-0699(2007)12:6(651)

Sklar A (1959) Fonction de repartition a n dimensions et leurs marges, Publications de L'Institute de Statistique, vol. 8 edn. Universite' de Paris, Paris, pp. 229–231

Song S, Singh V (2010) Meta-elliptical copulas for drought frequency analysis of periodic hydrologic data. Stoch Env Res Risk A 24:425–444

Sraj M, Bezak N, Brilly M (2015) Bivariate flood frequency analysis using the copula function: a case study of the Litija station on the Sava River. Hydrol Process 29:225–238

Todorovic P, Zelenhasic E (1970) A stochastic model for flood analysis. Water Resour Res 6(6):1641–1648

Xiong L, Yu K, Gottschalk L (2014) Estimation of the distribution of annual runoff from climatic variables using copulas. Water Resour Res 50:7134–7152

Yue S, Ouarda TBMJ, Bobée B, Legendre P, Bruneau P (1999) The Gumbel mixed model for flood frequency analysis. J Hydrol 226(1–2):88–100

Zhang D (2014) Vine copulas and applications to the European Union sovereign debt analysis. Int Rev Financ Anal 36:46–56

Zhang L, Singh VP (2006) Bivariate flood frequency analysis using copula method. J Hydrol Eng 11(2):150–164