Parametric Vine Copula Construction for Flood Analysis for Kelantan River Basin in Malaysia

Shahid Latif [a*], Firuza Mustafa [b]

[a] *Ph.D. Research Scholar, Department of Geography, University of Malaya, 50603, Kuala Lumpur Malaysia.*

[b] *Associate Professor, Department of Geography, University of Malaya, 50603, Kuala Lumpur, Malaysia.*

Abstract

The multivariate approach of flood characteristics such as flood peak flow (P), volume (V), and duration (D) is much beneficial in recognizing the critical behaviour of flood episodes at a river basin scale. The incorporation of 2-dimensional copulas for establishing bivariate flood dependency frequently appears, but it could be more comprehensive if we focus all the three flood characteristic simultaneously. In such circumstances, incorporation of vine or Pair-Copula Construction (PCC) could produce a better approximation of joint probability density and much practical approach in the uncertainty analysis, in comparison with conventional trivariate copula distribution. This study demonstrated the efficacy of parametric vine copula in the modelling of trivariate flood characteristics for the Kelantan River basin in Malaysia. The D-vine tree structure is selected where the Gaussian and Frank copula is recognized for bivariate flood pairs (P-V) and (P-D) pairs in the first stage, using the maximum-pseudo-likelihood (MPL) estimation procedure. Similarly, the Gumbel copula is selected in the modelling of conditioned data obtained through the conditional distribution function of bivariate copulas selected in the previous stage based on the partial differentiation, also called h-function. Finally, the full density function of the 3-dimension structure is derived and compared with the observed flood characteristics. Furthermore, tail dependence properties and behaviour of D-vine copula are also investigated, which reveals for well capturing the general behaviour of Gaussian and Frank copula fitted to flood pair (P-V) and (V-D) and reproduces the overall flood correlation structure fairly well. Both the primary 'OR' and 'AND' joint return periods for trivariate flood events are estimated which pointing that 'AND' joint case produces higher return value than 'OR' case.

Keywords: Flood; Kelantan River Basin; Copula Function; D-vine Model; h-function.

1. Introduction

The designing and planning of flood defence infrastructure system often demand the accurate estimation of the specified exceedance probabilities or design quantiles of the flow characteristics for assessing the hydrologic risk. Flood is a multivariate probabilistic consequence, usually exhibited a complex dependency shown among its triplet intercorrelated vectors, flood peak flow, volume and its durations series [1-3]. The frequency analysis of flood episodes is the statistical approach of establishing the interrelationship between the magnitude of flood events (or flood event quantiles) and their frequency of occurrence using the most logical and parsimonious probability distribution functions [1, 4, 5]. In the earlier studies, the univariate flood frequency approaches were frequently incorporated for estimating the univariate return periods or assessing flood risk [4]. However, on later, numerous studies recognized the drawback or unreliability of the single variable flood modelling approach which would be

---

attributed for underestimation or overestimation in the assessments of hydrologic risk and thus referred towards the multivariate joint probability analysis by integrating the multiple flood characteristics in visualizing the pair-wise dependency measures [5, 6].

Recently, the copula function is recognized as the most dynamic statistical tool in the establishment of joint probability distribution analysis in comparison with the traditional (or conventional) multivariate functions, due to its flexibility in the selection of the marginal distributions and their joint dependence structure separately [7-9]. De Michele [10], firstly incorporated copulas in establishing the joint dependence structure of the storm intensity and duration samples. After that, numerous kinds of literature incorporated the 2-dimensional copulas framework for tackling the different hydro-climatic observations such as flood, drought or rainfall [2, 11-13]. Readers are advised to follow Saklar [7], Nelsen [8] and Genest and Favre [14] for the detailed study about the essential mathematical theorems and terminologies associated with copula function and also, 'International Association of Hydrological Sciences (or IAHS)' which provided an extensive collection of literature related to copulas function in the field of hydro-climatological modelling.

Most of the existing copula-based methodology is frequently focused in the establishment of bivariate flood distribution analysis [11, 15]. However, the more comprehensive hydrologic risk analysis can be achieved through accounting all the relevant intercorrelated flood characteristics at the same time, i.e., using the trivariate joint copula distribution modelling. In actuality, the potential damage could likely depend upon the several associated flood variables such that ignorance of spatial dependency among its multiple vectors might be attributed for underestimation of uncertainty distributed over the estimated design quantiles and thus often demands more flood variables [3, 16]. Also, dealing with multiple flood design variables would limit the applicability of analysing only based on the bivariate joint analysis but, due to the trivariate flood characteristic, it could be demanding for simultaneous accountability of its all intercorrelated vectors [17-19]. Therefore, accountability of multiple relevant flood characteristics could provide a better demonstration of their mutual dependence structure and a logical risk-based decision making for solving several water-related queries. Minimal literature focused on the multivariate analysis through incorporating higher dimensional copula families by integrating more than two variables in defining the trivariate joint analysis and their associated return periods such as Zhang and Singh [2], Graler et al. [3], Reddy and Ganguli [18], Daneshkhan et al., [19], Madadgar and Moradkhani [20], Fan and Zheng [21] and references therein.

In the development of the higher dimension copula distribution framework, literature often targeted an interactive set of conventional trivariate copulas structure, but all such attempts usually surrounded with few statistical constraints and issues [19, 22, 23]. Few incorporations based on some standard trivariate copulas framework such as Grimaldi and Serinaldi [24] introduced the mono-parametric and asymmetric or fully nested Archimedean (or FNA) based Frank distribution function along with the Gumbel logistic distributions for flood characteristics. Similarly, Serinaldi and Grimaldi [25], also introduced and pointed out the significance and flexibility of FNA copula framework for the trivariate flood events. On another side, Genest et al. [26] pointed out the effectiveness of meta-elliptical copulas in the treatment of annual spring flood characteristics. They thus revealed that it could be flexible and effective in the handling of multidimensional observation because it exhibited a correlation matrix and which significantly preserved the pairwise dependencies exhibited among the random variables. This study also pointed out that the applicability of meta-elliptical copulas might be ineffective under the low probabilities unless the asymptotic properties of data will be justified through the strong arguments. Similarly, Reddy and Ganguli [18] study demonstrated the efficacy of the fully nested Archimedean or FNA class copula along with the Student's t copula function for the annual flood characteristics. The same study also pointed out the importance of considering the trivariate return period (derived from both joint and conditional distribution) in comparison with bivariate or univariate return periods in the hydrologic risk assessments.

Fan and Zheng [21] pointed the adequacy of the Entropy copula in conjunction with the Gibbs sampling procedure in the treatment of flood characteristics and concluded that this copulas function could be easily projected into any higher dimension distribution space. Besides the above incorporation, Kao and Govindaraju [27] demonstrated the modelling efficacy of Plackett family of copulas for the rainfall characteristics, Madadgar and Moradkhani [20] modelled the drought characteristics using the trivariate Gumble and t copula also, Song and Singh [28] performed a frequency analysis of drought events using meta-elliptical copulas.

According to Daneshkhan et al. [19], it might be quite ineffective to reflect and recognize all the possible joint concurrency exhibited among multivariate characteristics through conventional (or standard) trivariate copula distribution approach. Also, resolving the mutual dependencies of multivariate extreme via the conventional copula framework is quite complicated due to the higher degree of uncertainty as well as complex dependence structure of the hydrologic characteristics thus it often demands the estimation of tail dependence coefficient for the various tail dependency in a precise manner through adopting a comprehensive multivariate methodology [19, 23].

In such circumstances, the incorporation of the vine or pair-copula construction (or PCC) in the modelling of the multidimensional complex hydrologic phenomenon could be the right choice and an effective way of uncertainty

analysis in comparison with standard trivariate copulas [22, 23, 29]. The applicability of vine or pair-copula framework seems much popular in the field of financial data analysis and their risk management [30]. However, in the recent decades, such incorporation is also recognized and accepted in the treatment of hydrologic observations such as in the modelling of drought episodes [28, 31], flood episodes [3, 19], rainfall or storm distribution modelling [32, 33]. The implementation of vine or pair-copula framework is entirely based on the fact that it breaks or segregates the full multivariate density through the nesting of pair-copulae, which is based on the conditional mixing procedure [29]. In other words, the multiple best-fitted bivariate (or 2-dimension) copulas are often employed for each random pair and their conditional and unconditional distribution to form a nesting framework instead of selecting a fixed copula for all the intercorrelated flood vectors and thus often reveals for higher flexibility in the construction of higher dimension hydrologic framework. Under the regular vine framework, the canonical vine, also called C-vine and D-vine are two holistic decomposition approach of the pair-copula approximation, [30, 34].

The D-vine structure is highly flexible in comparison with C-vine, especially in the presence of a known variable that controls the level of mutual dependency within the samples [19, 23]. The strength of dependency or degree of mutual concurrencies among the multivariate characteristics further decided in the selection of either the D-vine or C-vine structure [3, 34]. According to Graler et al. [3] and Hobaek et al. [35], the manner through which the decomposition of vine copula will be performed, which will further decide their approximation capability with the multivariate observations and which further reveals that the conditioning process of the vine or PCC model is not rigid or fixed. Readers are advised to follow Czado et al. [34] for the extended details over the selection procedure of the regular vine framework also, Kurowicka and Cooke [36] and Aas and Berg [37] for the details study over the C- and D-vine copula structure.

Flood becoming a crucial and most severe natural disaster for the Kelantan River basin, especially during the wet monsoon circulation in Malaysia and perceiving for increasing in term of their frequency and magnitude [38]. Such that, in the year 2002, intense and prolonged precipitation attributed for flooding and affected 714, 287 people, similarly, in the year 2014, much heavy precipitation caused flooding in most of the part of the eastern coast of Kelantan river basin and which affected more than 200,000 people [39].

In actuality, the downstream area of the Kelantan River basin is at the highest risk of devastating flood events according to the study performed by Nashwan et al. [40]. Multivariate analysis of flood episodes could be a practical approach and much beneficial for hydrologist and water practioner in recognizing the critical behaviour of flood episodes at a river basin scale and for solving different water-related issues. In this literature, we incorporated and tested the vine or pair-copula framework to analyze the block (annual) maxima-based flood characteristics for the Kelantan river basin in Malaysia. The 3-dimensional joint probability framework is simulated using the parametric vine or pair-copula framework approach in the trivariate distribution modelling of flood characterization and estimating the multivariate (i.e., trivariate and bivariate) primary joint return periods.

The structure of this paper as follow, as a prerequisite, the brief details on the theoretical aspect in constructing the 3-dimension vine or pair-copula distribution framework is discussed in the next section (or Section-2). In the third section, the vine copula-based methodology is incorporated to a case study for the annual flood characteristics in the Kelantan River basin which further comprises, the selection of appropriate vine tree structure (i.e., D- or C-vine) and then simulation of vine structure by fitting most appropriate 2-dimensional copulas in a stage-wise hierarchical nesting manner and also, estimating the primary joint return periods for the trivariate (and bivariate) flood episodes. The fourth section provides the research conclusions. Figure 1 illustrates the methodological flow diagram, indicating the steps of analysis.
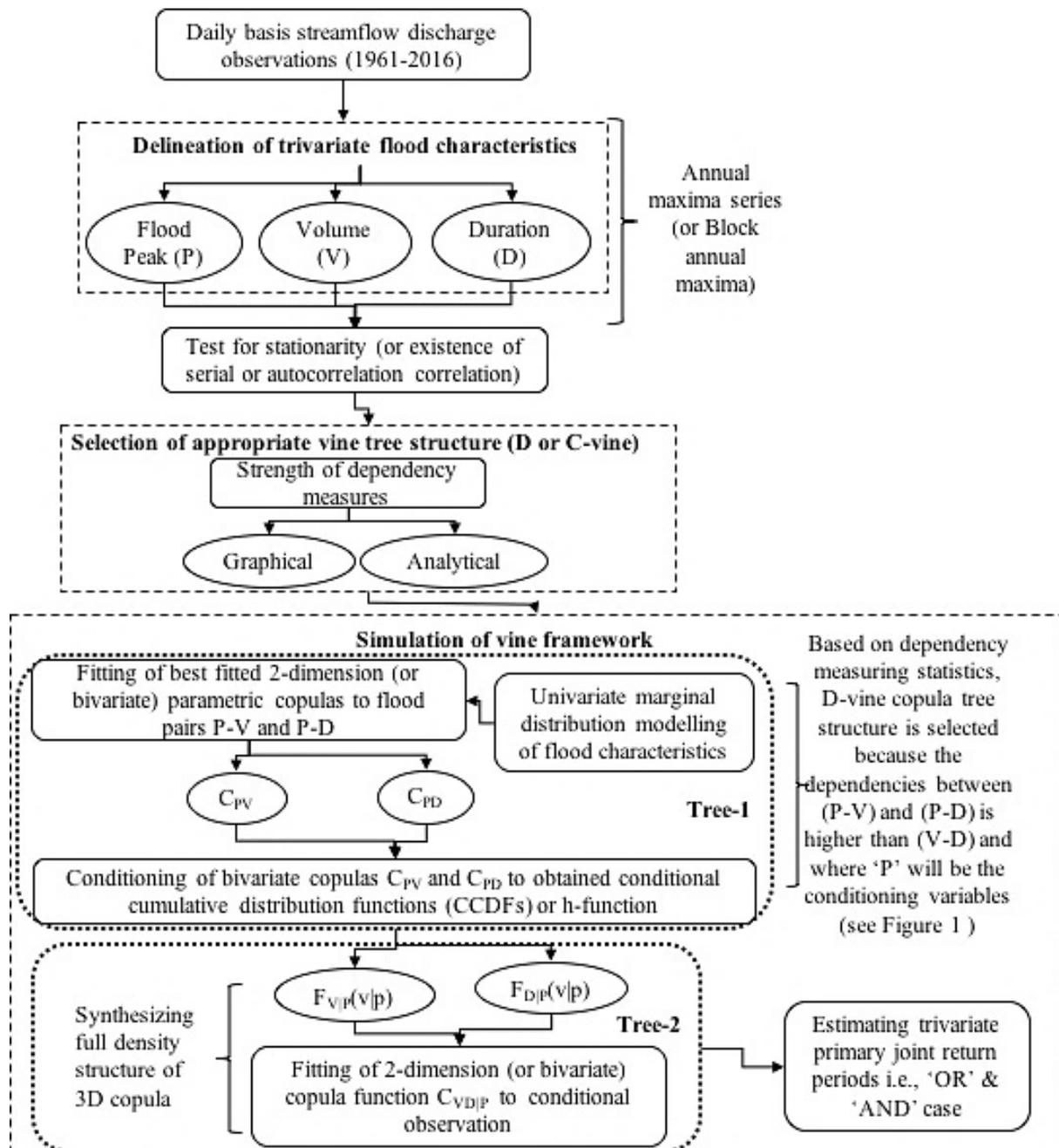
Figure 1. Methodological flow diagram of vine copula framework for trivariate flood distribution analysis

2. Theoretical Framework

2.1. Construction of 3-Dimension Vine or Pair-Copula

According to Nelsen [8], copula function interlinks the multivariate distribution to their univariate probability functions, also called the marginal distribution, and their idea was firstly given by Saklar [7]. Multivariate copula function modelled the univariate marginal density of individual intercorrelated random variables independently from their joint dependence structure. According to Sklar's theorem [7], if (X, Y) be bivariate random variables with continuous marginal distributions such as, $u_1 = F_X(x)$, and $u_2 = F_Y(y)$, then it can be characterized uniquely by using the Copula function or C, and defined on the unit square;

$$H_{X,Y}(x, y) = C[F_X(x), F_Y(y)] = C(u_1, u_2) \tag{1}$$

Where, C = bivariate copula under investigation; $F_X(x) = F_Y(y)$ = cumulative distribution function (or CDF) of univariate observations i.e., X and Y; $H_{X,Y}(x, y)$ = bivariate joint probability density function (or PDF), which can be expressed in terms of its univariate marginal functions and the associated dependence function C, as revealed from Equation 1. Also, if $F_X(x)$ and $F_Y(y)$ are continuous which indicates the copula C must be unique [11, 41].

Conversely, if $F_X(x)$, $F_Y(y)$ and the copula $C[x, y]$ is given, then Equation (1) must defines the bivariate joint distribution functions with its marginal distributions $F_X(x)$ and $F_Y(y)$. Similarly, if $f_X(x)$ and $f_Y(y)$ are the PDF of variable X and Y, then the joint probability density can be defined through Equation (2), as given below;

$$f_{X,Y}(x, y) = c(F_X(x), F_Y(y)) \, f_X(x) f_Y(y) \tag{2}$$

Where, c is the density function of bivariate copula C and which can be further expressed as

$$c(u_1 u_2) = \partial^2 c(u_1 u_2) / \partial u \partial v \tag{3}$$

The construction of vine or PCC framework is solely based on the principle of stagewise hierarchical nesting procedure, also called the conditional mixing of the bivariate (or 2-dimension) copulas where the decomposition of full multivariate density into a simple local building block are carried out through the conditional independence or pair-copula [3, 23, 29]. Thus, such conditional mixing approaches attributes for higher flexibility in the modelling of trivariate flood characteristics. At the beginning of vine construction, the marginal empirical distribution functions are obtained by transforming the samples of all the three flood characteristics, P, V, D, within the range [0, 1] using the rank order transformation S (see Equation 4).

Figure 2 illustrates the 3-dimensional parametric vine framework. In the first tree (or Tree-1) structure, all the relevant pairwise joint association between the flood characteristics are modelled using the best-fitted bivariate copula function, i.e., between (V-P) pair and (P-D) pair (where we placed the flood variable, P between the other two variables, V and D, and which is solely based on the strength of dependencies exhibited among the flood characteristics, see section 3.2.1 for further details). In this demonstration, we introduced the different varieties of bivariate copula families in the establishing the mutual dependency between flood characteristics such as the Archimedean copulas (i.e., Clayton, Frank, Gumbel, Joe) [10, 11, 15] and one Elliptical copula [i.e., Gaussian or normal copula] [42, 43]. In actuality, the Clayton and Gumbel family only capture the positive dependencies, but the Frank copula significantly captures the entire range of dependencies (i.e., $\tau_\theta \in [1, -1]$).

According to Aas et al. [23] and Graler et al. [3], the bivariate copulas fitted to flood pairs (V-P) and (P-D) in the first tree (or Tree-1) can be conditioned under the flood characteristics, P, using the partial differentiation which is also called the h-function (i.e., $h(v, p, \theta)$ and $h(d, p, \theta)$) to obtain the conditional cumulative distribution functions (CCDFs), $F_{V|P}$ and $F_{D|P}$ using the Equations 5 and 6 as given below;

$$S(y) = \frac{\text{rank}(y)}{n+1} \tag{4}$$

$$h(v, p, \theta) = F_{V|P}(v|p) = \frac{\partial C_{VP}(v, p)}{\partial p} \tag{5}$$

$$h(d, p, \theta) = F_{D|P}(d|p) = \frac{\partial C_{DP}(d, p)}{\partial p} \tag{6}$$

The copulas dependence parameter in the first tree of vine structure (see, Figure 2) i.e., $(\theta_{vp}, \theta_{pd})$) are estimated from the original flood characteristics P, V and D using the maximum-pseudo-likelihood (or MPL) estimation procedure, which is based on maximizing the pseudo log-likelihood function [15, 44, 45]. After that, the transformed data (obtained using Equation 5 and 6), are employed to determine the unknown copula dependence parameters in the second tree (or Tree-2) and thus used to fit another bivariate copula $C_{VD|P}$. The best-fitted bivariate copulas in both Tree-1 and Tree-2 are selected using the Cramer-von Mises (or CvM) test statistics which is further based on the parametric bootstrapping procedure that uses the CvM test statistics, S_n [46, 47]. Finally, the full density function $c_{PVD}(p, v, d)$ of the 3-dimensional copula can be derived using the Equation 7 [19, 23];

$$c_{PVD}(p, v, d) = c_{VD|P}\left(F_{P|V}(p|v), F_{D|V}(d|v)\right) . c_{VP}(v, p) . c_{PD}(p, d) \tag{7}$$
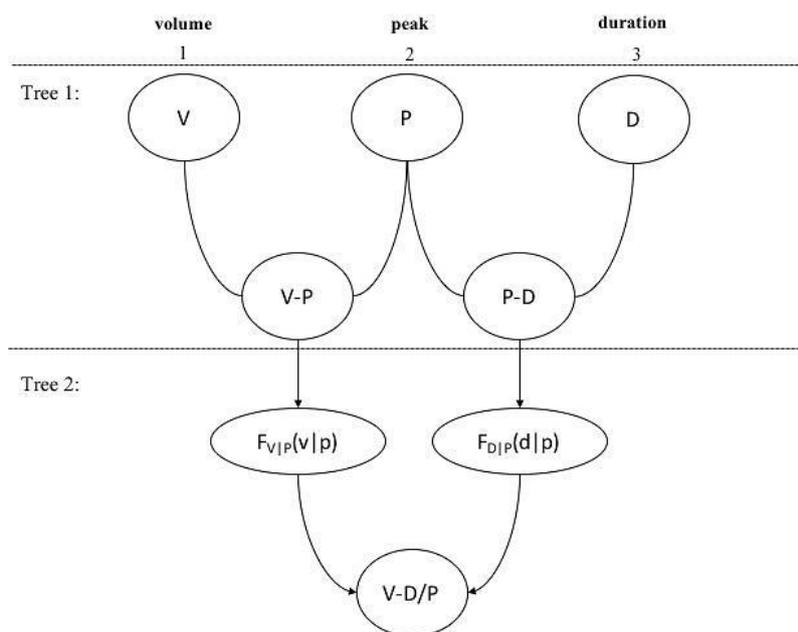
Figure 2. Construction of 3-dimensional vine or PCC framework based on stage-wise hierarchical nesting of bivariate copulas for the flood characteristics, flood volume (V), peak (P) and duration (D)

3. Applications

3.1. Description of Study Area and Flood Characteristics

This experiment targeted the daily basis streamflow discharge records from the period (1961-2016) of the Kelantan river basin, which are collected at the Gulliemard Bridge gauge station (located at the downstream region) [see Figure 3] and delivered by the Drainage and Irrigation Department, Malaysia. The geographical location of this river basin is Lat $4°$ $30′$ N to $6°$ $15′$ N and Long $101°$E to $102°$ $45′$ E. This river is originating from Tahan mountain range to the South China Sea (a north-eastern portion of Peninsular Malaysia). The drain area of this river is about 13100 km^2. According to DID report [38], the estimated runoff of this river basin is about 500 m^{3}sec^{-1} where the variations of annual precipitations are in between 0 mm (dry period)-1750mm (wet or north-eastern monsoonal period). According to Hussain and Ismail [48], the frequency of flood episodes has highest in the Gulliemard Bridge, Lebir and Galas gauge stations in compare with the Nenggiri gauge station. Nashwan et al., [40] literature also revealed that the risk of devasting flood episodes in the downstream portion of this river basin is higher.

Annual (Maximum) series or block (annual) maxima and Peak over Threshold (or POT) are the two holistic technique widely accepted in the frequency analysis of hydrologic extreme [49]. In this experiment, we adopted the block (annual) maxima approach in the delineation of the triplet intercorrelated flood vectors, i.e., P, V, D series from the daily basis streamflow discharge records collected at the Gulliemard Bridge gauge station. For each observation year, there are only one flood episodes at the target site which is characterized through pointing the maximum discharge flow value, also called flood peak flow, P and their hydrograph volume, V and duration, D series are extracted corresponding to each flood peak using the algorithm as described in the literature such as Zhang and Singh [2], Yue and Wang [6]. Table 1, defines the descriptive behaviour of the derived annual flood events characteristics and also Figure 4, illustrates the histogram distribution of the derived annual flood characteristics and which indicates the existence of positively skewed distribution behaviour.

Table 1. Summary statistics of the flood episodes characteristics

Descriptive statistics	P (m³/sec)	V (m³)	D (days)
Sample Size	50	50	50
Mean	6078	19122	19.04
Std. Deviation	4639	14623	10.851
Coef. of Variation	0.76324	0.76473	0.56993
Skewness (Pearson)	1.506	1.590	2.210
Kurtosis (Pearson)	1.883	2.864	6.252
Min	916.3	3182.3	7
Max	20586	74740	64

In the multivariate distribution analysis, it is often an essential concern to check the existence of stationarity within individual flood characteristics before introducing into univariate or multivariate copula framework. For this, Ljung and box [50] based hypothesis testing also called Q-statistics, are undertaken [51]. Therefore, the Q-statistics for the sample size 'm' with 't' defines the total no of lag being tested with sample autocorrelations at the specific lag, i.e., $\hat{\rho}_t$ are given below;

$$Q = m(m+2) \sum^{h} \hat{\rho}_t^2 / m - t \tag{8}$$

Where, Null hypothesis (H_0) = zero autocorrelation or independent distributions; and Alternative hypothesis (H_1) = existence of serial correlation (or autocorrelation). The Q-statistics and their associated p-value for different lag size (i.e., 30, 20, 15, 10 and 5) are listed in Table 2 and which indicates for almost zero first-order autocorrelations because their estimated statistics are less than the critical value for each of the univariate series of flood characteristics by accepting the null hypothesis (H_0) at 5% or 0.05 significance level against their alternative hypothesis (H_a).
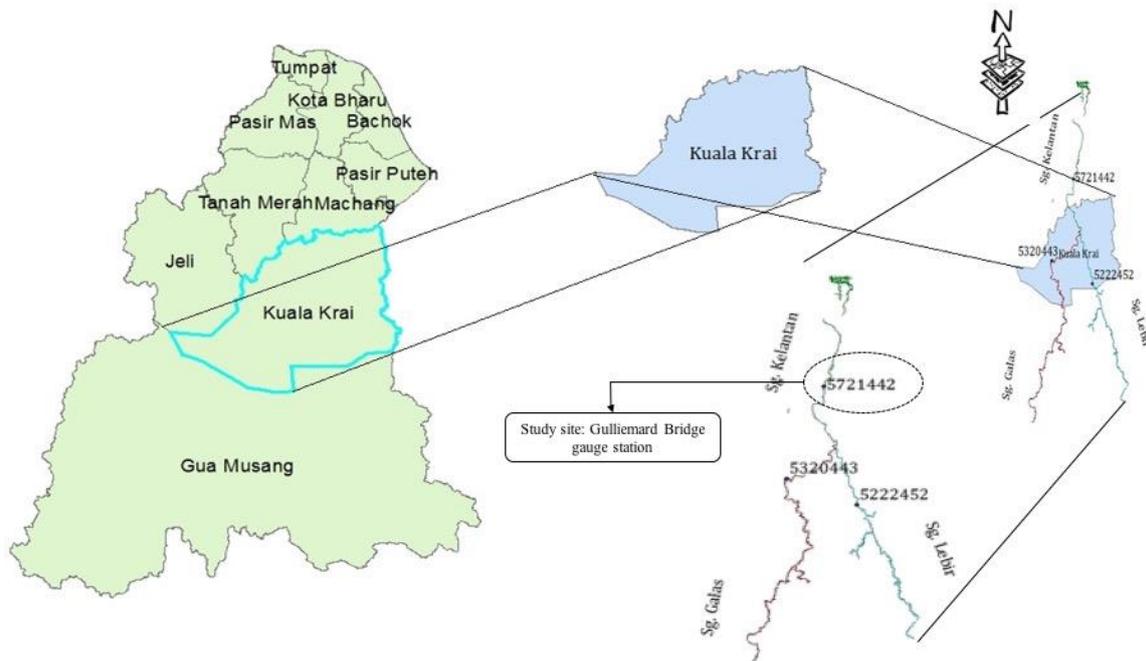
Figure 3. Geographical location of the study area (Gulliemard Bridge gauge station located at the downstream region of Kelantan River basin in Malaysia)
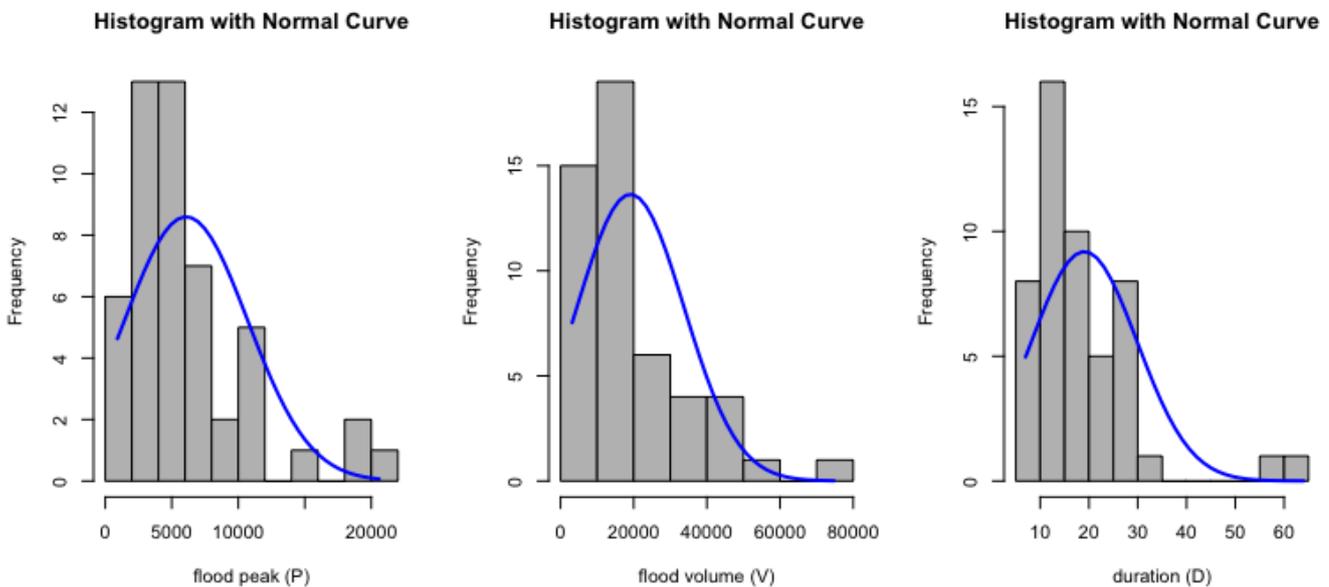
Figure 4. Histogram distribution of annual flood characteristics

Table 2. The Q-statistics and their corresponding p-value

Flood vectors	Box-Ljung test	Lag size (30)	Lag size (20)	Lag size (15)	Lag size (10)	Lag size (5)
	X-squared (Q-statistics)	18.87	14.88	13.34	9.72	4.89
P (peak)	df	30	20	15	10	5
	P-value	0.9427	0.7828	0.5762	0.4655	0.4295
	X-squared (Q-statistics)	14.72	11.98	10.68	7.49	0.49
V (volume)	df	30	20	15	10	5
	P-value	0.9912	0.9167	0.7747	0.6779	0.9922
	X-squared (Q-statistics)	23.34	15.70	13.96	6.17	3.03
D(duration)	df	30	20	15	10	5
	P-value	0.8009	0.7349	0.5287	0.8007	0.6946
Note	Critical value (at 5% level of significance)	43.77	31.41	25	18.31	11.07

3.2. Modelling Flood Characteristics using Vine (or PCC) Model

The fitting procedure of the vine or pair-copula distribution framework to the delineated trivariate flood characteristics, i.e., flood peak (P), volume (V) and duration (D) comprise a series of computation task. The first stage of modelling comprises the selection of an appropriate vine tree structure (i.e., either D- or C-vine) that will be solely depending upon the strength of dependencies among flood characteristics pair and then after, to select the most parsimonious bivariate copulas families for pair-copulas in establishing different joint distribution pairs of flood characteristics and estimation of their dependence parameters. Different statistical tools are available that can support in the selection of most logical vine tree structure such as, graphically based, i.e., scatter plot, Kendall's plot, chi-plot and also analytically based correlation coefficients measures. Such tools can also be employed to select the most parsimonious parametric bivariate copula family for the underlying flood characteristics for describing the dependence structure.

3.2.1. Selection of Appropriate Vine Tree Structure

Figures 5, 6 and 7 illustrated the scatter plots, chi plots [52] and Kendall's plots [53] for analyzing the degree of interassociation and strength of dependency between flood attribute pairs. In actuality, the Chi-plot uses data ranks, also λ_i values measure the distance of bivariate random observations (say $p_i v_i$) from the center of the data sets within the range of [-1, 1]. Thus, their values will tend to positive or negative for positively or negatively correlated random pairs, Also, the control limits χ_i are the another measuring factor in chi-plot that are placed at $\chi = \pm c_p/\sqrt{n}$ [52]. Therefore, in the case of stronger dependency exhibited, the random observations must be outside the control limit of chi-plot otherwise, inside the control limit region can be indicated for independence between random pairs. Similarly, Kendall's plot is analogous to quantile-quantile (Q-Q) plot such that, deviation of random observations from the main diagonal of K-plot reveals the existence of mutual correlation otherwise could be revealing for independence when the plot tends to be linear [14, 53]. Based on the above graphical investigation, it could be revealed that flood peak-volume (P, V) pair exhibited strong and positive correlation structure while the flood pairs volume-duration (V, D) and peak-duration (P, D), both exhibited weak and negative correlation. Table 3 indicates the correlation coefficient statistics and thus, which support the similar conclusion taken from the graphical-based flood characteristics plots. Therefore, it must be concluded that the dependencies between the flood pairs (P, V) and (P, D) is higher than (V, D) that means the flood peak, P must be placed between the other two variables flood volume, V and duration, D, as graphically illustrated in Figure 2. According to Graler et al. [3], the choice of the conditioning variable (i.e. P, in Figure 2) is not unique, and different choices might lead to different results also, according to Hobaek et al. [35], different vine copula decomposition procedure differently approximates the multivariate distribution. Overall, it must be concluded that the D-vine copula could be a right choice and thus employed in this study for constructing 3-dimensional flood characteristics and which are more flexible than the canonical or C-vine copula because the D-vine model construction provides higher flexibility in the selection of which flood attribute pairs to model than C-vine [23, 34].

Table 3. The correlation coefficients statistics between the flood characteristics

Dependence measure	Peak-Volume(P-V)	Volume-Duration(V-D)	Peak-Duration(P-D)
Pearson's correlation (r)	0.7387784	-0.1079511	-0.0061526
Kendall's correlation(τ)	0.60759499	-0.0225141	-0.0741828
Spearman's correlation (ρ)	0.79425677	-0.0343127	-0.094851

Scatter plot of flood charactreistics
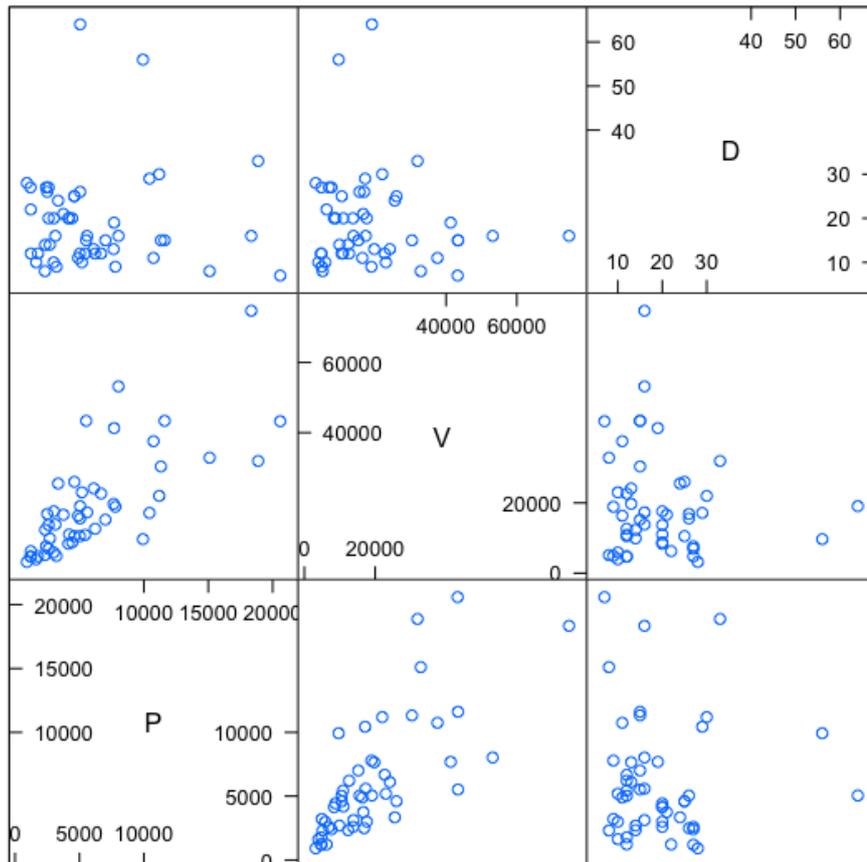
Figure 5. The scatter plots of the flood characteristics
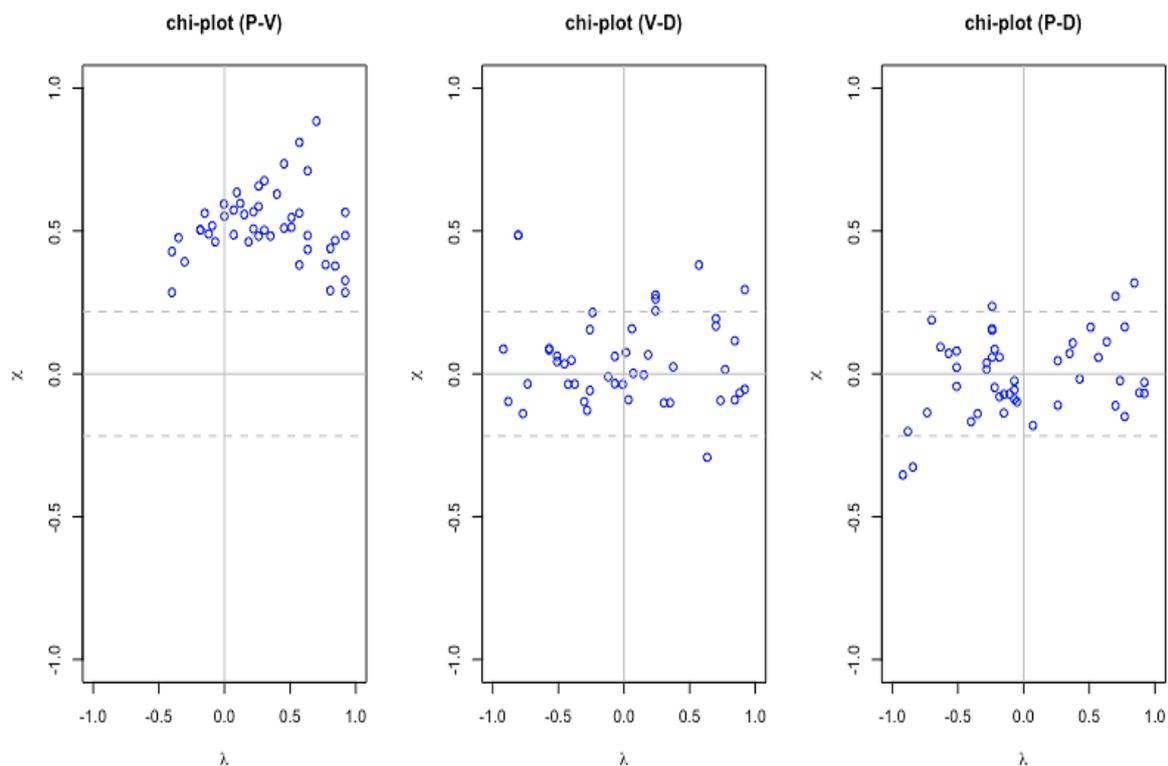
Figure 6. Chi plot of the pairwise flood characteristics; between flood peak-volume (P-V) pair which shows positive dependence structure and between volume-duration (V-D) pair and flood peak-duration (P-D) pair shows negative and weak dependence structure
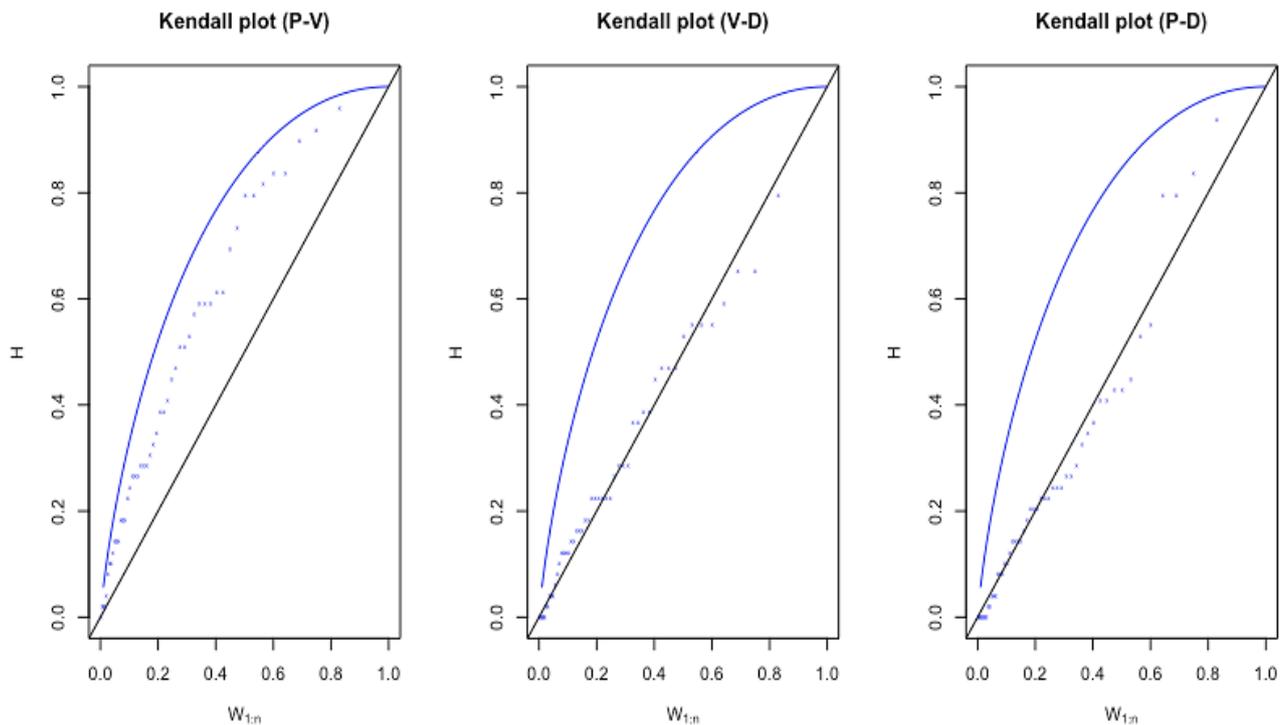
Figure 7. Kendall plot (or K-Plot) of the pairwise flood characteristics; between flood peak-volume (P-V) pair which shows positive dependent and between volume-duration (V-D) pair and flood peak-duration (P-D) pair shows negative and weak dependent

3.2.2. Simulation of Vine Tree by Fitting 2-Dimensional Copulas

After the selection of the appropriate vine framework, the bivariate joint dependence structure of the flood attribute pairs is defined through selecting the most appropriate 2-dimension (or bivariate) copulas function, such as C_{PV} for (P-V) flood pairs and C_{PD} for (P-D) pair copulas (referred to Figure 2). Similarly, the bivariate copulas, i.e., C_{PV} and C_{PD} are conditioned under the variable P (i.e., flood peak) [23] and thus the conditional CDFs, also called the h-functions are derived using the Equations 5 and 6, in defining the realization demands in the second tree (or Tree-2) of vine structure. Finally, a bivariate copulas $C_{VD|P}$ is fitted to the conditioned observations i.e., $F_{V|P}(v|p)$ and $F_{D|P}(d|p)$. Actually, the conditional cumulative distribution values are estimated for all the triplet flood characteristics in the vine tree structure.

Before initiating the copula fitting procedure for vine framework, we introduced commonly used univariate probability functions which belongs to parametric families such as Gamma-3P, Generalized Extreme Value (GEV), Generalized Gamma-3P, Inverse Gaussian-2P, Johnson SB-4P, Lognormal-2P and Weibull-2P, as a candidate function for defining the marginal structure of P, V and D series individually. All the univariate distribution fitting procedures along with their parameter estimations are carried out using the Easyfit software. The empirical observations are retrieved from commonly used Gringorten based position-plotting formula [11, 54]. The best-fitted probability function is selected for each individual flood characteristics using different goodness-of-fit (GOF) test statistics such as Kolmogorov-Smirnov (or K-S) test and Anderson-Darling (or A-D) test [55], based on information criteria statistics such as Akaike Information criteria (or AIC) [56], Schwartz's Bayesian Information criteria (or BIC) [57] and Hannan-Quinn Information criteria (HQC) [58]. Table 4 illustrates the GOF test statistics, which clearly pointed out that the flood peak flow is best modelled by the Lognormal-2P distribution, volume series by the Johnson SB-4P distribution and duration series by the Gamma-3P distribution.

Figure 2 illustrated the D-vine structure incorporated in the present demonstration. The pair copula exhibited three bivariate copula functions, i.e., C_{PV} and C_{PD} in the first tree (or Tree 1) and $C_{VD|P}$ in the second tree of the vine structure. In this demonstration, we introduced and investigated different copulas families which include one Gaussian copula that belongs to the Elliptical family and the Archimedean copula families such as the Gumbel, Frank, Clayton and Joe copula, as a possible candidate model. In actuality, the Gumbel, Clayton and Joe, copula cannot be used for negatively correlated flood characteristics (i.e., Kendall's tau < 0), which is only applicable in the modelling of positively correlated random variables. The copula dependence parameters fitted to the observed flood characteristics are estimated using the Maximum-Pseudo-Likelihood (MPL) estimation procedure [46, 47] and their estimated values are listed in Table 5. The Cramer-von Mises (CvM) test statistics 'S_n' with the parametric bootstrapping procedure is employed to evaluate the adequacy of hypothesized copulas fitted to bivariate flood pairs (P-V) and (P-D), i.e., C_{PV}

and C_{PD} copulas, which is often considered as one of the most potent model compatibility testing. Therefore, minimum the value of 'S_n' test statistics must indicate for minimum gap or dispensary between empirical and derived parametric copulas, in other words, this copula would be most justifiable and effective for establishing the bivariate joint relationship. Overall, from the Table 5, it must be concluded that the Gaussian copula and the Frank copula recognize as most justifiable for defining bivariate joint dependence structure of flood pairs (V, P) and (P, D) in the first tree (or Tree-1) of vine structure. Figure 8 illustrated the density and contour plot derived from the bivariate copulas fitted in the first tree between the flood pairs V-P and D-P while Figure 9 indicates the bivariate contour plot derived from joint CDFs of the Gaussian copula (C_{VP}) and the Frank copula (C_{DP}). Similarly, based on GOF test statistics, the Gumbel copula is recognized as the most parsimonious bivariate structure to model the variable (V, D) conditional on P, in the second tree (or Tree-2) (see Figure 10, which illustrates the surface plot of bivariate Gumble copula fitted with conditioned observation in the second tree of vine structure). Therefore, for the 3-dimensional case of flood characteristics, the three selected bivariate copulas such as C_{PV}, C_{PD}, and $C_{VD|P}$ are introduced and composed into the 3-dimension vine copula using the Equation 7. Simulation of the flood characteristics taken from the vine or pair-copula model is based on the simulation approach as discussed by the literature such as Aas et al. [23]. Figure 11 illustrates the comparison based on scatter plot between the observed and simulated flood characteristics and which indicates a good agreement based on the fitted models, in other words, the vine model reproduces the overall flood correlation structure reasonably well.

Table 4. Goodness of fit (a) based on Kolmogorov-Smirnov and Anderson-Darling test statistics (b) based on information criteria statistics such as AIC, BIC and HQC

(a)	Peak			Volume			Durations		
Functions	p-value	KSn (d- max)	Adn (d-max)	p-value	KSn (d-max)	ADn (d-max)	p-value	KSn (d- max)	ADn (d-max)
GEV(3P)	0.99655	0.05451	0.21667	0.99931	0.04897	0.24945	0.82259	0.086	0.35244
Gamma (3P) *	0.8802	0.08007	0.26953	0.98701	0.06089	0.21109	0.89254	0.07865	0.37708
Log-Normal (2P) *	0.9977	0.05293	0.19412	0.98539	0.06157	0.2338	0.60127	0.10511	0.4602
Weibull (2P)	0.81311	0.0869	0.73212	0.89172	0.07875	0.63575	0.23928	0.14235	1.5472
Inv. Gaussian (2P)	0.98175	0.06293	0.38095	0.81919	0.08633	0.48954	0.87056	0.08114	0.60496
Gen.Gamma (3P)	0.66896	0.09944	0.45939	0.89941	0.07782	0.36811	0.28097	0.13672	0.91168
Johnson SB (4P) *	0.84788	0.84788	14.822	0.99811	0.05222	0.17314	0.56249	0.1084	11.874

Notes. K-S test stands for Kolmogorov-Smirnov test; A-D test stands for Anderson-Darling test

* Indicates that Lognormal (2P), Johnson SB (4P) and Gamma (3P) distribution exhibited minimum test statistics i.e., K-S and A-D values for describing flood peak, volume and duration series.

(b)	Peak			Volume			Duration		
Functions	AIC	BIC	HQC	AIC	BIC	HQC	AIC	BIC	HQC
GEV(3P)	-374.335	-368.599	-372.15	-268.985	-263.249	-266.8	-336.32	-330.583	-334.135
Gamma (3P) *	-216.301	-210.565	-214.12	-210.107	-204.371	-207.92	-343.62	-337.88	-341.438
Log-Normal (2P) *	-379.344	-375.52	-377.89	-371.028	-367.204	-369.57	-327.46	-323.633	-326.001
Weibull (2P)	-329.681	-325.857	-328.23	-342.868	-339.044	-341.41	-292.91	-289.085	-291.453
Inv. Gaussian (2P)	-362.489	-358.665	-361.03	-344.722	-340.898	-343.27	-325.76	-321.938	-324.306
Gen.Gamma (3P)	-321.553	-315.817	-319.37	-338.918	-333.182	-336.73	-290.95	-285.21	-291.856
Johnson SB(4P) *	-340.899	-333.251	-337.99	-381.821	-374.173	-378.91	-223.65	-216.006	-220.742

Notes. AIC stands for Akaike information criteria; BIC stands for Bayesian information criteria; HQC stands for Hannan-Quinn information criteria

* Indicates that Lognormal (2P), Johnson SB (4P) and Gamma (3P) distribution exhibited minimum values of AIC, BIC and HQC test statistics for describing flood peak, volume and duration, thus could be further indicated for the better performance

Table 5. Copula dependence parameters based on the maximum pseudo likelihood (MPL) estimator and their corresponding goodness-of-fit (GOF) test statistics via parametric bootstrap technique for the (P-V) pair of Tree 1, (P-D) pair of Tree 2, (V-D|P) pair of Tree 2

Tree: 1 (C_{PV}), Between (P-V) pair		N=1000(No. of bootstrap sampling)	
Copula family	Parameter Estimates $\hat{\theta}$	S_n	(*p*-value) $_{Sn}$
Gaussian*	0.8333772	**0.013444**	**0.9356**
Clayton	2.600312	0.035144	0.1923
Gumbel-Hougaard (GH)	2.311711	0.027751	0.2063

Frank	7.878869	0.02383	0.464
Joe	2.553838	0.083346	0.0004995

[Note: Bold letter indicated via * denotes that the Gaussian copula exhibiting minimum S_n value, which means the performance of Gaussian copula for P-V is much consistence that the other copula functions. $(\tau *)$ in the last column of above table indicated the estimated kendall's tau value from derived copulas fitted to observed random series]

Tree: 1 (C_{PD}), Between (P-D) pair		N=1000(No. of bootstrap sampling)	
Copula family	Parameter Estimates $\hat{\theta}$	S_n	(*p*-value) $_{Sn}$
Gaussian	-0.1276312	0.032132	0.486
Clayton	NA	NA	NA
Gumbel-Hougaard (GH)	NA	NA	NA
Frank*	**-0.6942**	**0.031215**	**0.4001**
Joe	2.553838	NA	NA

[Note: Bold letter indicated via * denotes that the performance of Frank copula is most satisfactory that other copulas. NA denotes that for Gumbel-Hougaard, Clayton and Joe copulas can't be used for negatively dependent data (i.e., only positively correlated random variables can be simulated [i.e., Kendall's tau > 0]]

Tree: 2 ($C_{VD|P}$), Between (VD|P) pair		N=1000(No. of bootstrap sampling)	
Copula family	Parameter Estimates $\hat{\theta}$	S_n	(*p*-value) $_{Sn}$
Gaussian	0.58356	0.02131	0.558
Clayton	0.69281	0.083018	0.006494
Gumbel-Hougaard (GH)*	**1.637**	**0.015227**	**0.9456**
Frank	3.8077	0.025713	0.474
Joe	2.005	0.01943	0.7298

[Note: Bold letter indicated via * indicates that the performance of GH copula is most satisfactory that other copulas which exhibited the minimum S_n statistics value and highest p-value.
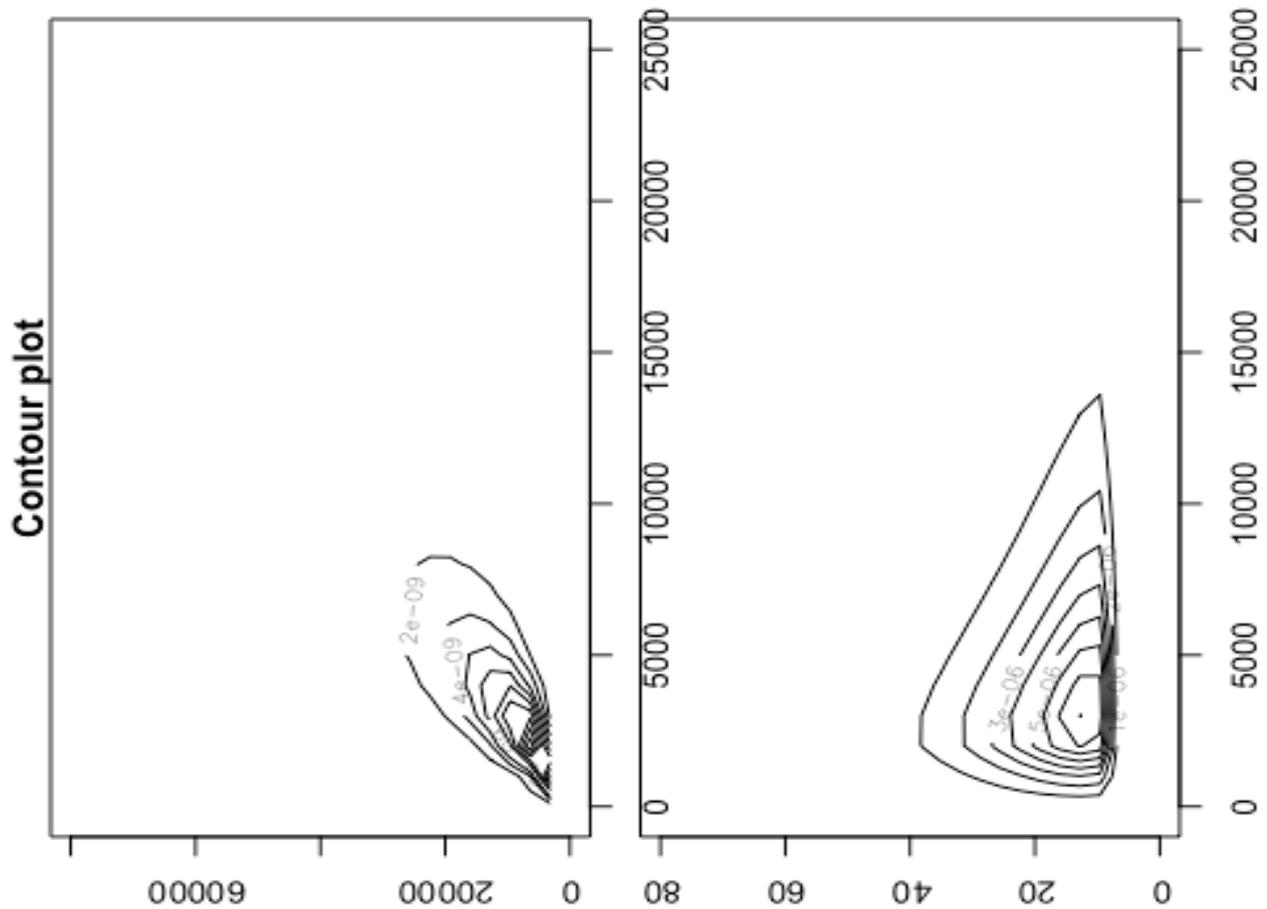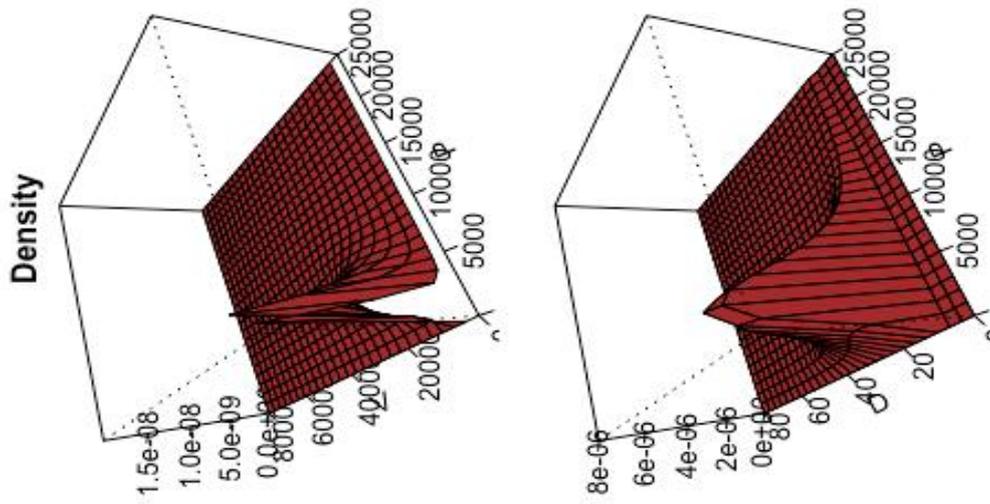
Figure 8. 3-dimension density and contour plot generated from the bivariate copula distribution between flood pairs, P-V and P-D in the first tree of vine structure
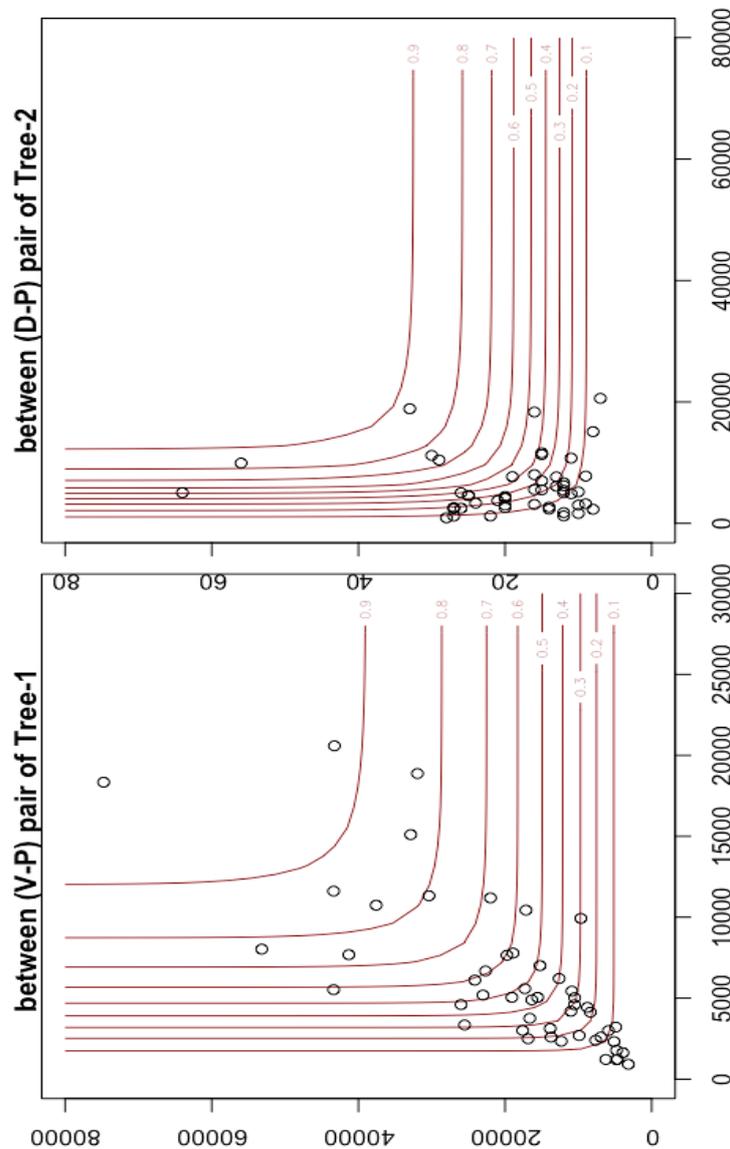
Figure 9. Bivariate contour plot derived based on the joint cumulative distribution of gaussian copula ($C_{V|P}$) and frank copula ($C_{D|P}$) fitted with flood characteristics in the first tree of vine structure (Note: black colour data represents the observed flood characteristics)
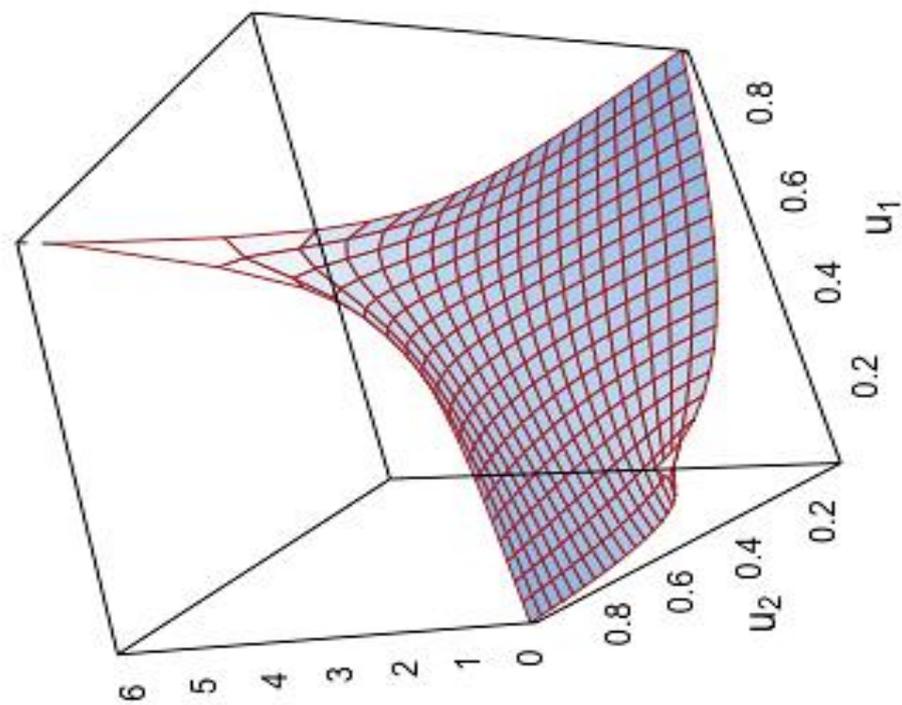
Figure 10. 3-D surface plot of bivariate copula fitted with the conditioned observations, $F_{V|P}(v|p)$ and $F_{D|P}(d|p)$ in the second tree of vine structure
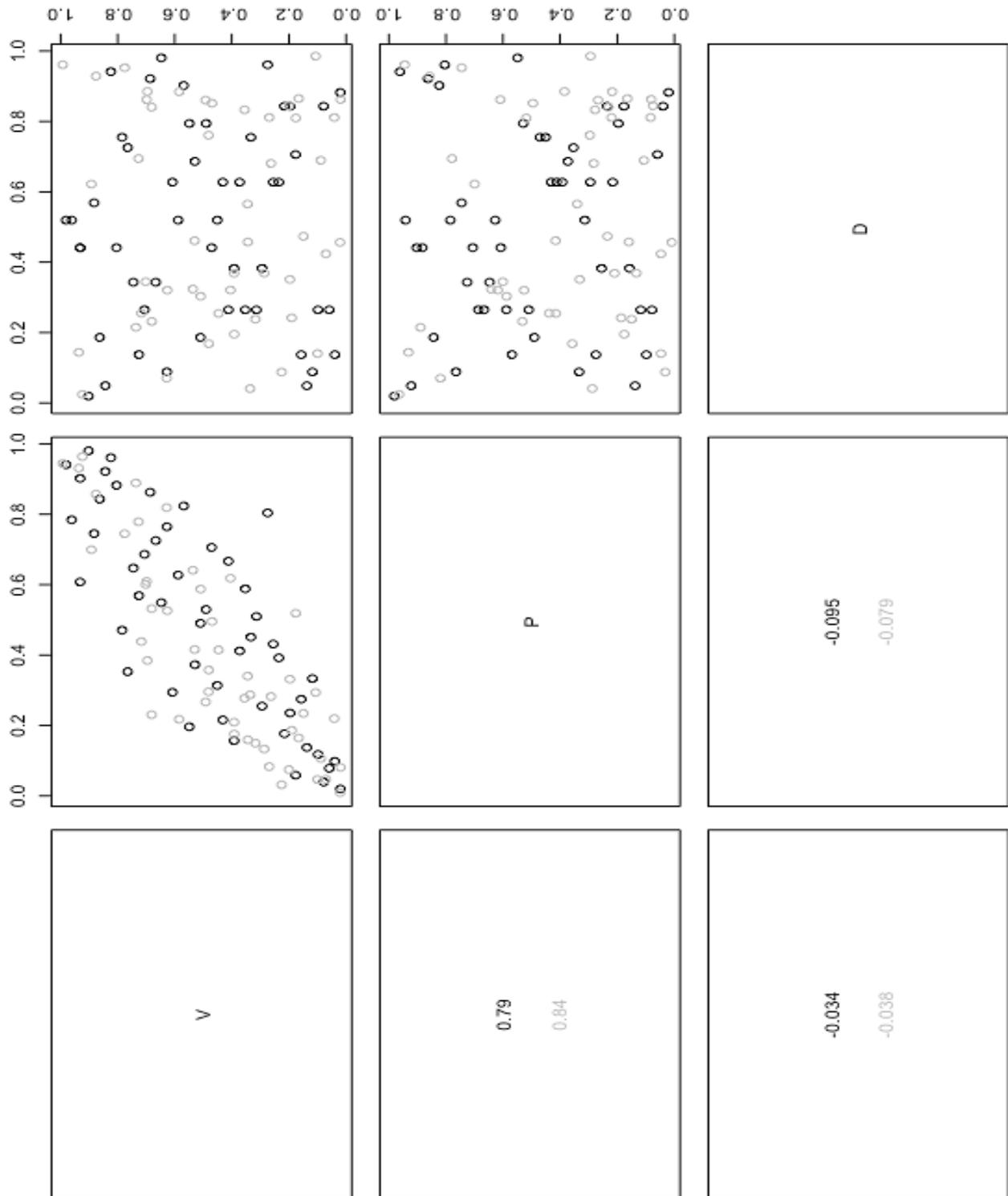
Figure 11. Comparing simulated from the vine or pair-copula models and observed flood characteristics [Note: light colour indicates simulated samples and dark colour indicates observed samples]

3.3. Estimating Return Period of Flood Characteristics

Accounting of multiple interactive flood characteristics, i.e., P, V, D through the multivariate analysis of extreme episodes is much beneficial and useful in recognizing the critical behaviour of flood episodes at a river basin scale and for making a justifiable risk-based decision making in both structural and non-structural water-related issues. According to Salvadori and De Michele [9] and Salvadori [59], the primary joint return periods are estimated using the inclusive probability such as, "OR" and "AND" case. The primary joint return period for "OR" case for the trivariate and bivariate events can be estimated as;

$$T_{(V,P,D)}^{OR}(v, p, d) = \frac{1}{1 - P(V \leq v, P \leq p, D \leq d)} = \frac{1}{1 - C(x_1, x_2, x_3)} \tag{9}$$

Where $x_1 = F_V(v), x_2 = F_P(p)$ and $x_3 = F_D(d)$ is the univariate marginal distributions of the flood characteristics and $C(x_1, x_2, x_3)$ is the trivariate copula function.

Similarly, the "OR-joint" case for bivariate events;

$$T_{(V,P)}^{OR}(v, p) = \frac{1}{1 - P(V \leq v, P \leq p)} = \frac{1}{1 - C(x_1, x_2)}, \text{ for V-P pair} \tag{10}$$

$$T_{(P,D)}^{OR}(p, d) = \frac{1}{1 - P(P \leq p, D \leq d)} = \frac{1}{1 - C(x_2, x_3)}, \text{ for P-D pair} \tag{11}$$

$$T_{(V,D)}^{OR}(v, d) = \frac{1}{1 - P(V \leq v, D \leq d)} = \frac{1}{1 - C(x_1, x_3)}, \text{ for V-D pair} \tag{12}$$

Similarly, the "AND-joint" return periods for trivariate and bivariate flood characteristics are estimated as;

$$T_{(V,P,D)}^{AND}(v, p, d) = \frac{1}{1 - F_V(v) - F_P(p) - F_D(d) + F_{V,P}(v, p) + F_{P,D}(p, d) + F_{V,D}(v, d) - F_{V,P,D}(v, p, d)}$$
$$= \frac{1}{1 - F_V(v) - F_P(p) - F_D(d) + C(x_1, x_2) + C(x_2, x_3) + C(x_1, x_3) - C(x_1, x_2, x_3)} \tag{13}$$

$$T_{(V,P)}^{AND}(v, p) = \frac{1}{1 - F_V(v) - F_P(p) + F_{V,P}(v, p)} = \frac{1}{1 - F_V(v) - F_P(p) - F_D(d) + C(x_1, x_2)} \tag{14}$$

$$T_{(P,D)}^{AND}(p, d) = \frac{1}{1 - F_P(p) - F_D(d) + F_{P,D}(p, d)} = \frac{1}{1 - F_P(p) - F_D(d) + C(x_2, x_3)} \tag{15}$$

$$T_{(V,D)}^{AND}(v, d) = \frac{1}{1 - F_V(v) - F_D(d) + F_{V,D}(v, d)} = \frac{1}{1 - F_V(v) - F_P(p) - F_D(d) + C(x_1, x_3)} \tag{16}$$

At first, the univariate return periods are obtained from the best fitted marginal distribution of the flood characteristics, i.e., lognormal-2P for peak(P), Johnson SB-4P for volume (V) and Gamma-3P for duration (V) series (see Table 6). The joint return periods for 'OR' and 'AND' case for the trivariate flood distributions using the PCC model as well as bivariate distribution are derived from the Equation (9-16), and their value are listed in the same Table 6. It is clearly revealed that the primary joint return periods in "AND" cases produce a higher return than the joint return in 'OR' cases for both the bivariate and trivariate distribution (based on PCC model). For example, consider a flood event of peak, P = 2599 m^3s^{-1}, volume, V= 6933.5 m^3 and duration, D =27 (days), the bivariate primary joint return period between P-V, P-D and V-D for 'OR' and 'AND' cases are T^{AND}_{PV} = 1.34 years and T^{OR}_{PV} = 1.146 years, T^{AND}_{PD} = 7.67 years and T^{OR}_{PD} = 1.198 years, T^{AND}_{VD} = 6.96 years and T^{OR}_{PD} = 1.16 years. Similarly, considering all the flood characteristics at the same time or simultaneously, their corresponding trivariate primary joint RPs for 'OR' and 'AND' cases are $T^{AND}_{(V,P,D)}$ = 74.04 years and $T^{OR}_{(V,P,D)}$ = 1.26 years thus clearly indicates that $T^{AND}_{(V,P,D)}$ is greater than $T^{OR}_{(V,P,D)}$.

Investigating the tail dependence properties and their coefficient is an essential aspect in the extreme value analysis otherwise might be attributed for underestimation of hydrologic risk [22]. In this study, we also investigated the tail dependence behaviour of the derived D-vine copula for the flood characteristics extracted from the fitted copulas through both the graphical and analytical observation. In the graphical approach, we visualized and compared the scatter plot, Chi-plot and Kendall or K-plot of the random observations derived from the bivariate Normal (or Gaussian) copula, C_{VP} (as fitted to (V, P) flood pair) and the Frank copula C_{PD} (as fitted to (P, D) flood pair) with the D-vine copula fitted to (V, P) and (P, D) pairs (refer to Figures 12 and 13). From the Figures 12 and 13, it is revealed that vine copula is much consistent and well capturing the general behaviour of both the Normal copula fitted for (V, P) pair and Frank copula fitted for (P, D) pair, based on the scatter plot (i.e., the first row of Figures 12 and 13). Also, based on Chi-plot and K-plot it indicates the quite similar tail dependencies behaviour of both the Normal copula and vine copula for (V, P) and the Frank copula and vine copula for (P, D) flood characteristics (see 2nd and 3rd row of Figures 12 and 13). The non-parametric estimation of tail dependence coefficient, based on the Caperaa-Fougeres-Genest estimator λ_U^{CFG} [60] is employed to estimate the tail dependence of the fitted vine model, as given below;

$$\lambda_U^{CFG} = 2 - 2\exp\left(\frac{1}{n}\Sigma_{i=1}^n \log\left[\frac{\sqrt{\log\left(\frac{1}{M_i}\right)\log\left(\frac{1}{N_i}\right)}}{\log\left(\frac{1}{\max(M_i,N_i)^2}\right)}\right]\right) \tag{17}$$

Where $(M_1, N_1), (M_2, N_2), \ldots\ldots, (M_i, N_i)$ are the random observations simulated from the PCC model fitted to the flood characteristics. Thus, using the Equation (16), the tail dependence coefficients λ_U^{PCC} (see Table 7) are estimated for each flood attributes pairs i.e., (V-P), (P-D), (V-D) based on the sample generated from fitted vine or pair-copula model.

Table 6. Comparison of return periods (RPs) for flood characteristics estimated based on the pair-copula model (trivariate RPs, i.e., T^{OR}_{PVD}, T^{AND}_{PVD}), 2-D copula models (bivariate RPs, i.e., T^{OR}_{PV}, T^{AND}_{PV} …..) and univariate marginal distributions (univariate RPs, i.e., T(P), T(V) and (TD))

P	V	D	T^{OR}_{PV}	T^{AND}_{PV}	T^{OR}_{PD}	T^{AND}_{PD}	T^{OR}_{VD}	T^{AND}_{VD}	T(P)	T(V)	T(D)	T^{OR}_{PVD}	T^{AND}_{PVD}
2597	13729.8	20	1.25	1.89	1.14	3.742	1.40	5.39	1.268	1.85	2.80	1.06	4.28
2599	6933.5	27	1.146	1.34	1.198	7.67	1.16	6.96	1.26	1.208	5.66	1.26	74.04
5189.7	23008.7	10	2.104	3.85	1.078	2.730	1.116	4.107	2.25	3.44	1.17	1.27	13.95
6111.8	24127.8	13	2.48	4.46	1.23	4.46	1.30	5.66	2.78	3.71	1.48	1.46	199.68
4895.2	16354.7	11	1.79	2.69	1.105	2.76	1.124	2.83	2.101	2.20	1.26	1.42	14.21
4191.5	11022.15	20	1.43	1.97	1.358	5.56	1.28	4.47	1.78	1.55	2.80	1.48	40.54
18339.4	74740	16	30.07	191.09	1.86	72.31	1.91	307.06	31.43	149.9	1.92	1.63	12.97

Table 7. Tail dependence coefficients for the flood attribute pairs based on the D-vine or PCC model

Flood pairs	λ_U^{PCC}
(V, P)	0.3542107
(P, D)	0.0131994
(V, D)	0.039963
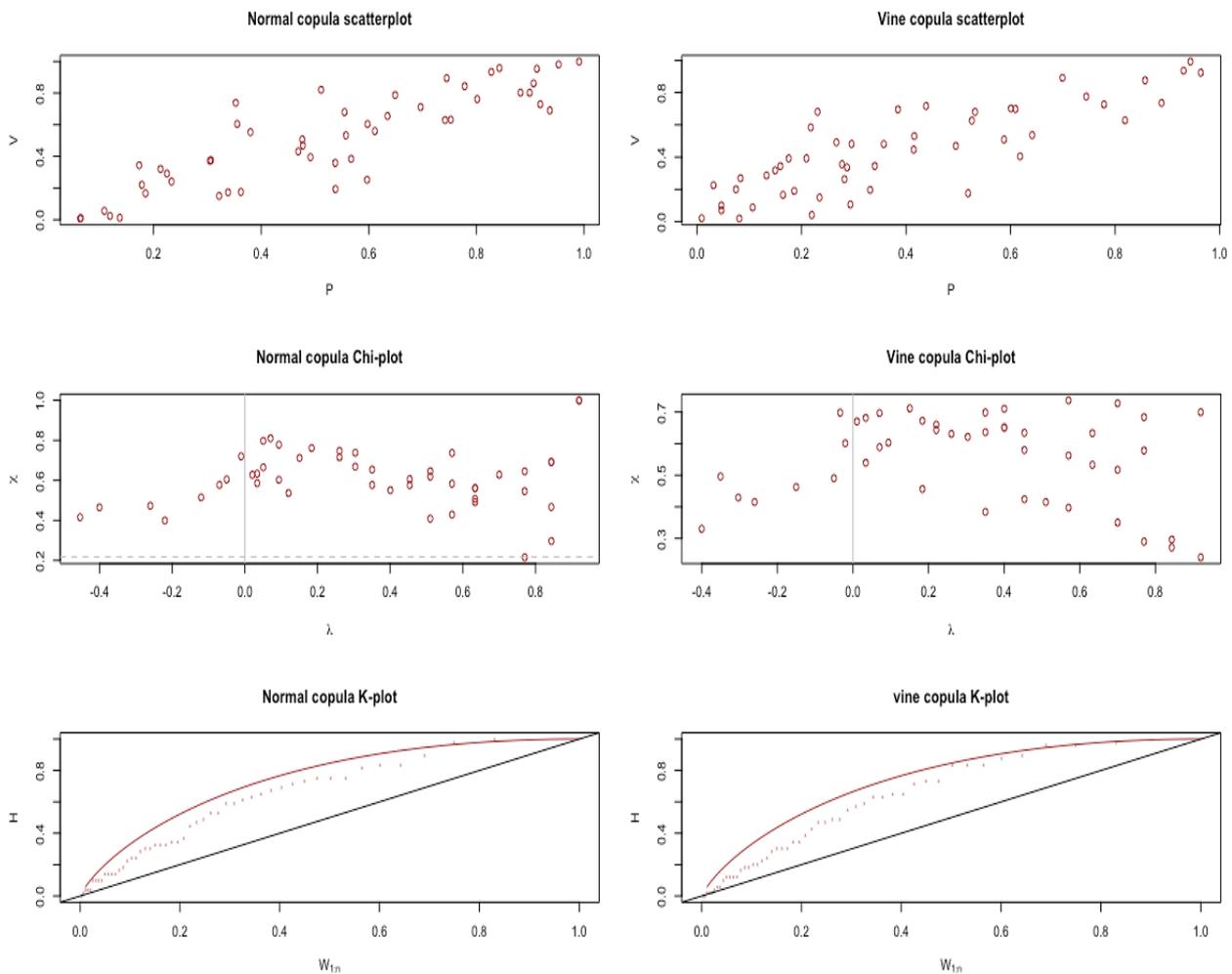
Figure 12. Scatterplot, Chi-plots and Kendall or K-Plots of the Gaussian (or Normal) copula and the D-vine copula fitted to flood characteristics (V-P)
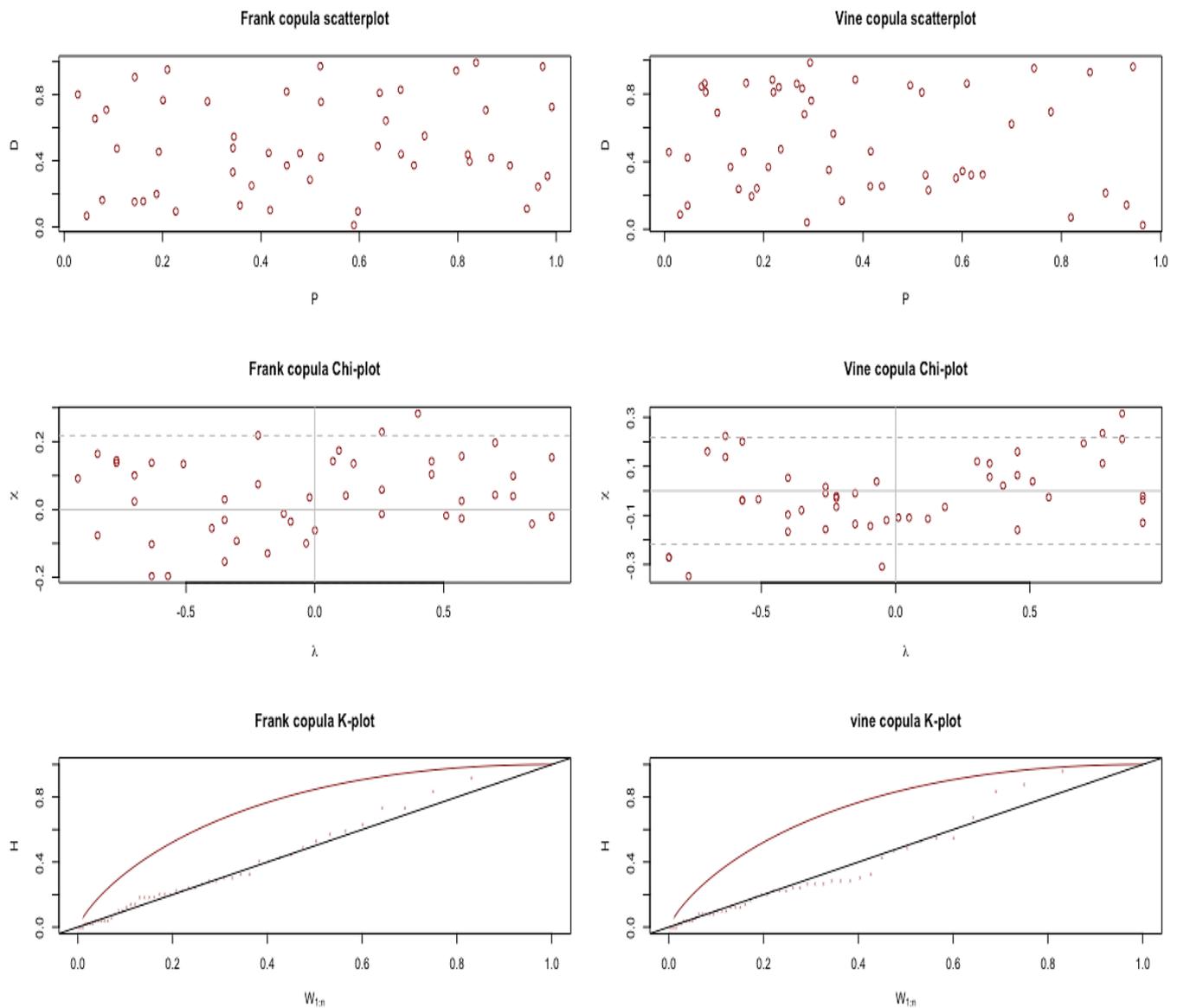
Figure 13. Scatterplot, Chi-plots and Kendall or K-Plots of the Frank copula and the D-vine copula fitted to flood characteristics (P-D)

4. Conclusions

The necessity of multivariate flood distribution analysis is often an essential concern among the hydrologist, which usually reveals an effective outcome during the assessments of hydrologic risk. The higher dimension flood distribution modelling in the light of the standard trivariate copula function usually surrounded with few statistical constraints and issues. In such circumstances, the incorporation of the vine or pair-copula construction (or PCC) could be the logical choice for revealing a comprehensive and practical approach in the uncertainty analysis and which is solely based on stagewise hierarchical nesting procedure. In this demonstration, the parametric vine copula distribution framework is incorporated for the block (annual) maxima flood characteristics collected at the Gulliemard Bridge gauge station in the Kelantan River basin in Malaysia. Tests for stationarity or existence of serial correlation (or autocorrelation) within the individual time series of flood characteristics is often a pre-requisite before introducing the random samples into a univariate or a multivariate framework. It was found that the time series of the flood characteristics do not exhibit any significant trend. Different dependency measures graphical approaches such as scatter plot, Kendall's plot, chi-plot and also analytically based correlation coefficients statistics are estimated to select the most appropriate vine tree structure. Investigation revealed that the D-vine copula could be a right choice and thus employed in constructing the hierarchical nesting structure of 3-dimensional flood characteristics where the flood variable, P is placed in between variable, V and D because the dependency between (P-V) and (P-D) is higher than (V-D). In actuality, the D-vine copula usually exhibited higher flexibility in the selection of which flood attribute pairs to model, in comparison with the canonical or C-vine copula [23, 34]. Different copulas families which include the

Gaussian copula that belongs to the Elliptical family and the Archimedean copula families such as the Gumbel, Frank, Clayton and Joe copula, are introduced and test as a possible candidate model in demonstrating the pair-wise (or bivariate) joint dependency modelling of flood characteristics, i.e., (P-V) and (P-D) in Tree-1 and also, the bivariate copulas fitted to the conditioned observations in the Tree-2 of the selected D-vine structure (see Figure 2). The vector of unknown copula dependence parameters, in both Tree-1 & Tree-2 are estimated using the Maximum-Pseudo-Likelihood (MPL) estimation procedure where the best-fitted bivariate copulas are selected using the Cramer-von Mises (CvM) test statistics 'S_n' with the parametric bootstrapping procedure.

The Lognormal-2P distribution, Johnson SB-4P distribution and Gamma-3P distribution are identified as possible univariate marginal distribution for the flood peak flow, volume and duration series based on the GOF test statistics. Firstly, the Gaussian copula and Frank copula are selected for defining the bivariate joint dependence structure of flood pairs (V, P) (i.e., 2-D copula fitted to peak-volume pair, C_{PV}) and (P, D) (i.e., 2-D copula fitted to peak-duration pair, C_{PD}) in the first tree of vine structure (see Figure 2). Similarly, the Gumbel copula is recognized as the most justifiable bivariate structure in the modelling of the variable (V, D) conditional on P (i.e., the conditioned observations, $F_{V|P}(v|p)$ and $F_{D|P}(d|p)$), in the second tree of vine framework. The scatterplot between observed and simulated flood characteristics are investigated and which indicating on an excellent agreement or in words the fitted D-vine distribution reproduces overall flood correlation structure reasonably well. The tail dependence behaviour of the fitted D-vine copula is also investigated for flood characteristics extracted from the fitted copulas through both the graphical and analytical observation and which reveals that vine copula is well capturing the general behaviour of both the fitted Normal and Frank copula for flood pairs (V, P) and (P, D). The primary joint return periods for trivariate (also, bivariate) flood events are examined and revealed that "AND" return period produces higher return value than 'OR' joint case. Unless the modelling efficacy of vine copula model is investigated in the above demonstration and from few existed kinds of literature, but it is still surrounded with the few modelling issues and constraints such that selection of best-fitted parametric copulas in the vine framework is often a tough challenge [19]. In such circumstances, the non-informative vine methodological approach, also called the minimum information-based vine construction could facilitate higher flexibility and accentuate the potential of traditional parametric vine approach, especially when the length of the observation samples under consideration are limited [13]. Daneshkhan et al. [19] study already demonstrated the performance of minimum information pair-copula and thus revealed that it could be possible to approximate multivariate density to any level of desired precision. All the above-raised issues can be considered as a future research purpose for the same river basin in order to achieve much comprehensive approximation of joint probability density and practical flood hazard assessments.

5. Conflicts of Interest

The authors declare no conflict of interest.

6. References

[1] Yue, Sheng. "The bivariate lognormal distribution to model a multivariate flood episode." Hydrological Processes 14, no. 14 (2000): 2575-2588. doi:10.1002/1099-1085(20001015)14:14<2575::AID-HYP115>3.0.CO;2-L.

[2] Zhang, L., and Vijay P. Singh. "Trivariate flood frequency analysis using the Gumbel–Hougaard copula." Journal of Hydrologic Engineering 12, no. 4 (2007): 431-439. doi:10.1061/(asce)1084-0699(2007)12:4(431).

[3] Gräler, B., M. J. van den Berg, S. Vandenberghe, A. Petroselli, S. Grimaldi, B. De Baets, and N. E. C. Verhoest. "Multivariate Return Periods in Hydrology: a Critical and Practical Review Focusing on Synthetic Design Hydrograph Estimation." Hydrology and Earth System Sciences 17, no. 4 (April 2, 2013): 1281–1296. doi:10.5194/hess-17-1281-2013.

[4] Cunnane, Conleth. "Review of Statistical Models for Flood Frequency Estimation." Hydrologic Frequency Modeling (1987): 49–95. doi:10.1007/978-94-009-3953-0_4.

[5] Nadarajah, S., and J. T. Shiau. "Analysis of Extreme Flood Events for the Pachang River, Taiwan." Water Resources Management 19, no. 4 (August 2005): 363–374. doi:10.1007/s11269-005-2073-2.

[6] Yue, S., and C. Y. Wang. "A Comparison of Two Bivariate Extreme Value Distributions." Stochastic Environmental Research and Risk Assessment (SERRA) 18, no. 2 (April 1, 2004): 61–66. doi:10.1007/s00477-003-0124-x.

[7] Sklar, M. "Fonctions de repartition an dimensions et leurs marges." Publ. inst. statist. univ. Paris 8 (1959): 229-231.

[8] Nelsen, R. B. "An Introduction to Copulas." Springer Series in Statistics (2006). doi:10.1007/0-387-28678-0.

[9] Salvadori, G., and C. De Michele. "Frequency Analysis via Copulas: Theoretical Aspects and Applications to Hydrological Events." Water Resources Research 40, no. 12 (December 2004). doi:10.1029/2004wr003133.

[10] De Michele, C. "A Generalized Pareto Intensity-Duration Model of Storm Rainfall Exploiting 2-Copulas." Journal of Geophysical Research 108, no. D2 (2003). doi:10.1029/2002jd002534.

[11] Zhang, L. S. V. P., and V. P. Singh. "Bivariate flood frequency analysis using the copula method." Journal of hydrologic engineering 11, no. 2 (2006): 150-164. doi:10.1061/(asce)1084-0699(2006)11:2(150).

[12] Salvadori, G., and C. De Michele. "Statistical Characterization of Temporal Structure of Storms." Advances in Water Resources 29, no. 6 (June 2006): 827–842. doi:10.1016/j.advwatres.2005.07.013.

[13] Wong, G., M. F. Lambert, M. Leonard, and A. V. Metcalfe. "Drought Analysis Using Trivariate Copulas Conditional on Climatic States." Journal of Hydrologic Engineering 15, no. 2 (February 2010): 129–141. doi:10.1061/(asce)he.1943-5584.0000169.

[14] Genest, Christian, and Anne-Catherine Favre. "Everything you always wanted to know about copula modeling but were afraid to ask." Journal of hydrologic engineering 12, no. 4 (2007): 347-368. doi:10.1061/(asce)1084-0699(2007)12:4(347).

[15] Reddy, M. Janga, and Poulomi Ganguli. "Bivariate Flood Frequency Analysis of Upper Godavari River Flows Using Archimedean Copulas." Water Resources Management 26, no. 14 (September 11, 2012): 3995–4018. doi:10.1007/s11269-012-0124-z.

[16] Renard, B., and M. Lang. "Use of a Gaussian Copula for Multivariate Extreme Value Analysis: Some Case Studies in Hydrology." Advances in Water Resources 30, no. 4 (April 2007): 897–912. doi:10.1016/j.advwatres.2006.08.001.

[17] Salvadori, G., C. De Michele, and F. Durante. "On the Return Period and Design in a Multivariate Framework." Hydrology and Earth System Sciences 15, no. 11 (November 4, 2011): 3293–3305. doi:10.5194/hess-15-3293-2011.

[18] Ganguli, Poulomi, and M. Janga Reddy. "Probabilistic Assessment of Flood Risks Using Trivariate Copulas." Theoretical and Applied Climatology 111, no. 1–2 (May 25, 2012): 341–360. doi:10.1007/s00704-012-0664-4.

[19] Daneshkhah, Alireza, Renji Remesan, Omid Chatrabgoun, and Ian P. Holman. "Probabilistic Modeling of Flood Characterizations with Parametric and Minimum Information Pair-Copula Model." Journal of Hydrology 540 (September 2016): 469–487. doi:10.1016/j.jhydrol.2016.06.044.

[20] Madadgar, Shahrbanou, and Hamid Moradkhani. "Drought Analysis Under Climate Change Using Copula." Journal of Hydrologic Engineering 18, no. 7 (July 2013): 746–759. doi:10.1061/(asce)he.1943-5584.0000532.

[21] Li, Fan, and Qian Zheng. "Probabilistic Modelling of Flood Events Using the Entropy Copula." Advances in Water Resources 97 (November 2016): 233–240. doi:10.1016/j.advwatres.2016.09.016.

[22] Joe, Harry. "Multivariate Models and Multivariate Dependence Concepts." C&H/CRC Monographs on Statistics & Applied Probability (May 1997). doi:10.1201/b13150.

[23] Aas, Kjersti, Claudia Czado, Arnoldo Frigessi, and Henrik Bakken. "Pair-Copula Constructions of Multiple Dependence." Insurance: Mathematics and Economics 44, no. 2 (April 2009): 182–198. doi:10.1016/j.insmatheco.2007.02.001.

[24] Grimaldi, Salvatore, and Francesco Serinaldi. "Asymmetric Copula in Multivariate Flood Frequency Analysis." Advances in Water Resources 29, no. 8 (August 2006): 1155–1167. doi:10.1016/j.advwatres.2005.09.005.

[25] Serinaldi, Francesco, and Salvatore Grimaldi. "Fully nested 3-copula: procedure and application on hydrological data." Journal of Hydrologic Engineering 12, no. 4 (2007): 420-430. doi:10.1061/(asce)1084-0699(2007)12:4(420)

[26] Genest, C., A.-C. Favre, J. Béliveau, and C. Jacques. "Metaelliptical Copulas and Their Use in Frequency Analysis of Multivariate Hydrological Data." Water Resources Research 43, no. 9 (September 2007). doi:10.1029/2006wr005275.

[27] Kao, Shih-Chieh, and Rao S. Govindaraju. "Trivariate Statistical Analysis of Extreme Rainfall Events via the Plackett Family of Copulas." Water Resources Research 44, no. 2 (February 2008). doi:10.1029/2007wr006261.

[28] Song, Songbai, and Vijay P. Singh. "Meta-Elliptical Copulas for Drought Frequency Analysis of Periodic Hydrologic Data." Stochastic Environmental Research and Risk Assessment 24, no. 3 (July 17, 2009): 425–444. doi:10.1007/s00477-009-0331-1.

[29] Bedford, Tim, and Roger M. Cooke. "Vines--a New Graphical Model for Dependent Random Variables." The Annals of Statistics 30, no. 4 (August 2002): 1031–1068. doi:10.1214/aos/1031689016.

[30] Czado, Claudia, and Aleksey Min. "Bayesian Inference for D-Vines: Estimation and Model Selection." Dependence Modeling (December 2010): 249–264. doi:10.1142/9789814299886_0012.

[31] Abdulrazzaq, Zaidoon T., Raghad H. Hasan, and Nadia A. Aziz. "Integrated TRMM Data and Standardized Precipitation Index to Monitor the Meteorological Drought." Civil Engineering Journal 5, no. 7 (July 21, 2019): 1590–1598. doi:10.28991/cej-2019-03091355.

[32] Gyasi-Agyei, Yeboah, and Charles S. Melching. "Modelling the Dependence and Internal Structure of Storm Events for Continuous Rainfall Simulation." Journal of Hydrology 464–465 (September 2012): 249–261. doi:10.1016/j.jhydrol.2012.07.014.

[33] Vernieuwe, H., S. Vandenberghe, B. De Baets, and N. E. C. Verhoest. "A Continuous Rainfall Model Based on Vine Copulas." Hydrology and Earth System Sciences 19, no. 6 (June 12, 2015): 2685–2699. doi:10.5194/hess-19-2685-2015.

[34] Czado, Claudia, Stephan Jeske, and Mathias Hofmann. "Selection strategies for regular vine copulae." Journal de la Société Française de Statistique 154, no. 1 (2013): 174-191.

[35] Hobæk Haff, Ingrid, Kjersti Aas, and Arnoldo Frigessi. "On the Simplified Pair-Copula Construction — Simply Useful or Too Simplistic?" Journal of Multivariate Analysis 101, no. 5 (May 2010): 1296–1310. doi:10.1016/j.jmva.2009.12.001.

[36] Kurowicka, Dorota, and Roger Cooke. "Uncertainty Analysis with High Dimensional Dependence Modelling." Wiley Series in Probability and Statistics (February 24, 2006). doi:10.1002/0470863072.

[37] Aas, Kjersti, and Daniel Berg. "Models for Construction of Multivariate Dependence – a Comparison Study." The European Journal of Finance 15, no. 7–8 (December 2009): 639–659. doi:10.1080/13518470802588767.

[38] Drainage, D. I. D. "Irrigation Department. Annual flood report of DID for Peninsular Malaysia." Unpublished report. Kuala Lumpur: DID (2004).

[39] Adnan, Nor Aizam, and Peter M. Atkinson. "Exploring the Impact of Climate and Land Use Changes on Streamflow Trends in a Monsoon Catchment." International Journal of Climatology 31, no. 6 (April 15, 2011): 815–831. doi:10.1002/joc.2112.

[40] Nashwan, Mohamed Salem, Tarmizi Ismail, and Kamal Ahmed. "Flood Susceptibility Assessment in Kelantan River Basin Using Copula." International Journal of Engineering & Technology 7, no. 2 (April 18, 2018): 584. doi:10.14419/ijet.v7i2.10447.

[41] Shiau, J. T. "Fitting Drought Duration and Severity with Two-Dimensional Copulas." Water Resources Management 20, no. 5 (June 27, 2006): 795–815. doi:10.1007/s11269-005-9008-9.

[42] Favre, Anne-Catherine, Salaheddine El Adlouni, Luc Perreault, Nathalie Thiémonge, and Bernard Bobée. "Multivariate Hydrological Frequency Analysis Using Copulas." Water Resources Research 40, no. 1 (January 2004). doi:10.1029/2003wr002456.

[43] Zhang, Runrun, Xi Chen, Qinbo Cheng, Zhicai Zhang, and Peng Shi. "Joint Probability of Precipitation and Reservoir Storage for Drought Estimation in the Headwater Basin of the Huaihe River, China." Stochastic Environmental Research and Risk Assessment 30, no. 6 (April 11, 2016): 1641–1657. doi:10.1007/s00477-016-1249-z.

[44] De Michele, C., G. Salvadori, M. Canossi, A. Petaccia, and R. Rosso. "Bivariate statistical approach to check adequacy of dam spillway." Journal of Hydrologic Engineering 10, no. 1 (2005): 50-57. doi:10.1061/(asce)1084-0699(2005)10:1(50).

[45] Kojadinovic, Ivan, and Jun Yan. "Modeling Multivariate Distributions with Continuous Margins Using thecopulaRPackage." Journal of Statistical Software 34, no. 9 (2010). doi:10.18637/jss.v034.i09.

[46] Genest, Christian, and Bruno Rémillard. "Validity of the Parametric Bootstrap for Goodness-of-Fit Testing in Semiparametric Models." Annales de l'Institut Henri Poincaré, Probabilités et Statistiques 44, no. 6 (December 2008): 1096–1127. doi:10.1214/07-aihp148.

[47] Genest, Christian, Bruno Rémillard, and David Beaudoin. "Goodness-of-Fit Tests for Copulas: A Review and a Power Study." Insurance: Mathematics and Economics 44, no. 2 (April 2009): 199–213. doi:10.1016/j.insmatheco.2007.10.005.

[48] Alamgir, Mahiuddin, Tarmizi Ismail, and Muhammad Noor. "Bivariate Frequency Analysis of Flood Variables Using Copula in Kelantan River Basin." Malaysian Journal of Civil Engineering 30, no. 3 (November 19, 2018). doi:10.11113/mjce.v30n3.515.

[49] Hosking, J. R. M., J. R. Wallis, and E. F. Wood. "Estimation of the Generalized Extreme-Value Distribution by the Method of Probability-Weighted Moments." Technometrics 27, no. 3 (August 1985): 251–261. doi:10.1080/00401706.1985.10488049.

[50] Ljung, G. M., and G. E. P. Box. "On a Measure of Lack of Fit in Time Series Models." Biometrika 65, no. 2 (August 1, 1978): 297–303. doi:10.1093/biomet/65.2.297.

[51] Benth, Fred Espen, and Jūratė Šaltytė‐Benth. "Stochastic Modelling of Temperature Variations with a View Towards Weather Derivatives." Applied Mathematical Finance 12, no. 1 (March 2005): 53–85. doi:10.1080/1350486042000271638.

[52] Fisher, N. I, and P Switzer. "Graphical Assessment of Dependence." The American Statistician 55, no. 3 (August 2001): 233–239. doi:10.1198/000313001317098248.

[53] Genest, Christian, and Jean-Claude Boies. "Detecting Dependence With Kendall Plots." The American Statistician 57, no. 4 (November 2003): 275–284. doi:10.1198/0003130032431.

[54] Gringorten, Irving I. "A Plotting Rule for Extreme Probability Paper." Journal of Geophysical Research 68, no. 3 (February 1, 1963): 813–814. doi:10.1029/jz068i003p00813.

[55] Anderson, T. W., and D. A. Darling. "A Test of Goodness of Fit." Journal of the American Statistical Association 49, no. 268 (December 1954): 765–769. doi:10.1080/01621459.1954.10501232.

[56] Akaike, H. "A New Look at the Statistical Model Identification." IEEE Transactions on Automatic Control 19, no. 6 (December 1974): 716–723. doi:10.1109/tac.1974.1100705.

[57] Schwarz, Gideon. "Estimating the Dimension of a Model." The Annals of Statistics 6, no. 2 (March 1978): 461–464. doi:10.1214/aos/1176344136.

[58] Hannan, E. J., and B. G. Quinn. "The Determination of the Order of an Autoregression." Journal of the Royal Statistical Society: Series B (Methodological) 41, no. 2 (January 1979): 190–195. doi:10.1111/j.2517-6161.1979.tb01072.x.

[59] Salvadori, G. "Bivariate Return Periods via 2-Copulas." Statistical Methodology 1, no. 1–2 (December 2004): 129–144. doi:10.1016/j.stamet.2004.07.002.

[60] Caperaa, P. "A Nonparametric Estimation Procedure for Bivariate Extreme Value Copulas." Biometrika 84, no. 3 (September 1, 1997): 567–577. doi:10.1093/biomet/84.3.567.

[61] Bedford, Tim, Alireza Daneshkhah, and Kevin J. Wilson. "Approximate Uncertainty Modeling in Risk Analysis with Vine Copulas." Risk Analysis 36, no. 4 (September 2, 2015): 792–815. doi:10.1111/risa.12471.