



## Bivariate Hydrologic Risk Assessment of Flood Episodes using the Notation of Failure Probability

Shahid Latif <sup>a\*</sup>, Firuza Mustafa <sup>b</sup>

<sup>a</sup> PhD Research Scholar, Department of Geography, University of Malaya, Kuala Lumpur 50603, Malaysia.

<sup>b</sup> Associate Professor, Department of Geography, University of Malaya, Kuala Lumpur 50603, Malaysia.

Received 13 July 2020; Accepted 15 September 2020

### Abstract

Floods are becoming the most severe and challenging hydrologic issue at the Kelantan River basin in Malaysia. Flood episodes are usually thoroughly characterized by flood peak discharge flow, volume and duration series. This study incorporated the copula-based methodology in deriving the joint distribution analysis of the annual flood characteristics and the failure probability for assessing the bivariate hydrologic risk. Both the Archimedean and Gaussian copula family were introduced and tested as possible candidate functions. The copula dependence parameters are estimated using the method-of-moment estimation procedure. The Gaussian copula was recognized as the best-fitted distribution for capturing the dependence structure of the flood peak-volume and peak-duration pairs based on goodness-of-fit test statistics and was further employed to derive the joint return periods. The bivariate hydrologic risks of flood peak flow and volume pair, and flood peak flow and duration pair in different return periods (i.e., 5, 10, 20, 50 and 100 years) were estimated and revealed that the risk statistics incrementally increase in the service lifetime and, at the same instant, incrementally decrease in return periods. In addition, we found that ignoring the mutual dependency can underestimate the failure probabilities where the univariate events produced a lower failure probability than the bivariate events. Similarly, the variations in bivariate hydrologic risk with the changes of flood peak in the different synthetic flood volume and duration series (i.e., 5, 10, 20, 50 and 100 years return periods) under different service lifetimes are demonstrated. Investigation revealed that the value of bivariate hydrologic risk statistics incrementally increases over the project lifetime (i.e., 30, 50, and 100 years) service time, and at the same time, it incrementally decreases in the return period of flood volume and duration. Overall, this study could provide a basis for making an appropriate flood defence plan and long-lasting infrastructure designs.

*Keywords:* Flood; Copula Function; Bivariate Hydrologic Risk; Return Period; Failure Probability.

### 1. Introduction

Nowadays, flood events are characterized as one of the most severe and disastrous naturally occurring hydrologic consequences across the world, and the risk of their occurrence will increase in the future due to the global and regional climate-changing scenario [1-3]. In the operational planning, management or flood defence infrastructure design of water resources, it is often demanding to accurately estimate the flow exceedance probability or return periods for assessing the hydrologic risk. Flood frequency analysis or FFA is a statistical approach to establishing an interlink between the magnitude of flood episodes (or flood design quantiles) and their return periods (or non-exceedance probability) using the most logical and parsimonious probability distribution functions [4-6]. Flood is a

\* Corresponding author: [macet.shahid@gmail.com](mailto:macet.shahid@gmail.com)

 <http://dx.doi.org/10.28991/cej-2020-03091599>



© 2020 by the authors. Licensee C.E.J, Tehran, Iran. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (<http://creativecommons.org/licenses/by/4.0/>).

multidimensional stochastic consequences, often defined thoroughly and comprehensively through peak discharge flow, volume and duration series [7-9]. Because of the multidimensional characteristics of a flood episode, the reliability of univariate return periods in the hydrologic risk assessments is questionable and thus point towards the feasibility and effectiveness of multivariate distribution modelling through integrating the multiple intercorrelated flood random vectors [10-12]. In actuality, the potential damage could likely be a function of several intercorrelated flood characteristics, such that ignoring the spatial dependency among them might contribute to underestimation of the uncertainty distributed over the estimated flood design quantiles, and thus, often demands more flood characteristics, i.e. based on the joint probability density function (or JPDF) and joint cumulative distribution functions (or JCDF) [13]. This is especially so from the prospect of hydraulic or flood defence infrastructure design procedures, where the accountability of multivariate design parameters could be a feasible desire [14, 15]. In recent decades, the copula function has been identified as the most dynamic multivariate statistical framework or tool in comparison with traditional multivariate functions, due to its flexibility to select univariate marginal distributions and their joint dependence structure separately [16-18]. Numerous studies have performed the copula-based multivariate joint analysis of the flood characteristics and have estimated the design variable quantiles under different notations of return periods, i.e., based on joint distribution, conditional joint distribution or Kendall's distribution [14, 15, 19 and references therein].

The hydrologist or water experts are often interested in the evaluation of the mean inter-arrival time between the two successive occurrences of hydrologic episodes, also called return periods (or RPs). According to Brunner [20] and Latif and Firuza [21], the selection of an appropriate or justifiable RPs in solving the different structural or either non-structural water-related queries is usually depending upon the importance of undertaken structure as well as its consequences of failure. On the otherside, the design quantiles define a higher RPs often seems a much practical approach in the hydraulic structure designs [22]. According to studies such as Salvadori et al. [23], Huang and Chen [24], Read and Vogel [25], and Moftakhari et al. [26], the traditional definition of the return periods (or RPs) would be ineffective and incapable of demonstrating the risk of potential flood hazards during the entire project design lifetime. Literature, such as Salvadori et al. [27] and Serinaldi [28] suggested the importance and necessity of the failure probability (FP) statistics, which can effectively quantify the risk of hydrologic episodes. Serinaldi [28] study pointed out that the definition of FP, which is usually defined over a given design life period, facilitates a more consistent and well-devised approach for assessing the risk of extreme flood episodes. Similarly, a study conducted by Read and Vogel [25] revealed that the definition and applicability of average return period (i.e. joint return period) did not account for the planning horizon and would be problematic. Besides this, a few other studies, such as Xu et al. [29], have performed the bivariate hydrologic risk of the multivariate flood characteristics for the Wei River basin using the concept of failure probability, which were derived using the bivariate copula distribution function. Moftakhari et al. [26] performed compound flood modelling caused by the coastal water level and fluvial flow (or river discharge) under different future sea-level scenarios, where the concept of failure probability was employed for assessing the risk. Recently, Xu et al. [30] performed copula-based joint modelling between rainfall and storm tide events, where the variation of bivariate flood risk was examined using the failure probability statistics.

Flooding has become one of the most critical hydrologic issues in the Kelantan River Basin, where the conditions are becoming more severe and intense year after year [31, 32]. Examples of extreme flood scenarios due to intense and prolonged precipitation occurred in 2002, which affected a total population of about 714,287 people, or in December 2014, on the east coast of the Kelantan River basin, which affected more than 200,000 people. This river basin receives an annual rainfall of about 2500 mm, much of which occurs during the north-east monsoon (or wet season) between mid-October and mid-January. A study conducted by Hussian and Ismail [33] pointed out that the flood frequency risk is higher at the Gulliemard Bridge, Lebir and Galas gauge station. Similarly, Nashwan et al. [34] also revealed that the downstream area of the Kelantan River Basin has the highest risk of devastating flood events. Similarly, a study performed by Adnan and Atkinson [35] pointed out the existence of a significant trend in streamflow samples for both the upstream (i.e., Galas River) and downstream (i.e., Kelantan River) sub-catchments such that in the downstream area streamflow increased in the north-east monsoon circulation. Also, the expectation of the occurrence of catastrophic flooding has increased from once in every 50 to 15 years from 2004 in the Kelantan region in Malaysia [36].

The cause of frequent failure of hydrologic or flood defence infrastructure in Malaysia due to the impact of moderately severe of flood episodes might be attributed due to the lack of complete flood hydrograph, where only flood peak discharge samples often targeted in deriving flood frequency curve during the structural development. The incorporation of the copula-based methodology is still very rare in Malaysia. Recently, few copula-based multivariate analysis and modelling of flood episodes are performed for this river basin such as Latif and Firuza [37], incorporated the vine or pair-copula construction (or PPC) in the trivariate distribution modelling of the flood characteristics. In this demonstration, the D-vine tree structure was selected in establishing the trivariate dependency analysis and deriving both the primary OR and AND-joint return periods. Similarly, Latif and Firuza [38] introduced the copula-based bivariate joint analysis of the annual flood characteristics under the semiparametric framework where a distinguish

varieties of nonparametric kernel density functions were introduced and tested in the modelling of the univariate marginal distribution of the flood characteristics. In another demonstration, Latif and Firuza [39] performed and tested the efficacy of the Archimedean copulas and Gaussian copula in the trivariate distribution analysis of flood characteristics for the same river basin. On the other side, Latif and Firuza [40], also performed the bivariate flood analysis using the nonparametric copula distribution framework. In such incorporation, a combination of both parametric and nonparametric marginal distribution separately conjoined by a nonparametric copula framework, which was based on the Beta kernel function. The Beta kernel copula function was incorporated to estimate bivariate copula density which further employed in deriving the joint cumulative density of flood peak-volume, volume-duration and peak-duration pairs and their associated joint as well as conditional return periods.

In this study, we incorporated and tested the Archimedean and Gaussian copulas in deriving the bivariate distribution analysis of the annual series between flood peak and volume (P-V) and flood peak flow and duration (P-D) pairs. Then the best-fitted parametric copula for each flood pair were employed in deriving the joint probability distribution and their associated return periods, i.e., ‘OR’ and ‘AND’ joint return periods, which were further employed in the estimation of failure probability or FP statistics. The FP statistics were estimated to assess the variations in the level of bivariate hydrologic risk for the flood episodes during the entire project design lifetime. In this study, the bivariate hydrologic risk analysis was carried out through two different investigation procedures: (1) demonstrating the hydrologic risk of the flood peak flow and volume pair (P-V) and flood peak flow and duration pair (P-D) in different return periods; and (2) visualizing the variation in hydrologic risk with changes of flood peak flow in different designed or synthetic flood volumes and durations under different service lifetimes. The copula-based bivariate distribution framework is applied as a case study for the daily basis streamflow discharge records, which were collected from the period 1961–2016 for the Kelantan River basin at the Gulliemard Bridge gauge station and delivered by the Department of Irrigation and Drainage (DID), Malaysia. This investigation will be useful for the hydrologist and water engineer in the prevention and management of flood episodes. Figure 1 illustrates the methodological flow diagram indicating the steps of the analysis in the present demonstration. The structure of this manuscript is organized as follows. Details of the study area and method for retrieving or delineating the trivariate flood characteristics are discussed in the next section (i.e. Section 2). A brief theoretical discussion about the copula-based bivariate joint distribution analysis of the flood characteristics also, briefing of the bivariate hydrologic risk using the concept of the failure probability (or FP) statistics are discussed in the third section (i.e. Section 3). In actuality, the FP statistics are derived for both the primary, OR-joint and AND-joint return period using the best-fitted bivariate copulas for the flood attribute pairs. The fourth section (i.e. Section 4) provides the result of the findings and discussions. The fifth section (Section 5) provides the research conclusions.

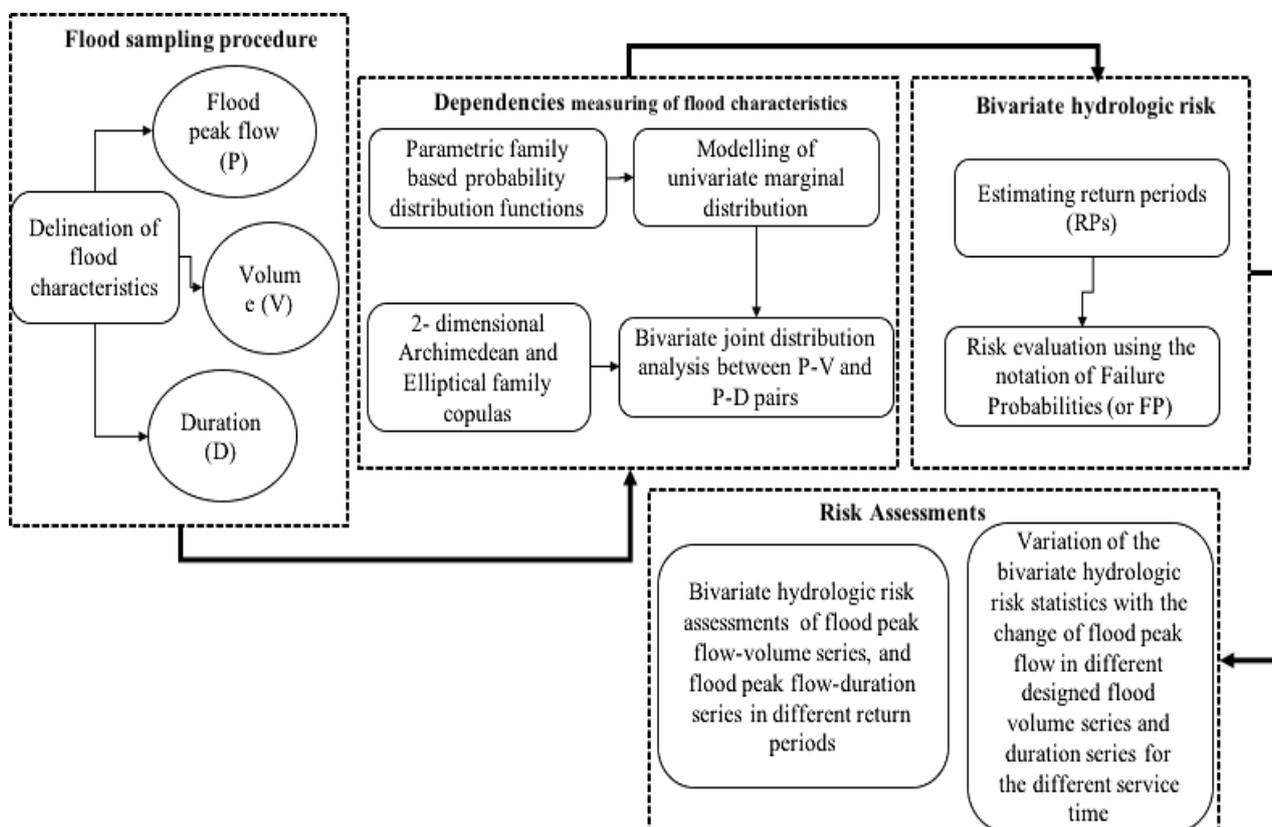


Figure 1. Methodological flow diagram of copula-based bivariate risk assessments

## 2. Details of Study Area and Delineation of Flood Characteristics

The origin of the Kelantan River basin is the Tahan mountain range, which finally drains into the South China Sea, located in the north-eastern part of Peninsular Malaysia and occupying more than 85% of the state of Kelantan. The length of this river basin is about 248 km long and the total drain area is about 13,100 km<sup>2</sup>. The geographical extent of this river basin is Lat 4° 30' N to 6° 15' N and Long 101°E to 102° 45' E. Annual precipitation varies between 0 mm (i.e. dry period) and 1750 mm (i.e. wet or north-eastern monsoonal period) and estimated runoff of this region is about 500 m<sup>3</sup>sec<sup>-1</sup> [41]. In this study, the daily streamflow discharge observations were collected at the Gulliemard Bridge gauge station, which is located near the Kuala Kari region (i.e. downstream) of the Kelantan River and provided by the DID (see Figure 2). Agriculture activities such as paddy, rubber and palm oil are the major land use in the midstream region, and forest in the upstream region of the river basin. Due to this rapid human intervention from natural to land-use activities would be responsible for these extreme hydrologic consequences [35, 36].

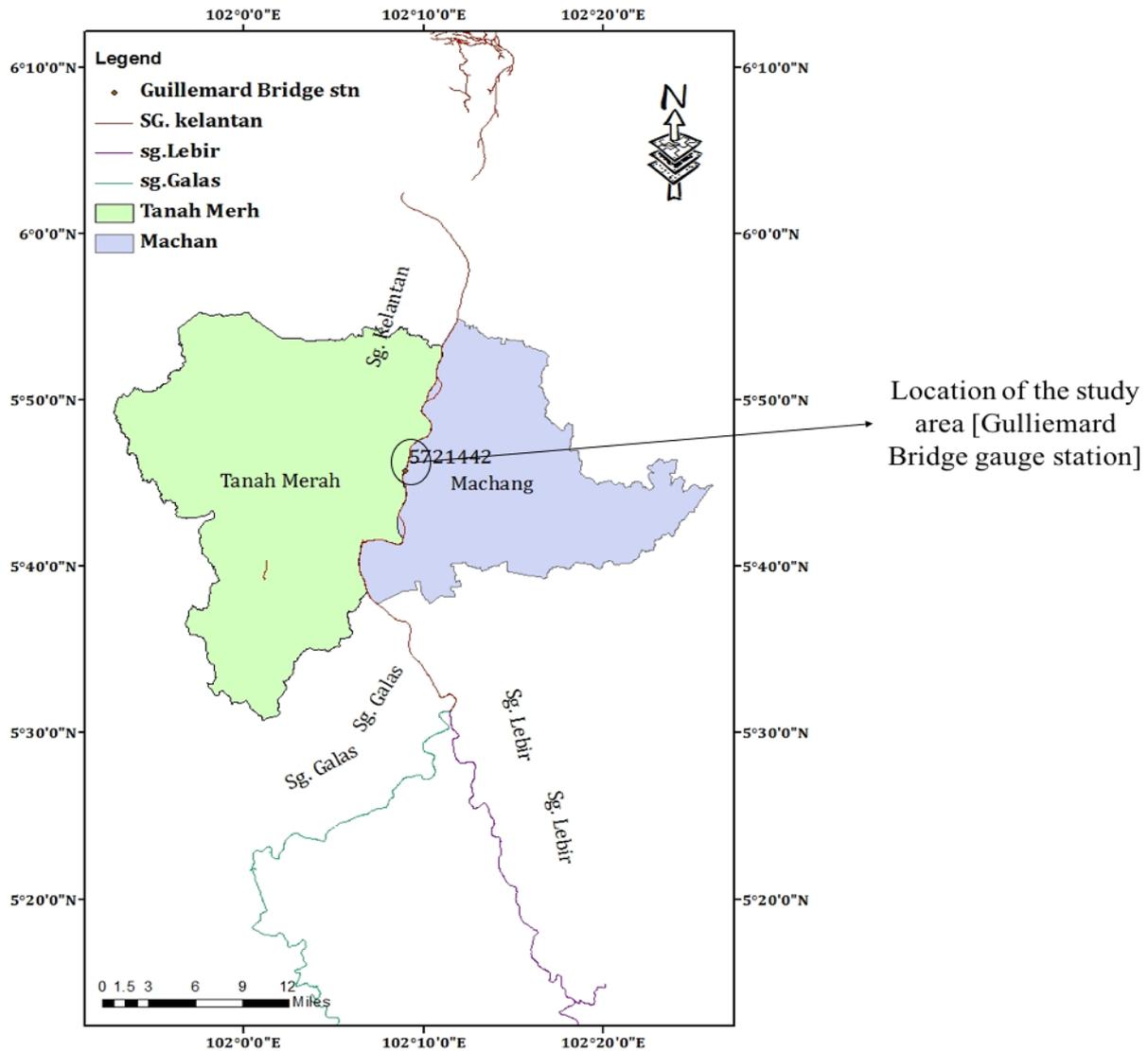


Figure 2. Geographical location of the study area (Gulliemard Bridge gauge station located at the downstream region of Kelantan River Basin in Malaysia)

The annual (maximum) series (AM), also called the block (annual) maxima, and peak over threshold (or POT) are the two holistic approaches of a partial data series that are widely accepted in extreme event modelling [42, 43]. In actuality, the partial data series only focuses on the extreme hydrograph portion, i.e. either high flow (i.e. for flood episodes) or low flow (i.e. for drought events), instead of visualizing the entire hydrograph. In this demonstration, we prefer the AM approach in the sampling of triplet flood characteristics, i.e., the flood peak discharge flow, volume and duration series, which are retrieved from the daily streamflow discharge records at the Gulliemard Bridge gauge station between the years 1960 and 2016, and delivered by the DID (Drainage and Irrigation Department) Malaysia. The characterization of flood peak flow values is based on maximum streamflow discharge values at an annual scale,

which means for each year there will be only one flood episode at the targeted site (i.e. [5, 10]), using the Equation 1. Similarly, corresponding to each flood peak value of the annual basis flood events, flood volume and duration series are derived from the streamflow hydrograph using the algorithm reported by studies such as Yue et al. [10], Zhang and Singh [8], and Reddy and Ganguli [3, 15], using Equations 2 and 3 as referring to Figure 3 (graphical illustration).

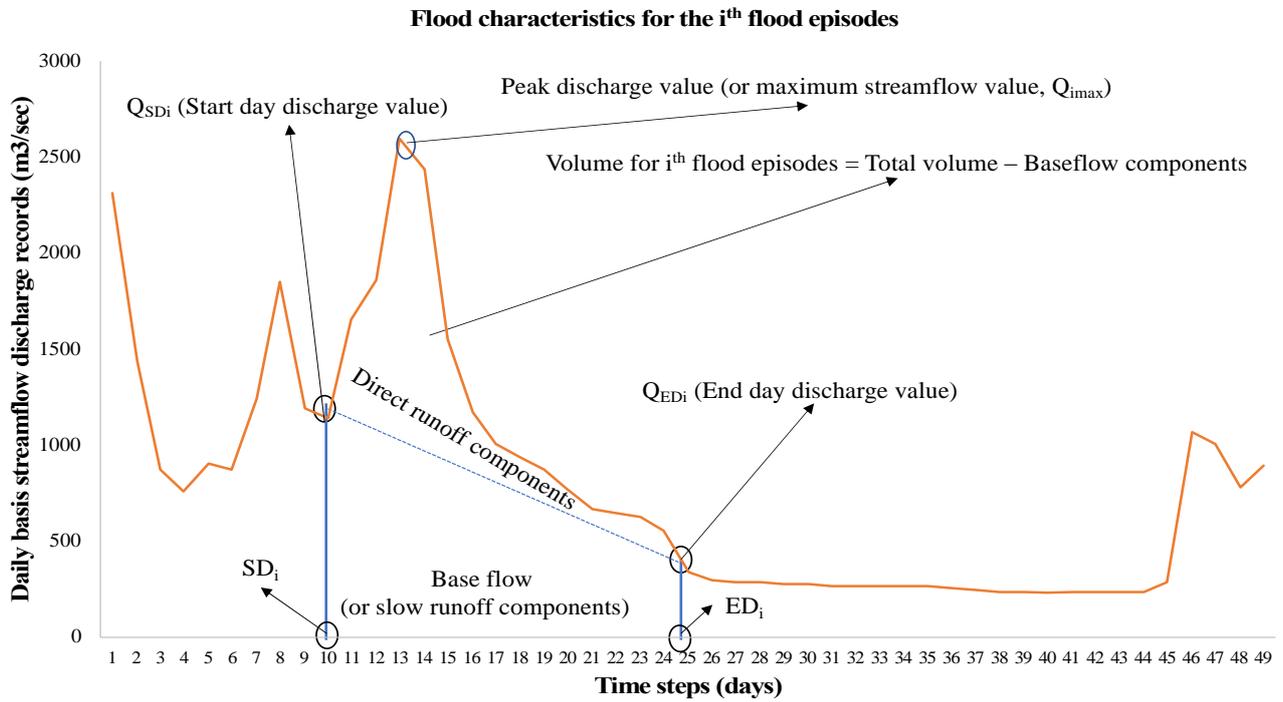


Figure 3. A graphical illustration of typical hydrograph of flood characteristic

$$\text{Annual flood peak series} = P_i = \max\{Q_{ij}, j = SD_i + SD_i + 1, \dots, ED_i\} \tag{1}$$

$$\text{Volume series} = V_i = V_i^{\text{total}} - V_i^{\text{Baseflow}} = \sum_{j=SD_i}^{ED_i} Q_{ij} - \frac{(1+D_i)(Q_{SD_i}+Q_{ED_i})}{2} \tag{2}$$

$$\text{Duration series} = D_i = ED_i - SD_i \tag{3}$$

Where,  $Q_{ij} = j^{\text{th}}$  days streamflow discharge value for the  $i^{\text{th}}$  year; and  $Q_{SD_i}$  &  $Q_{ED_i}$  = streamflow discharge value for the start date ‘ $SD_i$ ’ and end date ‘ $ED_i$ ’ of the flood runoff. Furthermore, according to Sraj et al. [9], flood peak discharge flow often attains its maximum value but this is not mandatory for the volume and duration series. The descriptive statistics of derived annual flood characteristics are shown in Table 1.

Table 1. Descriptive statistics of annual flood characteristics

| Descriptive measure      | P(m <sup>3</sup> /sec) | V(m <sup>3</sup> ) | D(days) |
|--------------------------|------------------------|--------------------|---------|
| Sample Size              | 50                     | 50                 | 50      |
| Range                    | 19670                  | 71558              | 57      |
| Min                      | 916.3                  | 3182.3             | 7       |
| Max                      | 20586                  | 74740              | 64      |
| Mean                     | 6078                   | 19122              | 19.04   |
| Variance                 | 2.15E+07               | 2.14E+08           | 117.75  |
| Std. Deviation           | 4639                   | 14623              | 10.851  |
| Skewness (Fisher)        | 1.5532                 | 1.6392             | 2.2793  |
| Skewness (Pearson)       | 1.506                  | 1.590              | 2.210   |
| Kurtosis (Pearson)       | 1.883                  | 2.864              | 6.252   |
| Excess Kurtosis (Fisher) | 2.2158                 | 3.3029             | 7.0557  |

### 3. Methods

#### 3.1. Bivariate Joint Distribution via Copula Function

The idea of the copula function was developed by Saklar [16]. It models the multivariate distributions to their univariate marginal distribution functions, which are not necessary from the same family of probability functions, capture a broader extent of dependencies (i.e., both linear and nonlinear), and preserve their joint dependence structure [6, 18, 19]. Mathematically, if  $(M, N)$  are the bivariate random vectors, such that  $u = F_M(m) = P(M \leq m)$  and  $v = F_N(n) = P(N \leq n)$  is the univariate marginal distribution of individual characteristics, then it can be demonstrated through a function called copula or  $C$ , which describes the dependence structure in between them and can be defined on the unit square. It can be expressed as:

$$H_{M,N}(m, n) = C[F_M(m), F_N(n)] = C(u, v) \quad (4)$$

Where  $C$  = any type of bivariate copulas under consideration;  $F_M(m) = F_N(n)$  = cumulative distribution functions of univariate vector 'M' and 'N';  $H_{M,N}(m, n)$  = bivariate joint distribution functions, which can be expressed in terms of their univariate marginal distributions and the associated dependence function  $C$ , as revealed from the above Equation 4. Similarly, if  $f_M(m)$  and  $f_N(n)$  are the probability density function or PDF of variables  $M$  and  $N$ , then the joint probability density of the two random variables can be expressed as:

$$f_{M,N}(m, n) = c(F_M(m), F_N(n)) f_M(m) f_N(n) \quad (5)$$

Where  $c$  is the density function of bivariate copula  $C$ . It can be defined as:

$$c(u, v) = \frac{\partial^2 c(u, v)}{\partial u \partial v} \quad (6)$$

Where  $u = F_M(m)$  and  $v = F_N(n)$

For the extended mathematical details and theorems about copula functions, readers are advised to refer to Saklar [16], Nelsen [18] and Genest and Favre [44], as well as the 'International Association of Hydrological Sciences (or IAHS)' for extended details and lists of their applicability in the field of hydro-climatological observations.

In this demonstration, the two-dimensional (2-D) copulas, such as the Archimedean family-based Clayton, Gumbel and Frank copula and also one Elliptical family called the Gaussian copula, are introduced and tested as candidate functions, where the best-fitted copulas are employed to model the bivariate joint dependency and estimation of joint return periods of the flood pairs (P-V) and (P-D). In actuality, different copula functions have different extents of dependency capturing capabilities. For example, both the Gumbel and Clayton copulas are only capable of modelling positively correlated random variables (i.e.,  $0 \leq \theta < \infty$  & Kendall's tau  $\tau_\theta \geq 0$ ), while the Frank family copula accommodates the entire range of dependencies among hydrologic characteristics (i.e.,  $\tau_\theta \in [1, -1]$ ) (i.e. [18, 45]). Mathematically, we can approximate the bivariate Archimedean class copula based on the two-dimensional copula framework (i.e.  $[C: [0,1]^2 \rightarrow [0,1]]$ ) as given below;

$$C(u, v) = \phi^{-1}(\phi(u) + \phi(v)) \text{ for } u, v \in [0,1] \quad (7)$$

Where  $\phi(\cdot)$  and  $\phi^{-1}$  represent the generator function of the specified Archimedean copulas and their inverse, such that the generator ( $\phi: I \rightarrow R^+$ ) signifies the positive, convex and decreasing function and can be approximated for ( $\phi(1) = 0$  &  $\phi(1) = \infty$ ) [18]. The mathematical expression for bivariate Archimedean copulas and their associated properties (i.e., parameter range ( $\theta$ ), generating function (or generator)  $\phi(t)$ , and relationship of Kendall's  $\tau$  and  $\theta$  ( $\tau_\theta$ )) are listed in Table 2. On other side, the Gaussian copula is an implicit copula that can easily capture both the positive as well as the negative dependency, and has almost no dependence on the tails of distribution, which are mostly distributed around the centre and can be mathematically expressed as [46, 47].

$$C_\theta(u, v) = \int_{-\infty}^{-\phi^{-1}(u)} \int_{-\infty}^{-\phi^{-1}(v)} \frac{1}{2\pi(1-\theta^2)^{1/2}} \exp\left[-\frac{s^2 - 2\theta st + t^2}{2(1-\theta^2)}\right] ds dt \quad (8)$$

**Table 2. Mathematical expressions for bivariate Archimedean copulas and their associated properties**

| Copula family   | Bivariate copula $C_\theta(u, v)$                                                                           | Parameter range ( $\theta$ ) | Generating function (or generator) $\phi(t)$                    | Relation of Kendall's $\tau$ and $\theta$ ( $\tau_\theta$ )                                                                                                                                                                              |
|-----------------|-------------------------------------------------------------------------------------------------------------|------------------------------|-----------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Clayton         | $[\max\{u^{-\theta} + v^{-\theta} - 1; 0\}]^{-1/\theta}$                                                    | $0 \leq \theta < \infty$     | $\frac{1}{\theta}(t^{-\theta} - 1)$                             | $\frac{\theta}{\theta + 2}$                                                                                                                                                                                                              |
| Frank           | $\frac{-1}{\theta} \ln \left( 1 + \frac{(e^{-\theta u} - 1)(e^{-\theta v} - 1)}{(e^{-\theta} - 1)} \right)$ | $-\infty < \theta < \infty$  | $-\ln \left( \frac{e^{-\theta t} - 1}{e^{-\theta} - 1} \right)$ | $1 + 4 \left( \frac{D_1(-\ln \theta) - 1}{\ln \theta} \right)$<br>where $D_k(x)$ is the Debye function, for any positive integer k,<br>$D_k(x) = \frac{k}{x^k} \int_0^x t^k / (e^t - 1) dt$ (Zhang and Singh 2006 and Wang et al., 2009) |
| Gumbel-Hougaard | $\exp \left\{ -[(-\ln(u))^\theta + (-\ln(v))^\theta]^{1/\theta} \right\}$                                   | $1 \leq \theta < \infty$     | $(-\ln t)^\theta$                                               | $\frac{\theta - 1}{\theta}$                                                                                                                                                                                                              |

**3.1.1. Estimating Copula Dependence Parameters and their Fitness Investigation**

Different approaches are often motivated in estimating the copula dependence parameters such as the method of moment (MOM) [48], exact maximum likelihood or EML [46], inference functions for marginal (IFM) [45], canonical maximum likelihood (CML) [49], Maximum pseudo likelihood estimations (MPL) [3, 49]. In this demonstration, the copula dependence parameters are derived using the method-of-moments (or MOM) estimators based on the inversion of Kendall's tau ( $\tau$ ) and are highly flexible and very popular, especially among the Archimedean copula families [12, 49]. The MOM estimation approach is solely based on establishing an interlink between the unknown dependence parameter ' $\theta$ ' and the sample ranked correlation coefficient, i.e., Kendall's tau ( $\tau$ ), and if a correlation is exhibited between them, then the copula parameter can be estimated as [3, 12, 18];

$$\hat{\theta} = f(\hat{\tau}) \tag{9}$$

$$\tau = 4 \int_{[0,1]^2} C(u, v) dC(u, v) - 1 = 1 + 4 \int_0^1 \frac{\phi(t)}{\phi'(t)} dt \tag{10}$$

Where  $\phi(\cdot)$  and  $\phi'(t)$  indicate the generator function and their first derivatives for the Archimedean copula family function. In addition, Kendall's tau ( $\tau$ ) can be statistically defined as [50];

$$\left(\frac{n}{2}\right)^{-1} \sum_{i < j} \text{sign}[(x_i - x_j)(y_i - y_j)] \text{ where; sign} = 1, \text{ if } [(x_i - x_j)(y_i - y_j) > 0] \text{ and sign} = -1, \text{ if } [(x_i - x_j)(y_i - y_j) < 0] \text{ for } i = j = 1, 2, \dots, n \tag{11}$$

The mathematical range of Kendall's tau ( $\tau$ ) is  $[-1,1]$ , such that  $\tau = 1$  represents concordance,  $\tau = -1$  represents discordance, and  $\tau = 0$  represents no concordance exhibited between random pairs. The adequacy of hypothesized copulas fitted to bivariate flood characteristics are identified using the Cramer-von Mises statistics, which are based on the parametric bootstrapping procedure that makes use of the Cramer-von Mises (or CvM) statistic,  $S_n$  (i.e. [3, 51, 52]). The simulation of copula function (i.e. estimating copula dependence parameters and their fitness test) was carried out using the R statistics platform [53]. The CvM test statistics can be mathematically expressed as;

$$S_n = n \int_{[0,1]^2} \{C_n(u, v) - C_\theta(u, v)\}^2 dC_n(u, v) = \sum_{i=1}^n \{C_n(U_{i,n}, V_{i,n}) - C_\theta(U_{i,n}, V_{i,n})\}^2 \tag{12}$$

Where  $C_n$  and  $C_\theta$  are the empirical copula (estimated using n observational flood characteristics) and parametric copula (which is derived under the null hypothesis). Under this test, the p-value is estimated using the parametric bootstrapping procedure (i.e. [51]). Overall, a minimum value of the  $S_n$  test statistic must indicate a minimum gap and deviation between an empirical and derived parametric copula.

**3.2. Bivariate Risk Evaluation using the Notation of Failure Probability (FP)**

In the multivariate risk framework, the return periods can be derived from the exceedance probabilities of the flood attributes pair [3, 54]. In actuality, hydrology and hydraulic applications are mostly interested in the evaluation of the mean inter-arrival period between two design events, which are usually defined in a year called the return period [14, 55]. The OR-joint return period of the annual flood episodes, where either of the flood vectors in a bivariate combination exceeds a specific threshold value, can be estimated [7, 14] as:

$$T_{P,V}^{OR} = \frac{1}{P(P \geq p \text{ OR } V \geq v)} = \frac{1}{(1-H(p,v))} = \frac{1}{(1-C(F(p),F(v)))}, \text{ for flood pair (P - V)} \tag{13}$$

$$T_{P,D}^{OR} = \frac{1}{P(P \geq p \text{ OR } D \geq d)} = \frac{1}{(1-H(p,d))} = \frac{1}{(1-C(F(p),F(d)))}, \text{ For flood pair (P - D)} \tag{14}$$

Where  $H(p, v)$  and  $H(p, d)$  are the joint cumulative densities estimated using the best-fitted bivariate copula  $C(F(p), F(v))$  and  $C(F(p), F(d))$  with marginal distribution  $F(p)$ ,  $F(v)$ , and  $F(d)$  series. On the other side, the joint return period, where both of the flood characteristics simultaneously exceed a certain threshold value, also called primary ‘AND’ joint return periods, can be expressed as:

$$T_{P,V}^{AND} = \frac{1}{P(P \geq p \text{ AND } V \geq v)} = \frac{1}{(1-F(p)-F(v)+C(F(p),F(v)))}, \text{ for flood pair (P - V)} \tag{15}$$

$$T_{P,D}^{AND} = \frac{1}{P(P \geq p \text{ AND } D \geq d)} = \frac{1}{(1-F(p)-F(d)+C(F(p),F(d)))}, \text{ for flood pair (P - D)} \tag{16}$$

The notation of failure probability or FP can be practical in assessing the variation of bivariate flood risk during the entire project lifetime. The FP statistics usually define the chance of a potential flood event occurring at least once in a given project design lifetime (i.e. [26-27, 30]). Mathematically, the failure probability can be expressed as:

$$p_T = 1 - P\{(M_T, N_T) \notin S_T\} \tag{17}$$

Where  $S_T$  is the series of bivariate flood hazard scenario. In addition, the failure probability for the independent and identical distribution series can be estimated using Equation 11 [27];

$$p_T = 1 - (1 - p)^T \tag{18}$$

However, the failure probability for both the primary ‘OR’ and ‘AND’ joint return periods can be estimated as;

$$p_T^{OR} = 1 - (1 - P(M \geq m \text{ OR } N \geq n))^T, \text{ for OR - joint case} \tag{19}$$

$$p_T^{AND} = 1 - (1 - P(M \geq m \text{ AND } N \geq n))^T, \text{ for AND - joint case} \tag{20}$$

## 4. Results and Discussion

### 4.1. Bivariate Dependency Modelling using Two-dimensional Copula

Approximating the most parsimonious probability density functions or PDFs for defining the univariate marginal distribution of the individual flood characteristics is often a mandatory pre-requisite before introducing the random observation into the multivariate or copula distribution framework. In this demonstration, few commonly used one-dimensional parametric family functions, such as Gamma-3P, Generalized Extreme Value (GEV), Generalized Gamma-3P, Inverse Gaussian-2P, Johnson SB-4P, Lognormal-2P and Weibull-2P, are selected and tested as candidate models, and all the distribution fitting procedures are carried out using the Easyfit software (MathWave Technologies 2004–2017). In reality, no universally accepted distributions are assigned or in favour of any probability distribution functions from any literatures to model any extreme hydrologic observations (i.e. [10, 56, 57]). The Gringorten-based position-plotting formula is used in estimating the empirical probabilities [58]. Different goodness-of-fit test statistics, such as the Kolmogorov–Smirnov (or K–S) test and the Anderson–Darling (or A–D) test [59], based on information criteria statistics such as Akaike information criteria (or AIC) [60], Schwartz’s Bayesian information criteria (or BIC) [61] and Hannan–Quinn Information criteria (HQC) [62] are used in visualizing the best-fitted marginal distribution to each individual flood characteristic. Therefore, the minimum test statistics value must indicate the best-fitted model. Table 3 illustrates the fitness test statistics of the candidate function and it is shown that 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.

**Table 3. Goodness-of-fit statistics of univariate distribution fitted with annual flood characteristics (a) based on Kolmogorov–Smirnov and Anderson–Darling test (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)<br>Functions   | Peak     |          |         | Volume   |          |         | Duration |          |          |
|--------------------|----------|----------|---------|----------|----------|---------|----------|----------|----------|
|                    | AIC      | BIC      | HQIC    | AIC      | BIC      | HQIC    | 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; HQIC 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 series, thus could be further indicated for the better performance

The strength of dependency among the triplet flood characteristics were investigated analytically using the Pearson’s linear correlation ( $r$ ), Kendall’s tau ( $t$ ) and Spearman’s rho ( $\rho$ ), and their estimated statistics are shown in Table 4. Based on the estimated statistics, it is clearly inferred that there is a strong positive correlation between flood pair (P-V), but a weak and negative correlation between flood pair (P-D). Three Archimedean copula families, the Clayton, Gumbel and Frank copula, and one Elliptical copula, the Gaussian, were selected. The copulas dependence parameters were derived using the method-of-moment (MOM) estimators, which are based on the inversion of Kendall’s tau ( $\tau$ ) using Equation 9, 10 and 11, and their estimated statistics are listed in Table 5. The fitness level and adequacy of the candidate copulas for both flood pair (P-V) and (P-D) were investigated using the CvM distance statistics with the parametric bootstrap procedure, where test statistic ‘ $S_n$ ’ and its associated p-values were estimated from 1000 and 500 simulated random samples by the mean of the parametric bootstrap procedure (Table 5). The analytical fitness investigation reveals that the Gaussian copula exhibited the minimum ‘ $S_n$ ’ test statistics value and also the highest p-value for both the (P-V) and (P-D) pairs, and thus was identified as the most justifiable copula for describing their joint dependence structure. A graphical visual inspection was also carried out using the scatter plot (refer to Figure 4) of the 1000 flood pairs, for both (P-V) and (P-D) observations, which were simulated from the joint distribution using the Gaussian copula function with best-fitted marginal distributions. From the comparison scatter plots of the simulated data (blue colour) with overlapped observed flood characteristics (red colour), it can be revealed that the Gaussian copula performed satisfactorily for capturing observed dependence for both flood pair (P-V) and (P-D), which means the simulated data adequately overlapped with the natural dependence of observed flood characteristics. Figure 5 and 6 illustrate the joint probability density and cumulative distribution plot of the bivariate Gaussian copula fitted to flood pair (P-V) and (P-D).

**Table 4. The correlation coefficient statistics for flood characteristics**

| Dependence measure                | Peak-Volume(P-V) | Peak-Duration(P-D) |
|-----------------------------------|------------------|--------------------|
| Pearson’s correlation ( $r$ )     | 0.7387784        | -0.0061526         |
| Kendall’s correlation( $\tau$ )   | 0.60759499       | -0.0741828         |
| Spearman’s correlation ( $\rho$ ) | 0.79425677       | -0.094851          |

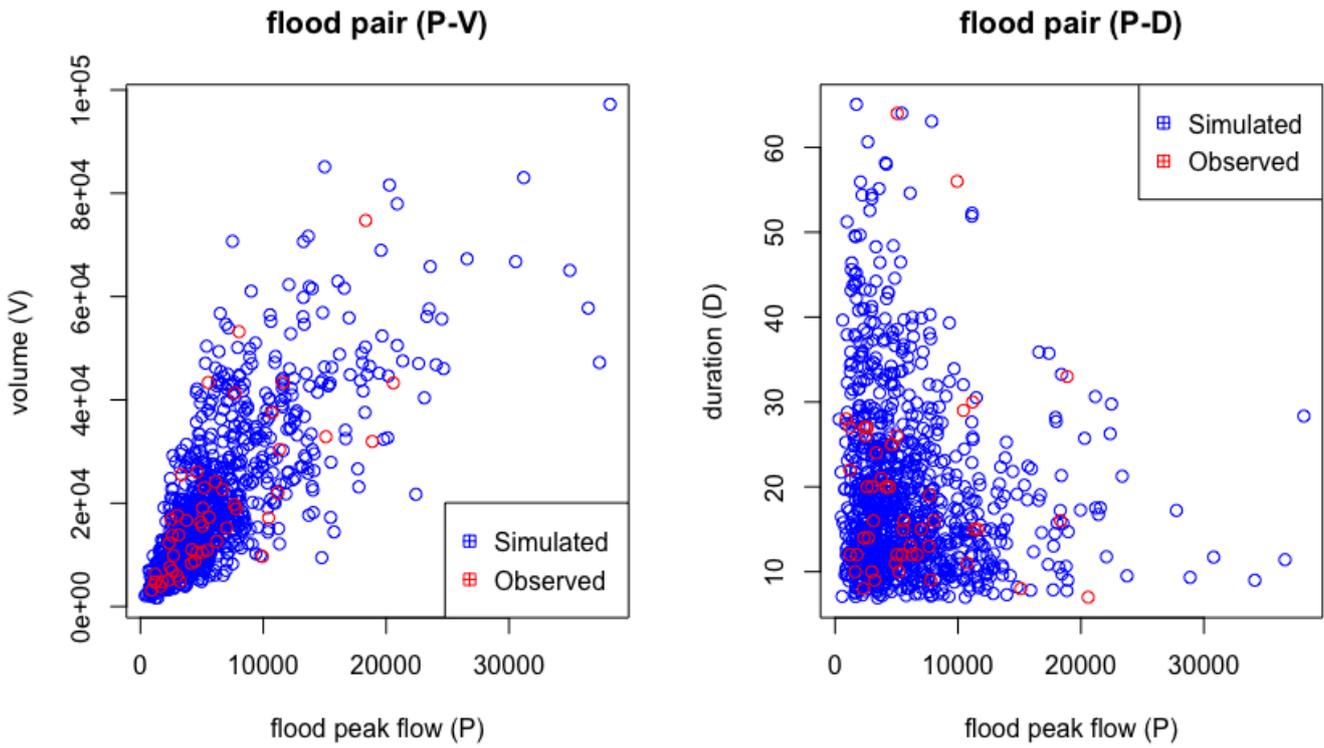


Figure 4. Scatter plot for comparing the observed (red colour) flood peak-volume and peak-duration series with the set of 1000 simulated samples (blue colour) from best-fitted Gaussian copula functions via the MOM estimation procedure

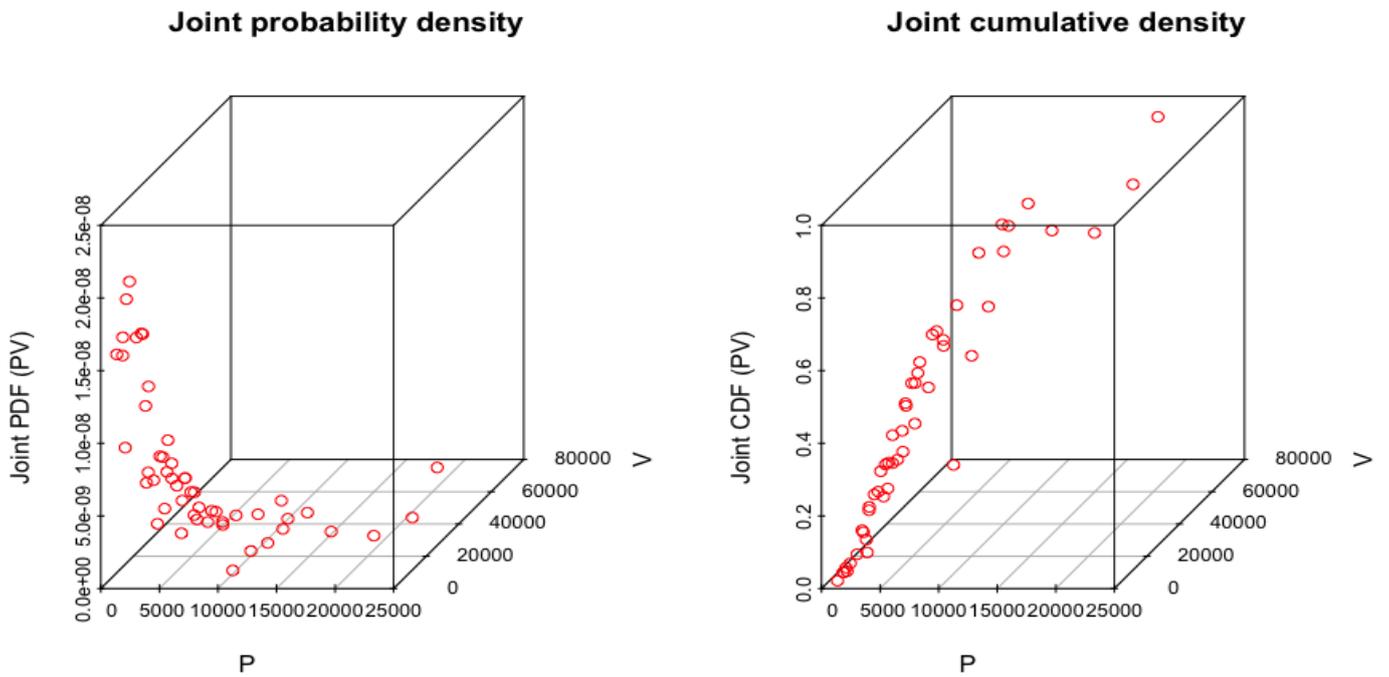


Figure 5. Joint probability density and joint cumulative distribution plot derived from the Gaussian copula fitted to flood peak flow-volume pair

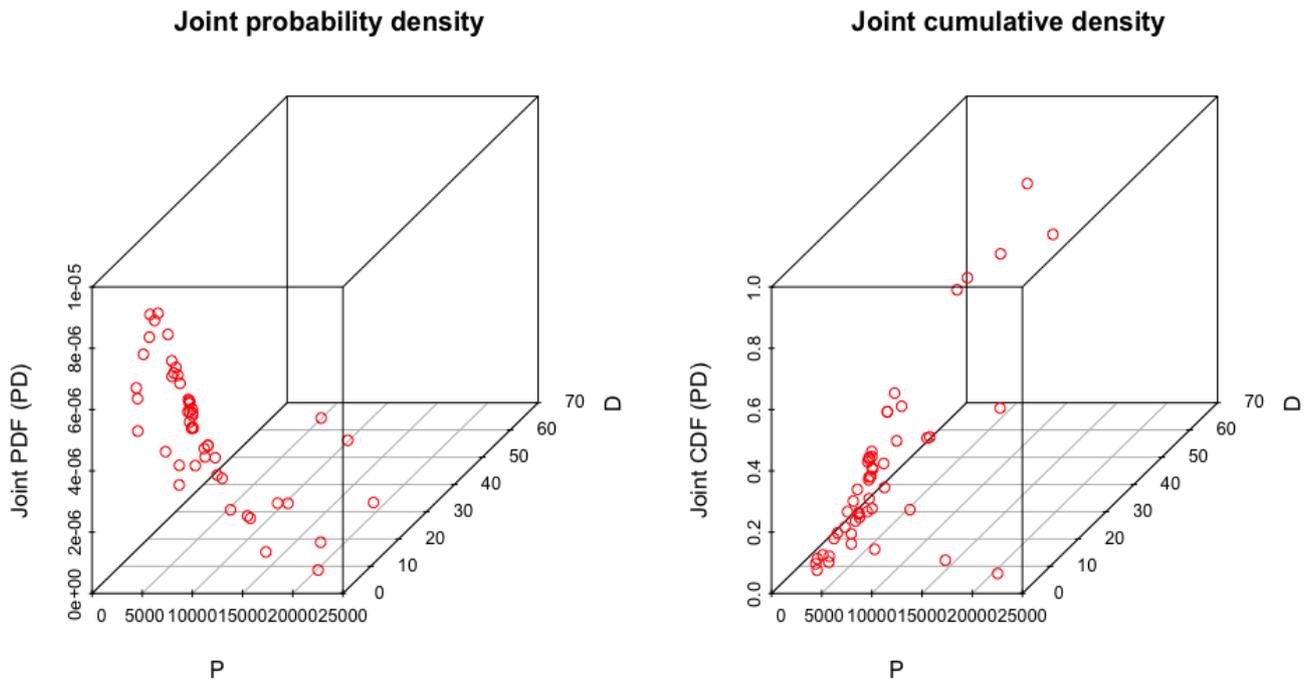


Figure 6. Joint probability density and joint cumulative distribution plot derived from the Gaussian copula fitted to flood peak flow-duration pairs

Table 5. Estimated copula dependence parameters and goodness of fit of the bivariate joint distribution for (a) flood peak flow and volume (b) for flood peak flow and duration pairs

| Copula family        | Parameter Estimates $\theta^*$ | Standard Error SE | N=1000<br>(No. of bootstrap sampling) |                    | N=500<br>(No. of bootstrap sampling) |                    |
|----------------------|--------------------------------|-------------------|---------------------------------------|--------------------|--------------------------------------|--------------------|
|                      |                                |                   | $S_n$                                 | (p-value) $_{S_n}$ | $S_n$                                | (p-value) $_{S_n}$ |
| (a) For (P-V) pairs  |                                |                   |                                       |                    |                                      |                    |
| Gaussian*            | 0.8159716                      | 0.054             | 0.01579                               | 0.9156             | 0.015797                             | 0.9012             |
| Clayton              | 3.097                          | 0.774             | 0.027508                              | 0.2073             | 0.027508                             | 0.1727             |
| Gumbel-Hougaard (GH) | 2.548387                       | 0.387             | 0.019973                              | 0.6199             | 0.019973                             | 0.5818             |
| Frank                | 8.135964                       | 1.009             | 0.022772                              | 0.4461             | 0.022772                             | 0.508              |
| (b) For (P-D) pairs  |                                |                   |                                       |                    |                                      |                    |
| Gaussian*            | -0.1163                        | 0.178             | 0.030559                              | 0.4101             | 0.030559                             | 0.3982             |
| Frank                | -0.6706                        | 0.003             | 0.030644                              | 0.3881             | 0.030644                             | 0.4401             |
| Clayton              | NA                             | NA                | NA                                    | NA                 | NA                                   | NA                 |
| Gumbel-Hougaard (GH) | NA                             | NA                | NA                                    | NA                 | NA                                   | NA                 |

Note: \*, represents that Gaussian copula poses the minimum value of  $S_n$  statistics in compare with the other bivariate copula functions for both flood pair (P-V) and (P-D),

NA denotes that the Clayton and Gumbel-Hougaard cannot be used for negative dependence modelling between the random variables, i.e., for flood pair (P-D), which is only applicable for positively correlated flood characteristics, i.e., for flood pair (P-V).

### 4.2. Bivariate Hydrologic Risk Estimation

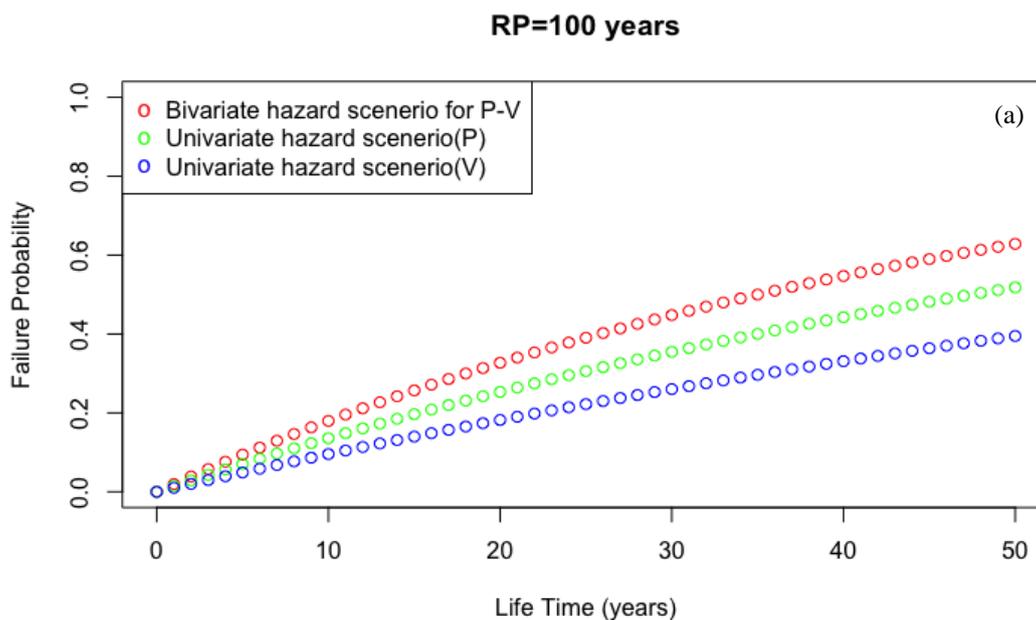
The selection of an appropriate or justifiable return period solely depends upon the importance of undertaken structure as well as its consequences of failure [20]. The best-fitted Gaussian copulas were employed in deriving primary ‘OR’ and ‘AND’ joint return periods or RPs for flood pairs (P-V) and (P-D) using Equations 13 to 16 for the given flood characteristics (see Table 6). It can be noticed that for the same values of the peak flow and volume pair (P-V) or the flood peak flow and duration pair (P-D), the joint return period in the ‘AND’ case is much greater than ‘OR’ joint cases, i.e.,  $T_{OR}$  (OR-joint case) <  $T_{AND}$  (AND-joint case. For example, the flood episodes with peak flow,  $P = 8545.246 \text{ m}^3\text{s}^{-1}$ , volume,  $V = 28,604.77 \text{ m}^3$  and duration,  $D = 25.78303$  (days), the joint return period between flood

pair P-V for ‘OR’ and ‘AND’ cases were  $T_{PV}^{AND} = 7.42953642$  years and  $T_{PV}^{OR} = 3.67338333$  years and between P-D pairs were  $T_{PD}^{AND} = 30.8281362$  years and  $T_{PD}^{OR} = 2.6710215$  years. Similarly, for  $P = 14,725.45 \text{ m}^3\text{s}^{-1}$ ,  $V = 49,059.65 \text{ m}^3$  and  $D = 12.2961453$  (days),  $T_{PV}^{AND} = 35.1571701$  years and  $T_{PV}^{OR} = 12.2961453$  years for P-V pairs, and  $T_{PD}^{AND} = 563.126478$  years and  $T_{PD}^{OR} = 9.25889919$  years for P-D pairs. In addition, from the same results it was also concluded that the univariate RP is greater than the OR-joint case but smaller than the AND-joint case. We can further conclude that the univariate approach in the hydrologic risk would be inappropriate and ineffective.

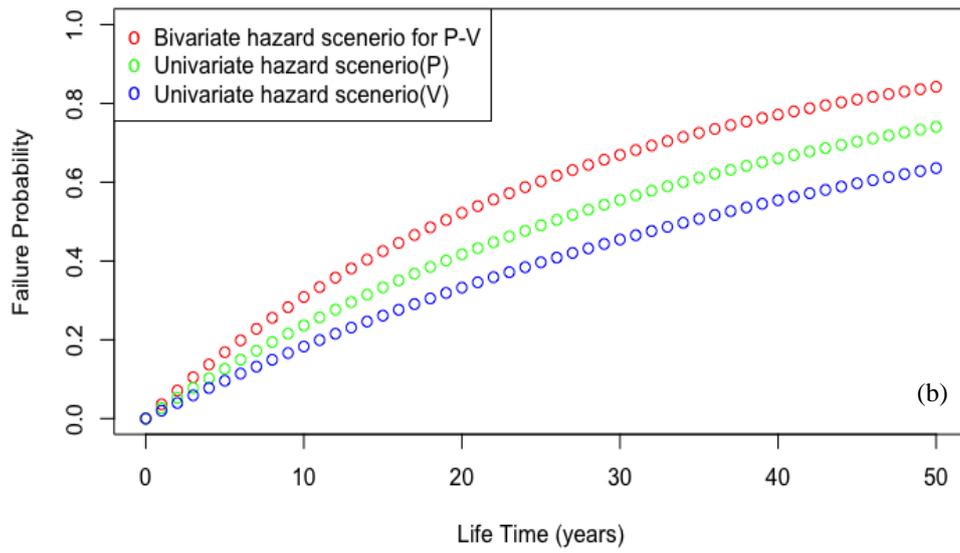
**Table 6. The univariate return period and the joint return period of flood characteristics (i.e., P-V and P-D)**

| Return period (year) | Flood peak (P) | Volume (V) | Duration (D) | OR RP (year) between P-V | AND RP (year) between P-V | OR RP (year) between P-D | AND RP (year) between P-D |
|----------------------|----------------|------------|--------------|--------------------------|---------------------------|--------------------------|---------------------------|
| 5                    | 8545.246       | 28604.77   | 25.78303     | 3.67338333               | 7.42953642                | 2.6710215                | 30.8281362                |
| 10                   | 11512.52       | 38973.7    | 32.43877     | 6.69856523               | 16.2590258                | 4.92082636               | 132.06376                 |
| 20                   | 14725.45       | 49059.65   | 38.91135     | 12.2961453               | 35.1571701                | 9.25889919               | 563.126478                |
| 50                   | 19425.96       | 61537.15   | 47.29428     | 27.6002926               | 96.2602878                | 21.5702727               | 3847.63371                |
| 100                  | 23366.43       | 70112.74   | 53.5464      | 50.9990718               | 40.9401496                | 204.42373                | 13495.2767                |

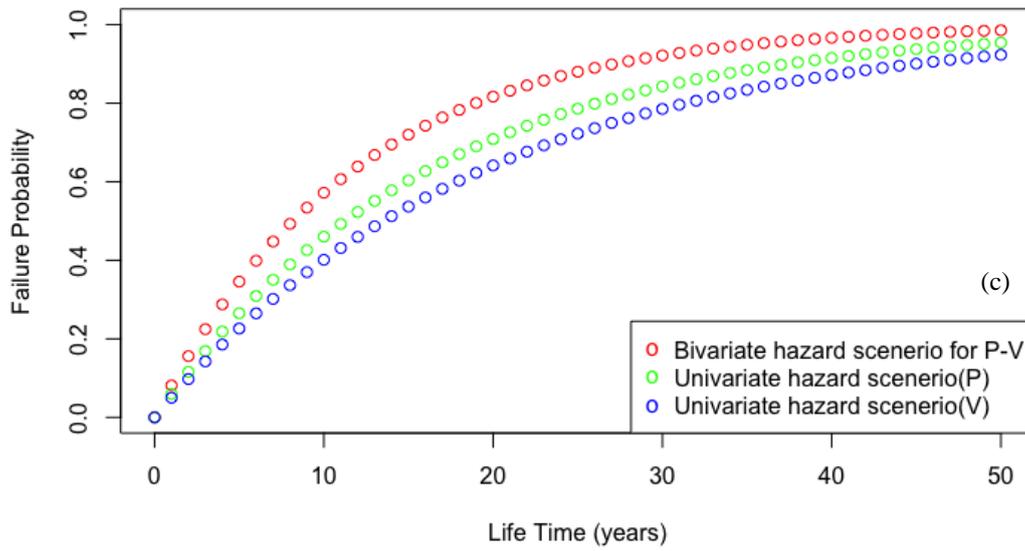
The failure probability statistics (see section 3.2) were estimated to demonstrate the bivariate hydrologic risk of flood characteristics for flood pair (P-V) and (P-D), which usually measure the chance of flood episodes potentially occurring once in a given project design lifetime. Figures 7 and 8 illustrate the risk statistics of the flood episodes for different return periods (i.e., 100, 50, 20, 10, 5 years) where the failure probabilities are estimated based on the univariate approach (indicated by blue and green colour), and bivariate events based on the OR-joint scenario for the flood pair (P-V) (Figure 7 a, b, c, d, and e) and flood pair (P-D) (Figure 8 a, b, c, d, and e). The results revealed that the risk statistics value would increase by an increment in the service lifetime. For example, for the 100-year univariate return period events, the failure probability attained the value,  $FP = 0.179654$  when the service lifetime is 10 years, which is lower in comparison with 50-year RP events, i.e.,  $FP = 0.3086$ , at the same service lifetime. Similarly, for the 5-year RP events, the FP is 0.95831, which is higher than the FP statistics estimated for 10- and 20-year RP events at the same service lifetime. On the other side, the univariate events produce lower FP values than the bivariate events (see both Figure 7 and 8). Thus, it could be revealed that neglecting the joint impact or mutual dependencies between multiple flood characteristics would be problematic and underestimate the FPs in the assessments of flood risk. From the same figure, it is also inferred that with the increase in the value of return period statistics, both the bivariate and univariate hydrologic risk would be decreased.



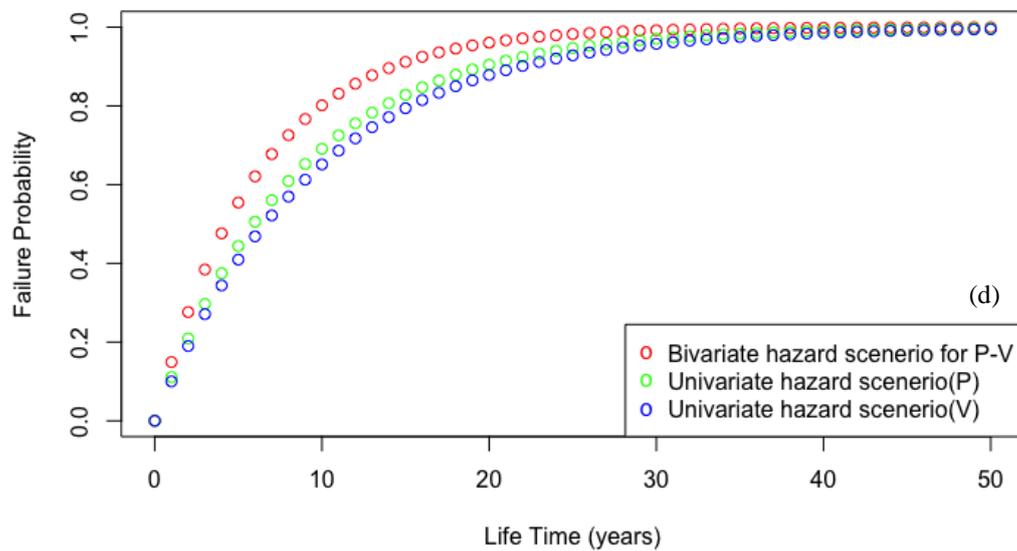
**RP=50 years**



**RP=20 years**



**RP=10 years**



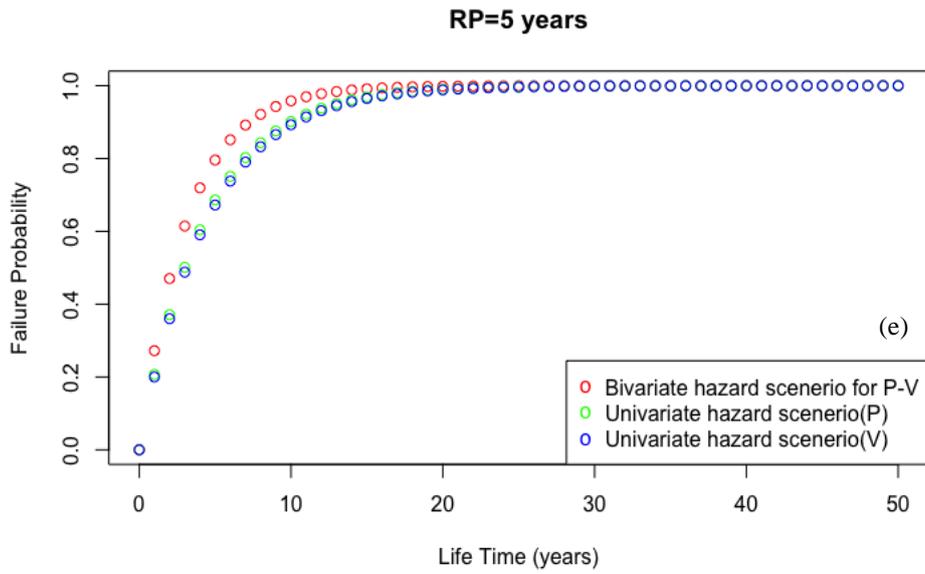
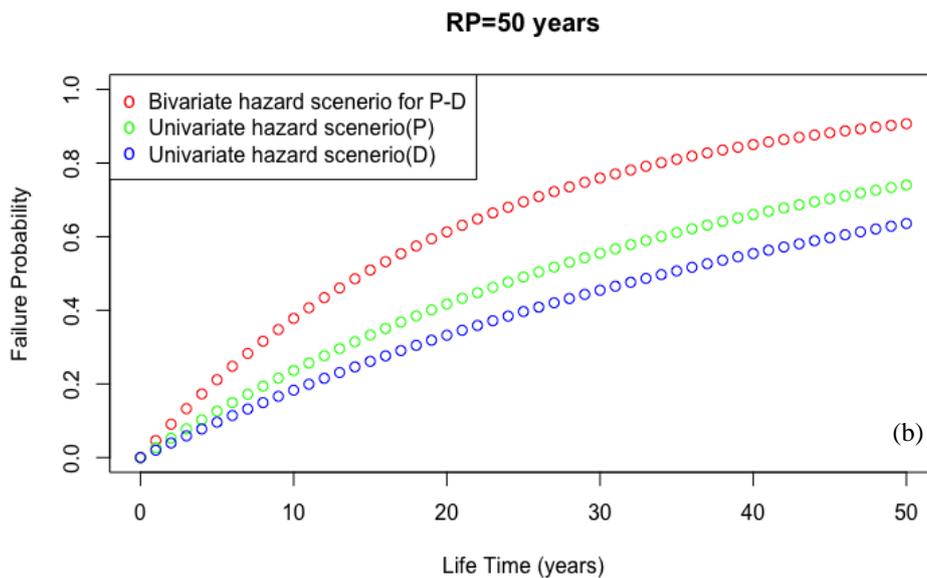
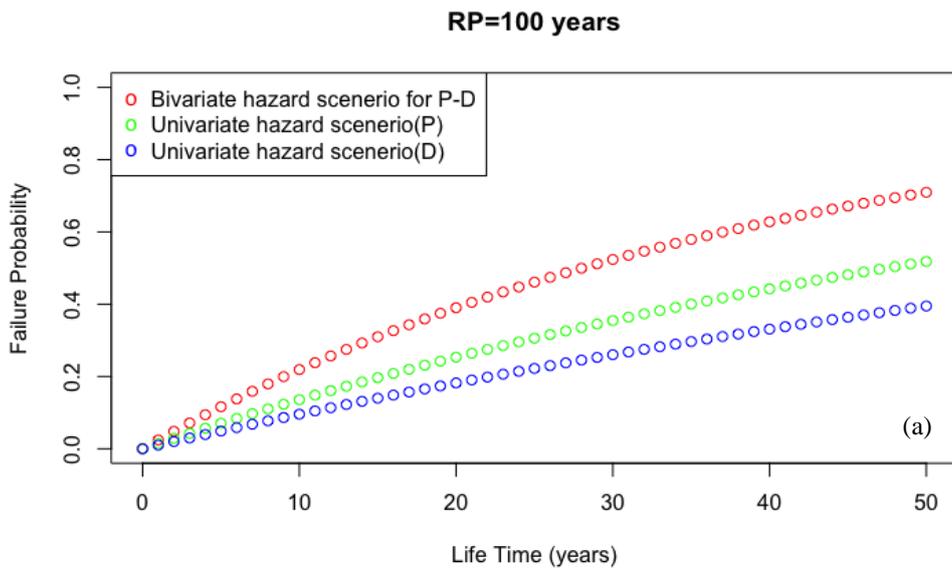


Figure 7. Bivariate hydrologic risk of flood peak flow-volume series in different return periods a) 100-year, b) 50-year, c) 20-year, d) 10-year, e) 5-year. The red colour plot indicates the estimated failure probability for bivariate events based on the OR-joint probability scenario, the blue and black colour plot indicates the estimated failure probability for the univariate events



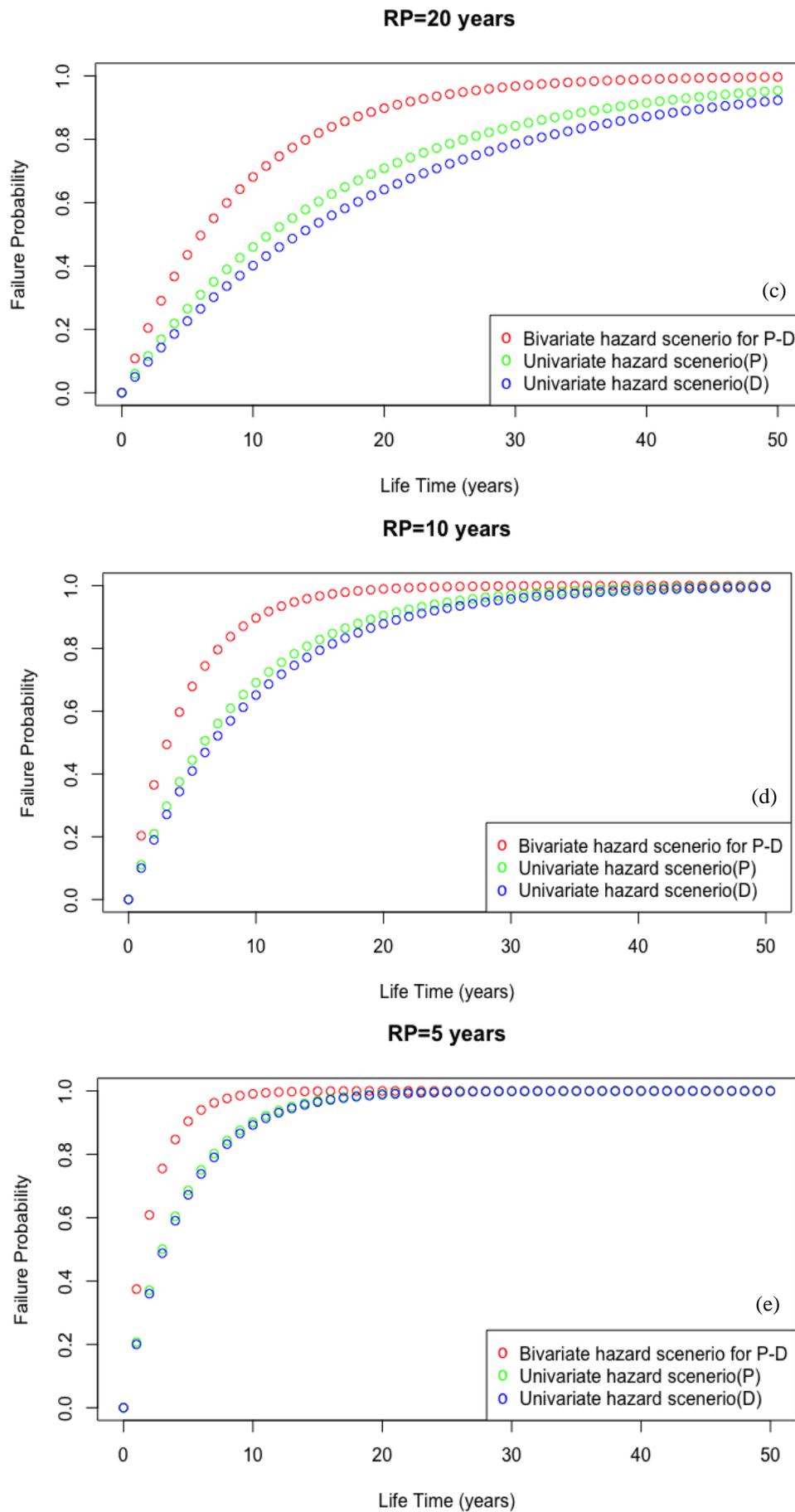
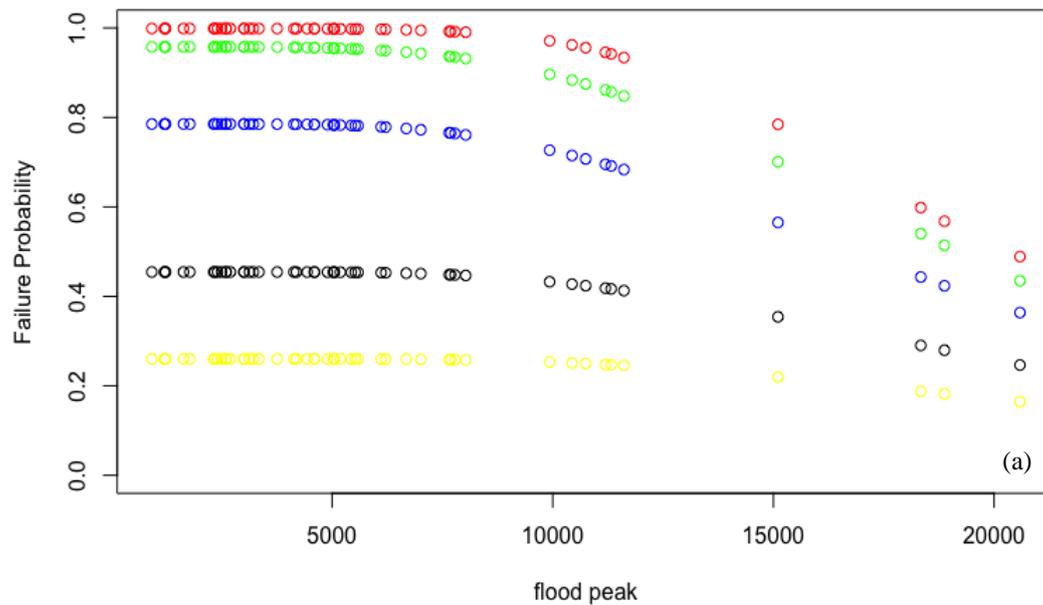


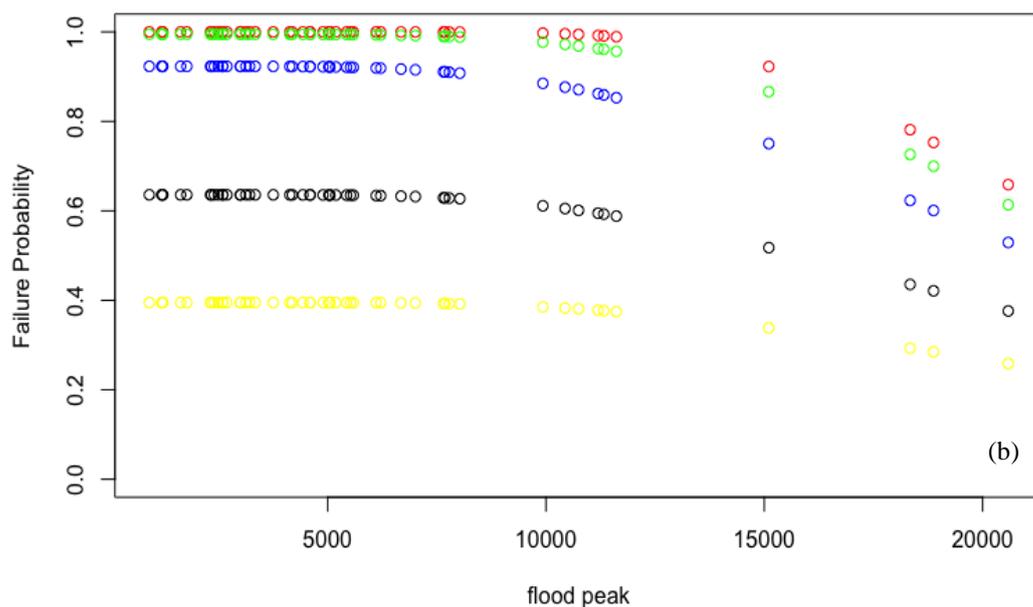
Figure 8. Bivariate hydrologic risk of flood peak flow and duration series in different RPs a) 100-year, b) 50-year, c) 20-year, d) 10-year, e) 5-year

Similarly, Figures 9 and 10 illustrates bivariate hydrologic risk statistics for the flood peak flow and volume pair (P-V) and the flood peak flow and duration pair (P-D) under the different scenarios where we considered the 5-year, 10-year, 20-year, 50-year, and 100-year RP of the flood volume and duration series. In this analysis, variation of the bivariate hydrologic risk with the change of flood peak flow in different synthetic or designed flood durations was studied for different project lifetimes (i.e. 30, 50, and 100 years) service time. The investigation revealed that the value of the bivariate hydrologic risk statistics is increasing with the increment in the project lifetime and at the same time, it is decreasing with the increment in the RP of flood volume and duration.

Service time-30 years



Service time-50 years



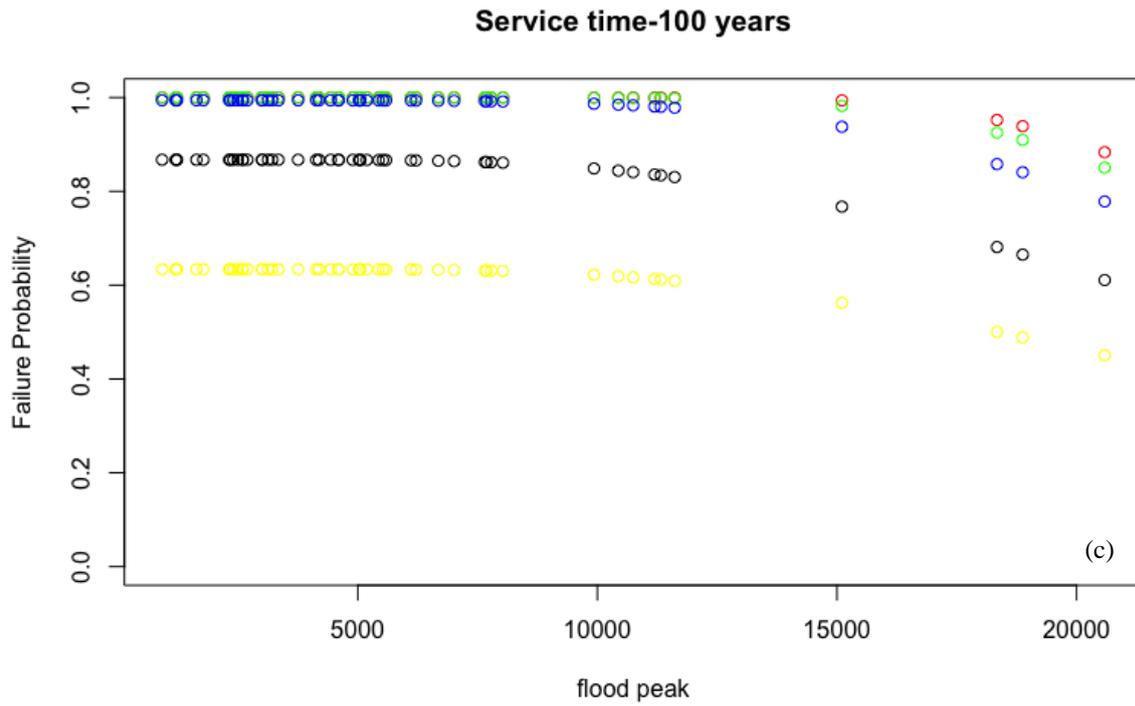
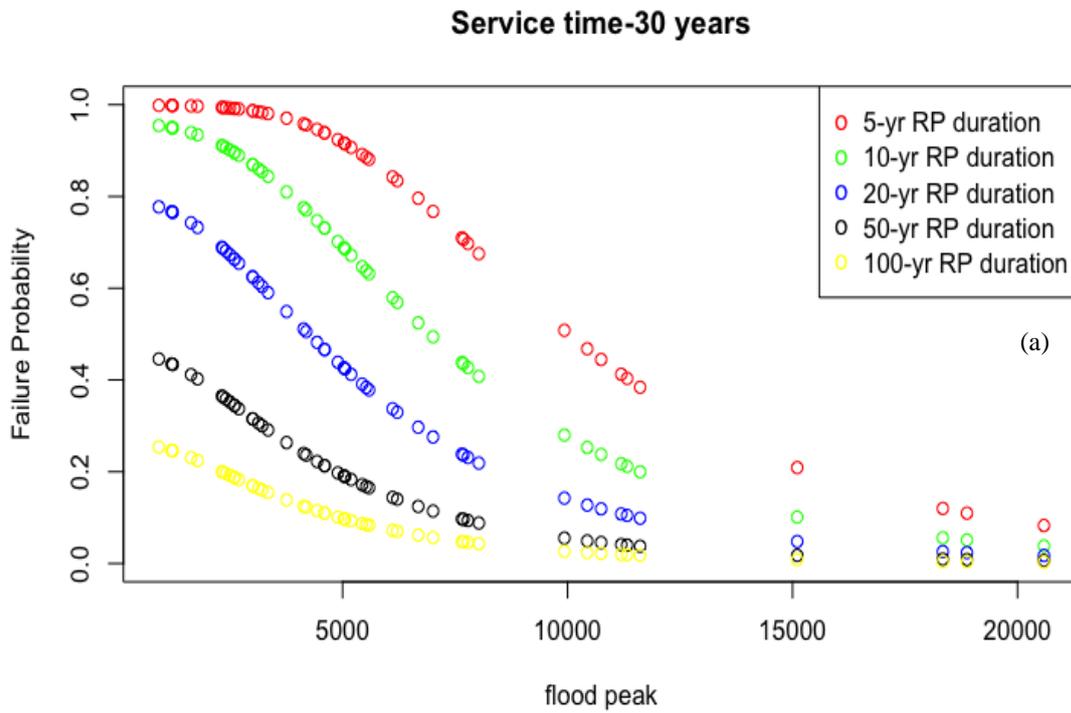


Figure 9. Graphical illustration indicating the variation of the bivariate hydrologic risk statistics with the change of flood peak flow in different synthetic or designed flood volume series a) 30-year service time b) 50-year service time c) 100-year services time (Red colour indicates 5-yr RP volume, Green colour indicates 10-yr RP volume, Blue colour indicates 20-yr RP volume, Black colour indicates 50-yr RP volume and Yellow colour indicates 100-yr RP volume series)



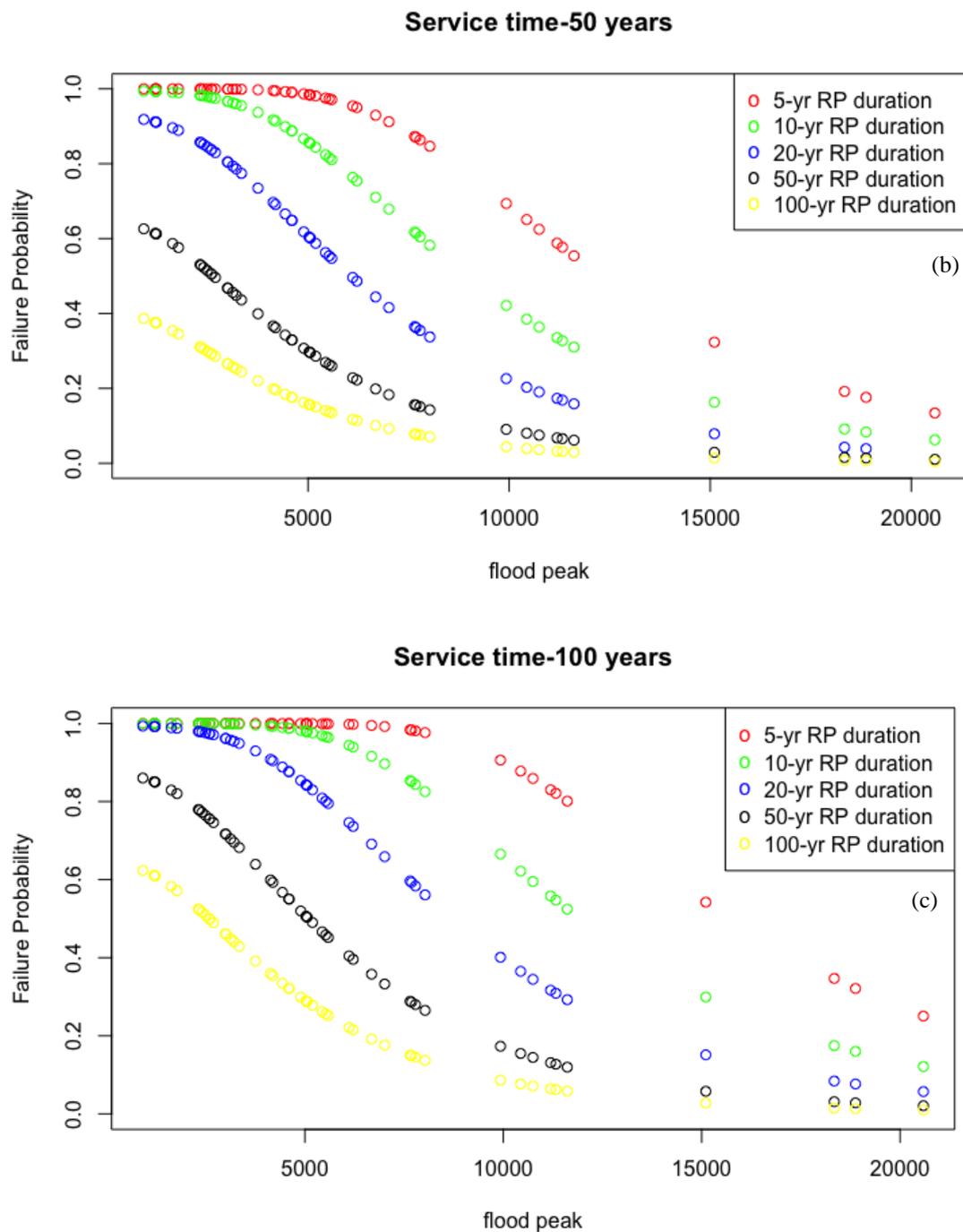


Figure 10. The variation of the bivariate hydrologic risk statistics with the change of flood peak flow in different synthetic or designed flood duration a) 30-year service time b) 50-year service time c) 100-year of service time

### 5. Conclusion

Water resources operational planning, managements or either flood defence infrastructure designs often demand the accurate estimations of flow exceedance probability for visualizing the risk of flood episodes. The Kelantan River basin in Malaysia is often affected by the most intensive and severe monsoonal flooding episodes. In this study, we incorporated the copula function in deriving the joint distribution of annual flood characteristics and failure probabilities for assessing the bivariate hydrologic risk. The Gaussian copula is recognized as the most parsimonious function for the flood pairs (P-V) and (P-D). The selected copula was then employed in deriving the joint return periods (RPs) for both the ‘OR’ and ‘AND’ case for a different possible combination of flood characteristics. It is revealed that the univariate RPs are higher than the ‘OR’ joint case but less than the ‘AND’ joint case, which further reveals that the RPs derived from univariate flood characteristics do not facilitate a sufficient and comprehensive approach to flood risk assessments. The bivariate hydrologic risk of flood episodes based on the joint relationship between flood pairs (P-V) and (P-D) was estimated using the definition of the failure probability (FP), which could be

useful in determining the risk level of extreme floods during the entire project design lifetime. The hydrologic risk analysis was carried out through two different investigation procedures. From the first investigation, it is inferred that the bivariate hydrologic risk statistics would increase by an increment in the service lifetime for the hydraulic facilities and at the same instant, both the bivariate and univariate hydrologic risk would incrementally decrease in return periods value. However, it could be revealed that the definition of return period is not capable of recognizing the probability or chance of the occurrence of extreme hydrologic events during the project lifetime.

In addition, univariate events produce lower FP statistics values in comparison with the bivariate events. Thus, it could be revealed that neglecting the mutual concurrency between multiple flood characteristics would be problematic and might underestimate the FP in the assessments of flood risk. Similarly, the variation of the bivariate hydrologic risk with the change of flood peak flow in different synthetic or designed flood volumes and flood durations for different service times are also demonstrated. It is concluded that hydrologic risk statistics incrementally increase by the project lifetime and at the same instant, incrementally decrease in the RP of the designed flood volume and duration series. Overall, this study provides a basis for making an appropriate flood defence plan and long-lasting flood defence infrastructure designs. According to Zhang [6], Bender et al. [63], Sarhadi et al. [64] the validity or accuracy in the estimated flood design quantiles under the time-invariant or stationary risk framework might be questionable, due to the ignorance of the accountability of the changing environment (Climate change or LULC change) in the flood distribution analysis. The existence of the non-stationarity usually tries to interrupt the hydrological behaviour within the catchment region thus might alter the expectation of such extreme happening under the stationary hydrologic risk assessment [65]. In other words, the actual associated flood risk either greater or smaller than the hazards statistics accounted under stationary risk concept and might reveal for under-dimensioned or over-dimensioned in the designing of hydraulic structure [66]. All the above raised issues can be considered as a future research purpose for the same river basin in order to achieve much practical and justifiable flood hazard assessments.

## 6. Conflicts of Interest

The authors declare no conflict of interest.

## 7. References

- [1] Milly, P. C. D., R. T. Wetherald, K. A. Dunne, and T. L. Delworth. "Increasing Risk of Great Floods in a Changing Climate." *Nature* 415, no. 6871 (January 2002): 514–517. doi:10.1038/415514a.
- [2] Gaume, Eric, Valerie Bain, Pietro Bernardara, Olivier Newinger, Mihai Barbu, Allen Bateman, Lotta Blaškovičová, et al. "A Compilation of Data on European Flash Floods." *Journal of Hydrology* 367, no. 1–2 (March 2009): 70–78. doi:10.1016/j.jhydrol.2008.12.028.
- [3] 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.
- [4] Goel, N. K., S. M. Seth, and Satish Chandra. "Multivariate Modeling of Flood Flows." *Journal of Hydraulic Engineering* 124, no. 2 (February 1998): 146–155. doi:10.1061/(asce)0733-9429(1998)124:2(146).
- [5] 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.
- [6] Zhang, L. "Multivariate hydrological frequency analysis and risk mapping". Doctoral dissertation, Beijing Normal University (2005).
- [7] Yue, Sheng, and Peter Rasmussen. "Bivariate Frequency Analysis: Discussion of Some Useful Concepts in Hydrological Application." *Hydrological Processes* 16, no. 14 (2002): 2881–2898. doi:10.1002/hyp.1185.
- [8] Zhang, L., and Vijay P. Singh. "Trivariate Flood Frequency Analysis Using the Gumbel–Hougaard Copula." *Journal of Hydrologic Engineering* 12, no. 4 (July 2007): 431–439. doi:10.1061/(asce)1084-0699(2007)12:4(431).
- [9] Sraj, Mojca, Nejc Bezak, and Mitja Brilly. "Bivariate Flood Frequency Analysis Using the Copula Function: a Case Study of the Litija Station on the Sava River." *Hydrological Processes* 29, no. 2 (January 30, 2014): 225–238. doi:10.1002/hyp.10145.
- [10] Yue, S, T.B.M.J Ouarda, B Bobée, P Legendre, and P Bruneau. "The Gumbel Mixed Model for Flood Frequency Analysis." *Journal of Hydrology* 226, no. 1–2 (December 1999): 88–100. doi:10.1016/s0022-1694(99)00168-7.
- [11] 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.
- [12] Bačová Mitková, Veronika, and Dana Halmová. "Joint Modeling of Flood Peak Discharges, Volume and Duration: a Case Study of the Danube River in Bratislava." *Journal of Hydrology and Hydromechanics* 62, no. 3 (September 1, 2014): 186–196. doi:10.2478/johh-2014-0026.

- [13] 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.
- [14] 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.
- [15] 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.
- [16] Sklar, M. "Fonctions de repartition an dimensions et leurs marges." *Publ. inst. statist. univ. Paris* 8 (1959): 229– 231.
- [17] 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.
- [18] Nelsen, R. B. "An Introduction to Copulas." *Springer Series in Statistics* (2006). doi: 10.1007/0-387-28678-0.
- [19] 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.
- [20] Brunner, Manuela Irene, Jan Seibert, and Anne-Catherine Favre. "Bivariate Return Periods and Their Importance for Flood Peak and Volume Estimation." *Wiley Interdisciplinary Reviews: Water* 3, no. 6 (September 4, 2016): 819–833. doi:10.1002/wat2.1173.
- [21] Latif, Shahid, and Firuza Mustafa. "Copula-Based Multivariate Flood Probability Construction: a Review." *Arabian Journal of Geosciences* 13, no. 3 (January 30, 2020). doi:10.1007/s12517-020-5077-6.
- [22] Requena, Ana I., Isabel Flores, Luis Mediero, and Luis Garrote. "Extension of Observed Flood Series by Combining a Distributed Hydro-Meteorological Model and a Copula-Based Model." *Stochastic Environmental Research and Risk Assessment* 30, no. 5 (August 7, 2015): 1363–1378. doi:10.1007/s00477-015-1138-x.
- [23] 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.
- [24] Qiang, HUANG, and CHEN Zishen. "Multivariate Flood Risk Assessment Based on the Secondary Return Period." *Journal of Lake Sciences* 27, no. 2 (2015): 352–360. doi:10.18307/2015.0221.
- [25] Read, Laura K., and Richard M. Vogel. "Reliability, Return Periods, and Risk Under Nonstationarity." *Water Resources Research* 51, no. 8 (August 2015): 6381–6398. doi:10.1002/2015wr017089.
- [26] Moftakhari, Hamed R., Gianfausto Salvadori, Amir AghaKouchak, Brett F. Sanders, and Richard A. Matthew. "Compounding Effects of Sea Level Rise and Fluvial Flooding." *Proceedings of the National Academy of Sciences* 114, no. 37 (August 28, 2017): 9785–9790. doi:10.1073/pnas.1620325114.
- [27] Salvadori, G., F. Durante, C. De Michele, M. Bernardi, and L. Petrella. "A Multivariate Copula-Based Framework for Dealing with Hazard Scenarios and Failure Probabilities." *Water Resources Research* 52, no. 5 (May 2016): 3701–3721. doi:10.1002/2015wr017225.
- [28] Serinaldi, Francesco. "Dismissing Return Periods!" *Stochastic Environmental Research and Risk Assessment* 29, no. 4 (July 9, 2014): 1179–1189. doi:10.1007/s00477-014-0916-1.
- [29] Xu, Ye, Guohe Huang, and Yurui Fan. "Multivariate Flood Risk Analysis for Wei River." *Stochastic Environmental Research and Risk Assessment* 31, no. 1 (December 17, 2015): 225–242. doi:10.1007/s00477-015-1196-0.
- [30] Xu, Hongshi, Kui Xu, Jijian Lian, and Chao Ma. "Compound Effects of Rainfall and Storm Tides on Coastal Flooding Risk." *Stochastic Environmental Research and Risk Assessment* 33, no. 7 (June 20, 2019): 1249–1261. doi:10.1007/s00477-019-01695-x.
- [31] Drainage, D. I. D. "Irrigation Department. Annual flood report of DID for Peninsular Malaysia." Unpublished report. Kuala Lumpur: DID (2004).
- [32] MMD (Malaysian Meteorological Department). "Report on Heavy Rainfall that Caused Floods in Kelantan and Terengganu." Unpublished report. MMD: Kuala Lumpur (2007).
- [33] Hussain, S.T.P.R, and Ismail H. "Flood frequency analysis of Kelantan River Basin, Malaysia." *World Applied Science Journal*, 28(12) (2013): 1989-1995. <https://doi.org/10.11113/mjce.v30n3.515>.
- [34] 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.
- [35] 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.

- [36] Chan, Ngai Weng. "Institutional Arrangements for Flood Hazards in Malaysia: An Evaluation Using the Criteria Approach." *Disasters* 21, no. 3 (September 1997): 206–222. doi:10.1111/1467-7717.00057.
- [37] Latif, Shahid, and Firuza Mustafa. "Parametric Vine Copula Construction for Flood Analysis for Kelantan River Basin in Malaysia." *Civil Engineering Journal* 6, no. 8 (August 1, 2020): 1470–1491. doi:10.28991/cej-2020-03091561.
- [38] Latif, Shahid, and Firuza Mustafa. "Bivariate Joint Distribution Analysis of the Flood Characteristics Under Semiparametric Copula Distribution Framework for the Kelantan River Basin in Malaysia." *Journal of Ocean Engineering and Science* (July 2020). doi:10.1016/j.joes.2020.06.003.
- [39] Latif, Shahid, and Firuza Mustafa. "Trivariate Distribution Modelling of Flood Characteristics Using Copula function—A Case Study for Kelantan River Basin in Malaysia." *AIMS Geosciences* 6, no. 1 (2020): 92–130. doi:10.3934/geosci.2020007.
- [40] Latif, Shahid, and Firuza Mustafa. "A Nonparametric Copula Distribution Framework for Bivariate Joint Distribution Analysis of Flood Characteristics for the Kelantan River Basin in Malaysia." *AIMS Geosciences* 6, no. 2 (2020): 171–198. doi:10.3934/geosci.2020012.
- [41] Drainage, D. I. D. "Irrigation Department. Annual flood report of DID for Peninsular Malaysia." Unpublished report. Kuala Lumpur: DID (2000).
- [42] 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.
- [43] Madsen, Henrik, Peter F. Rasmussen, and Dan Rosbjerg. "Comparison of Annual Maximum Series and Partial Duration Series Methods for Modeling Extreme Hydrologic Events: 1. At-Site Modeling." *Water Resources Research* 33, no. 4 (April 1997): 747–757. doi:10.1029/96wr03848.
- [44] Genest C and Favre A.-C. (2007). Everything you always wanted to know about copula modelling but were afraid to ask. *J. Hydrol. Engg.*, 12, 347-368.
- [45] 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.
- [46] 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.
- [47] Alina B. "Copula Modeling for World's Biggest competitors". Master Thesis, Amsterdam School of Economics, Universiteit Van Amsterdam (2018).
- [48] Genest, Christian, and Louis-Paul Rivest. "Statistical Inference Procedures for Bivariate Archimedean Copulas." *Journal of the American Statistical Association* 88, no. 423 (September 1993): 1034–1043. doi:10.1080/01621459.1993.10476372.
- [49] GENEST, C., K. GHOUDI, and L.-P. RIVEST. "A Semiparametric Estimation Procedure of Dependence Parameters in Multivariate Families of Distributions." *Biometrika* 82, no. 3 (1995): 543–552. doi:10.1093/biomet/82.3.543.
- [50] Schweizer, B., and E. F. Wolff. "On Nonparametric Measures of Dependence for Random Variables." *The Annals of Statistics* 9, no. 4 (July 1981): 879–885. doi:10.1214/aos/1176345528.
- [51] 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.
- [52] 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.
- [53] R Core Team. "R: A language and environment for statistical computing. R Foundation for Statistical Computing." Vienna, Austria. URL <http://www.R-project.org/>. (2013).
- [54] Zhang, L., and V. P. Singh. "Bivariate Flood Frequency Analysis Using the Copula Method." *Journal of Hydrologic Engineering* 11, no. 2 (March 2006): 150–164. doi:10.1061/(asce)1084-0699(2006)11:2(150).
- [55] Shiau, J. T. "Return Period of Bivariate Distributed Extreme Hydrological Events." *Stochastic Environmental Research and Risk Assessment (SERRA)* 17, no. 1–2 (May 1, 2003): 42–57. doi:10.1007/s00477-003-0125-9.
- [56] Adamowski, Kaz. "Nonparametric Kernel Estimation of Flood Frequencies." *Water Resources Research* 21, no. 11 (November 1985): 1585–1590. doi:10.1029/wr021i011p01585.
- [57] Dooge, James C. I. "Looking for Hydrologic Laws." *Water Resources Research* 22, no. 9S (August 1986): 46S–58S. doi:10.1029/wr022i09sp0046s.

- [58] 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.
- [59] 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.
- [60] 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.
- [61] Schwarz, Gideon. "Estimating the Dimension of a Model." *The Annals of Statistics* 6, no. 2 (March 1978): 461–464. doi:10.1214/aos/1176344136.
- [62] 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.
- [63] Bender, Jens, Thomas Wahl, and Jürgen Jensen. "Multivariate Design in the Presence of Non-Stationarity." *Journal of Hydrology* 514 (June 2014): 123–130. doi:10.1016/j.jhydrol.2014.04.017.
- [64] Sarhadi, Ali, María Concepción Ausin, and Michael P. Wiper. "A New Time-Varying Concept of Risk in a Changing Climate." *Scientific Reports* 6, no. 1 (October 20, 2016). doi:10.1038/srep35755.
- [65] Latif, Shahid, and Firuza Mustafa. "Copula-Based Multivariate Flood Probability Construction: a Review." *Arabian Journal of Geosciences* 13, no. 3 (January 30, 2020). doi:10.1007/s12517-020-5077-6.
- [66] Lima, Carlos H.R., Upmanu Lall, Tara J. Troy, and Naresh Devineni. "A Climate Informed Model for Nonstationary Flood Risk Prediction: Application to Negro River at Manaus, Amazonia." *Journal of Hydrology* 522 (March 2015): 594–602. doi:10.1016/j.jhydrol.2015.01.009.