REVIEW PAPER

Copula-based multivariate flood probability construction: a review

Shahid Latif [1] · Firuza Mustafa [1]

Abstract

Basin perspective hydrology and hydraulic water-related queries often demanding an accurate estimation of flood exceedance probabilities or return periods for assessing hydrologic risk. The research on the advancements of flood probability modelling contributed to reduction of flood risk, damage property and human life losses associated with the occurrence of flood events. Higher degree of uncertainty and complex flood dependence structure did not facilitate for their accurate prediction through deterministic approaches, which often demand a probability distribution framework. Unreliability of univariate frequency analysis under parametric or non-parametric framework would be an attribute for underestimation or overestimation of flood risk. Multivariate distribution framework facilitating a comprehensive understanding of flood structure for various possible occurrence combinations among the flood-related random vectors (i.e. flood peak flow, volume and duration). In this literature, copula function is recognized as a highly flexible tool for establishing multivariate joint dependency and their associated return periods in comparison with traditional multivariate functions. The incorporation of vine or pair-copula constructions (or PCC) further exaggerated the efforts of higher dimension copula construction, in terms of precision level in their estimated quantiles, under the minimum information concept. This review explored the efficacy of copula-based methodology for tackling multivariate design problems and can be used as a guideline for water practioner and hydrologist.

Keywords Flood frequency analysis (or FFA) · Copulas · Joint distribution · Conditional distribution · Return period · Pair-copula constructions (or PCC)

Introduction

Risk-based hydrology and hydraulic water-related queries, i.e. engineering-based flood defence infrastructure designs or the non-structural assessments (i.e. flood control pressure or flood diversion practices) often demanding an accurate estimation of flood exceedance probability or flow quantiles through extrapolating long-term catchments, have streamflow characteristics in the light of probability distribution framework (Bobee 1975; Rao 1980; Singh and Singh 1988; Chow et al. 1988; Bobee and Ashkar 1989; Adamowski, 1989; Adamowski and Feluch 1990; Cunnane 1987, 1988, 1989; Adamowski and

Responsible Editor: Tajudeen Iwalewa

✉ Shahid Latif
macet.shahid@gmail.com

Firuza Mustafa
firuza@um.edu.my

[1] Department of Geography, University of Malaya, 50603 Kuala Lumpur, Malaysia

Feluch 1990; Choulakian et al. 1990; Bras 1990; Yue et al. 1999; Yue 1999, 2000; Rao and Hameed 2000; Shiau 2003; Salvadori 2004; Sraj et al. 2014; Serinaldi 2015; Sarhadi et al. 2016). In actuality, the higher degree of uncertainty and complexness usually distributed over the hydrological or flood characteristics did not facilitate for their exact or accurate prediction in the light of any physical or deterministic framework but could be demanding to establish a probabilistic framework (Sen 1999; Requena et al. 2016). Therefore, several mathematical or statistical strategies are often motivated from past decades towards the incorporations of probability distribution framework of the hydroclimatic or flood observations series (i.e. Hosking et al. 1985; Adamowski 1985; Silverman 1986; Bardsley 1988; Adamowski and Labatiuk 1987; Bobee and Rasmussen 1994; Goel et al. 1998; Yue et al. 2001; Yue 2001a, 2001b; Coles 2001; Kartz et al. 2002 and references therein). Flood signifies for an inundation attributed through an overflowing of river water from their banks, due to an abnormality in hydrometeorological consequences, such as intensive precipitation structure (Reddy and Ganguli 2012a). Flood frequency analysis or FFA statistically defines the inter-association between extreme event quantiles and their non-

exceedance probabilities by fitting a probability distribution functions or pdf (i.e. either flood peak or volume as functions with their non-exceedance probabilities) (Yue 1999; Yue 2001a; Yue and Rasmussen 2002; Yue and Wang 2004; Xu et al. 2015).

Hydrometeorological stimulations either via the extension of historical rainfall samples, in order to recognize catchments profile, or through joint probability simulations in conjunction with univariate or multivariate statistical framework over the variables of interest are the two distinct ways to address the risk assessments for an extreme flood scenario. Numerous attempts, i.e. Calver and Lamb (1995), Boughton et al. (2002), Blazkova and Beven (2004) and Lawrence et al. (2014), retrieved flood frequency curve through integrating hydrological models in conjunction with probabilistic rainfall models for demonstrating the catchment's rainfall-runoff profile. Such incorporation usually adapted conventional based lumped and distributed models or via continuous or event-based hydroclimatic simulations. But the requirement of longer computational analysis due to demands of high spatial and temporal resolutions, in order to reveal a satisfactory demonstration of flood stimulation procedure, would attribute for an ineffective characterization of catchments behaviour (Requena et al. 2016). Similarly, few other approaches in flood analysis, which is based on data-driven method of flood predictions using stochastic or time series model, i.e. AR (or autoregressive), ARMA (or autoregressive moving average) and ARIMA (or autoregressive integrated moving average), such as generation of synthetic flow series and their forecasting using ARMA model (O'Connel 1977), time series modelling of annual maximum observation using ARIMA model (Shakeel et al. 1993), forecasting of rainfall and runoff using stochastic time series modelling using AR model (Sherring et al. 2009), generation and forecasting of annual inflow observations using ARMA model (Vijayakumar and Vennila 2016) demonstrate the efficacy of ARIMA model for flood forecasting using the annual streamflow (i.e. peak and maximum discharge) observations for Karkheh River basin in the west of Iran (Machekposhti et al. 2017). Besides this literature, few other studies, such as Ghanbarpour et al. (2010), Tian et al. (2011), Huang et al. (2016) and references therein, explored the efficacy of stochastic time series modelling approaches for solving several water resources problems. 'Development of few ML algorithms in the flood prediction' discussed over the developments of few other machine learning (ML) algorithms in the field of flood modelling and their forecasting. Overall, in the above demonstrations, their concern is only limited to target single flood vector, i.e. annual peak or maximum discharge values. Flood is a trivariate stochastic consequences, usually characterized completely through its intercorrelated random vectors, i.e. flood peak discharge, volume and duration of flood hydrograph (Zhang and Singh 2007b; Veronika and Halmova 2014). Thus, it could

limit the reliability of univariate design estimations or return periods, which would be insignificant to providing full screen of flood hydrograph and might be attributes for the underestimations (i.e. low design value might increase the risk of failure) or overestimations (i.e. increasing hydraulic construction cost) of hydrologic risk (Grimaldi and Serinaldi 2006; Serinaldi and Grimaldi 2007; Genest et al. 2007; Grimaldi et al. 2013; Fan and Zheng 2016). Such that, flood events with a peak flow of 100-year recursion interval could be less intensive and damaging than the same events described based on the joint occurrence between multiple flood vectors, i.e. between peak-volumes or peak-durations or volume-durations. In actuality, the potential damage that could likely be a function of several associated random variables as well as ignorance of spatial dependency among multiple flood vectors might be attributes for underestimation of uncertainty distributed over the estimated design quantiles and thus often demand more flood variables through the joint distributional assessments for revealing much insightful understanding of flood structure (Renard and Lang 2007; Graler et al. 2013; Vernieuwe et al. 2015). Especially, from the prospects of hydraulic designing procedures where an accountability of multivariate design parameters could be a feasible desire based on their multivariate exceedance probabilities (Brunner et al. 2016; Reddy and Ganguli 2013).

In multivariate risk statistics, return periods are usually associated with certain exceedance probabilities that will demonstrate the risk of extreme scenario through multiple aspects, i.e. based on the joint, conditional or Kendall's distribution relation (Shiau 2003; Salvadori 2004; Zhang and Singh 2006; Kao and Govindaraju 2008; Salvadori et al. 2011; Salvadori et al. 2015; Serinaldi 2015; Tosunoglu and Kisi 2016). According to Salvadori et al. (2011), under the hydraulic design facilities, selection of appropriate concurrence probabilities is a function of undertaken structure and consequences of its failure. The selection of return period is not an arbitrary process which solely based on the nature of work assessments that will further decide the importance of design vectors into considerations (Salvadori 2004; Serinaldi 2015; Brunner et al. 2016). Multivariate constructions usually comprise a combination of the joint probability density functions or pdfs and joint cumulative distribution functions or cdfs, where cdf statistically defines the probability of event 'X' less than their pre-defined critical or threshold values 'x', i.e. $P(X \leq x)$ (Yue and Rasmussen 2002; Veronika and Halmova, 2013). This literature intended towards overviewing the practices of copula-based stochastically synthetizations of flood consequences in the light of multivariate probability distribution framework. In this review, different methodological attempts in the light of bivariate and trivariate copula distribution analysis are pointed for tackling multivariate design problems or estimating design variable quantiles under different notations of return periods. Second section pointed the

different attempts and strategies towards the incorporations of univariate frequency analysis or defining the marginal distribution structure which often is a mandatory pre-requisite desire in copula distribution framework. Figure 1 illustrates research methodological flowchart of the literature review. Distinguished varieties of the one-dimensional parametric functions and also efficacy of non-parametric distributions for the treatment of hydroclimatic samples are reviewed in 'Flood frequency analysis via one-dimensional probability distribution framework or approximation of marginal distributions' under two different sub-sections. The necessity of establishing multivariate joint distribution of flood samples is pointed in the third section which are further divided into different sub-sections such that the applicability and flexibility of copula distribution for establishing bivariate joint relationship over the traditional multivariate functions, desire towards capturing flood design hydrograph by introducing all the relevant flood vectors simultaneously in the light of trivariate or 3-dimensional copulas construction, discussion over the distinguish varieties of some standard trivariate copulas and their efficacy for establishing joint distribution are reviewed. This section also pointed the flexibility of vine or PCC methodologies as well as minimum information PCC model for revealing many comprehensive attempts in the uncertainty analysis

of flood episodes in comparison with traditional trivariate copula functions. 'Return periods under multivariate settings' reviewed the importance of different notations of return period, i.e. joint return period, conditional return period, Kendall's return or survival return period which solely depends upon the nature of work assessments in the water-related issues. Development of few machine learning (or ML) algorithm in the field of flood prediction and forecasting is discussed separately in 'Development of few ML algorithms in the flood prediction'. 'Research discussion' and 'Research conclusion' comprise the research discussions and conclusions. Lastly, few ideas to strengthen the current attempts of multivariate practices in the light of time-varying copula framework are discussed in the last section of this literature.

Flood frequency analysis via one-dimensional probability distribution framework or approximation of marginal distributions

An approach via parametric distribution function

Hydrological episodes can be characterized through rare and extreme consequences according to prospects of time and
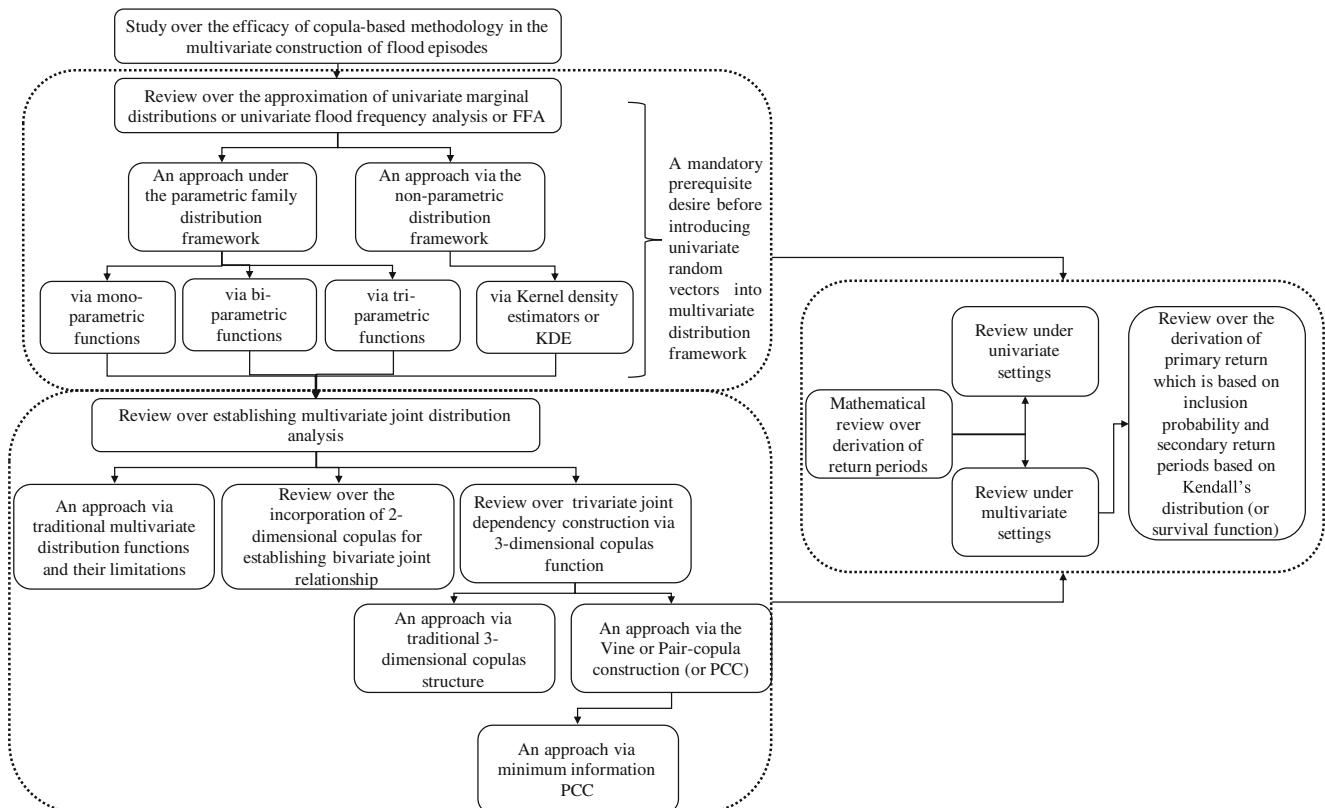

Fig. 1 Methodological flowchart of the literature review

magnitude scale. Three different approaches of flood modelling are usually motivated over the works of literature, i.e. regional-based analysis, stream-based and time-series analysis in which flood frequency estimations via annual peak discharge series could be effective for longer data length availability (Rao and Hameed 2000). Regional based hydrological modelling, which is also called pooling group analysis, targeted the data from multiple gauge site to derive a regional distribution of multivariate extreme which might reduce the chance of sampling variations in model parameters and, thus, would be effective for ungauge stream in comparison with the at-site frequency analysis (Burn 1990; Hosking and Wallis 1997; Viglione et al. 2007; Kyselý et al. 2011). Region-of-influence or ROI technique, i.e. based on the unique flexible pooling group for each targeted site (i.e. Burn 1990), and Hosking-Wallis or HW, i.e. based on delineating fixed regions and where each site characterized with same weight within the targeted region (i.e. Hosking and Wallis 1997) are the two-distinct variants or approaches of the regional based frequency analysis.

Conventional flood frequency practices frequently motivated either through block (annual) maxima (i.e. high flood peak) or peak over threshold on the partial series of data with an assumption of stationary, independent and identically (or i.i.d) distributions of historical samples (Hosking et al. 1985; Bras 1990; Coles 2001; Kartz et al. 2002). Annual maxima records often signify a justifiable basis of design problems such that the expected structural design life establishes a simple relation of its magnitude as well as their distributional structure and thus forms a basis to estimates the design quantiles or event exceedance by selecting an appropriate distributional structure of the given targeted maxima (Bardsley and Manly 1987). An interactive sets of univariate parametric families functions often targeted for univariate density modelling or defining marginal distributions of extreme random vectors such as 3-parameter generalized extreme value distribution (GEV) (i.e. Jenkinson 1955; Ouarda et al. 2001; Yue and Wang 2004), 2-parameter gamma distribution (i.e. Yevjevich 1972; Yue 2001a), 2-parameter with light tailed Gumbel distribution or extreme value type-1 distribution (or EV-1) (i.e. Adamowski, 1989; Yue et al. 1999), 2-parameter with bounded upper tailed or Weibull distribution (i.e. Johnson 1994; Zhang et al. 2016), 1-parameter exponential distribution (i.e. Choulakian et al. 1990; Bacchi et al. 1994; Karmakar and Simonovic 2008), 2-parameter log-normal distribution (i.e. Yue 2000; Xu et al. 2015), normal or Gaussian distribution (Goel et al. 1998; Yue 1999), log-logistic distribution (Bobee and Ashkar 1989), generalized logistic (or GLO) distribution (i.e. Requena et al. 2016), 3-parameter with heavy tailed Freshet distribution (i.e. Graler et al. 2013; Reddy and Ganguli 2013), 3-parameter general Pareto (or GP) distribution (i.e. Johnson 1994; Zhang et al. 2016), 3-parameter log-gamma distribution (i.e. Veronika and Halmova 2014) and log-Pearson type-3 distribution (i.e. Bobee 1975).

Generalized extreme value or GEV distribution exhibited a significant relation with hydrologist during extreme value practices which is further encompassed into three distinct functions such as Gumbel, the (Reversed) Weibull and Frechet distributions (i.e. Jenkinson 1955; Coles 2001; Khaliq et al. 2006). Each function attributed for the different tail behaviour based on their shape parameter 'ξ', i.e. Gumbel characterized with light tail behaviour, Frechet with heavy tail and bounded upper tail for the Weibull distribution (Graler et al. 2013). If the shape parameter 'α' is equal to 0 and correspond to thin upper and unbounded tail for GEV distribution then, it signifies for the Gumbel function and for $\alpha > 0$ termed for Frechet distribution, which signifies for the long and heavy tailed due to unbounded with decreasing behaviour, polynomially (Khaliq et al. 2006; Graler et al. 2013; Reddy and Ganguli 2013). The flexibility of available univariate models exhibited a control to justify an appropriate fit with distribution samples such that it depends upon its associated vectors of unknown statistical parameters or model parameters, i.e. 3-parameter log-gamma distribution extensively employed in flood modelling over many regions due to its capability of adjustments in their shape in accordance with the flood series (Veronika and Halmova 2014). Also, different density structures attribute different estimation of design quantiles, especially in the distribution tail structure (Karmakar and Simonovic 2008, 2009). Readers are advised to follow Coles (2001), Kartz et al. (2002) and Khaliq et al. (2006) for the extended details of the varieties of univariate models for hydrological observations.

An approach via non-parametric distribution framework

The above-cited literatures are frequently adapted the parametric distribution functions to approximating probability density or marginal distribution of flood characteristics. Simulations via the parametric functions often imposed an assumption that random samples are drawn from the population whose density structure is pre-defined, i.e. the marginal distribution of flood characteristics is assumed to follow some specific family of parametric density functions (Silverman 1986; Adamowski 1985, 1990, 1996; Botev et al. 2010). In actuality, no specific models are categorized and opted universally for any specific hydrologic variables, which would follow different distributions or, in other words, the best-fitted marginal distributions were not from the same probability distribution family (Adamowski 1985; Kim and Heo 2002; Karmakar and Simonovic 2008; Santhosh and Srinivas 2013). Dooge (1986) already pointed out that no amount of statistical refinement can overcome the consequences due to lack of prior probability distribution information of the observed random samples. Also, approximation of any distribution tail beyond the largest value under parameter distribution

framework would be difficult (Bardsley 1988; Bardsley and Manly 1987). More especially, in case of multimodal or skewed distributions where parametric functions might be incompatible and attribute for inconsistencies in the estimated quantiles. Therefore, from the last few decades, few demonstrations such as Schwartz (1967), Duins (1976), Singh (1977), Bowman (1984), Silverman (1986), Scott (1992), Lall et al. (1993), Lall (1995), Wand and Jones (1995); Jones and Foster (1996), Lall et al. (1996), Adamowski (1996, 2000), Bowman and Azzalini (1997), Efromovich (1999), Duong and Hazelton (2003), Kim et al. (2003), 2006), Ghosh and Mujumdar (2007) and Santhosh and Srinivas (2013) pointed the flexibility of non-parametric probability concept in the light of Kernel density estimations or kde. Kernel estimator is recognized as a much stable data smoothing procedure in the field of hydrologic or flood frequency analysis and which yields a bonafide density. Enumerations of the alternate theoretical overview for non-parametric setting are conducted in the earlier literature such as Rosenblatt (1956), Parzen (1962) and Bartlett (1963). Actually, the non-parametric framework does not require any prior distribution assumptions and will be directly retrieved from the distribution series with higher extent of flexibility as compared with parametric density estimators (Adamowski 1989 and Moon and Lall 1994).

Unless, the univariate approach defines the general concept of non-exceedance probability or return period via cumulative distribution function or cdf, but it might be unsatisfactory when the requirement demands the consideration of multivariate design parameters, which often reveals an essential concern in the water-related queries. Flood is a multidimensional phenomenon which often characterizes comprehensively by accounting its triplet intercorrelated random vectors and thus could demand the necessity of multivariate constructions for estimating the design hydrograph instead of just estimating design quantiles by targeting single flood vectors, i.e. univariate frequency analysis or return period (Choulakian et al. 1990; Bacchi et al. 1994; Goel et al. 1998; Yue 1999, Yue 2001a; Nadarajah and Shiau 2005). Actually, the selection of suitable recursion interval depends upon the selected design variable quantiles (Brunner et al. 2016) or, in other words, the importance of different notations of return periods, i.e. joint, conditional, Kendall's or survival functions solely depend upon the nature of assessments going to tackle in the water-related issues (Salvadori 2004; Serinaldi 2015). For examples, in the non-structural water-related queries, i.e. flood control and mitigation practices, demonstrating the mutual concurrency of flood peak with their volume extents would be a defensive approach in flood diversion practices or the joint dependency between flood peak and duration of events for flood controlling pressure practical (Fan et al. 2015; Xu et al. 2015).

Bivariate joint distribution framework of flood characteristics

Limitation of traditional multivariate distribution framework

Actually, capturing of correlation structure among the multiple hydrologic or flood vectors under classical statistical formulations such as Pearson correlation coefficient ('ρ') or Kendall's tau ('τ') would be ineffective to characterize co-movements tendencies of extreme vectors (Poulin et al. 2007). The unreliability and impractical consequences of univariate frequency analysis motivated numerous demonstrations towards multivariate joint probability constructions to investigate the mutual concurrency among flood vectors (Sackl and Bergmann 1987; Krstanovic and Singh 1987; Singh and Singh 1991; Raynal-Villasenor and Salas 1987; Cuadras 1992; Bacchi et al. 1994; Goel et al. 1998; Choulakian et al. 1990; Yue et al. 1999; Yue 1999, 2000, 2001a, 2001b; Yue and Rasmussen 2002; Durrans et al., 2003; Yue and Wang 2004; Nadarajah and Shiau 2005; Escalante 2007 and references therein). Distinguished varieties of traditional multivariate functions are incorporated for establishing bivariate joint relations and frequencies between flood peak-volume, volume-durations or peak-durations such as bivariate normal, lognormal and gamma functions (i.e. Yue 1999, 2000 and 2001), bivariate exponential distributional (i.e. Singh and Singh 1991; Choulakian et al. 1990), generalized extreme value distributions (i.e. Yue et al. 1999; Yue 2001b; Nadarajah and Shiau 2005), Pearson type III distribution (i.e. Durrans 1992), Gumbel mixed and Gumbel logistic functions (i.e. Yue and Wang 2004).

Multivariate practices via traditional probability functions often attribute several statistical constraints and shortcomings during the joint dependency measure, such that each individual hydrological entities or flood vectors will have an identical marginal structure or assumed to have Gaussian or normal distributions or either transformed or forced to have normal distribution through the data transformation procedure, which might be following the different marginal structures and would desire to model separately (Zhang 2005; Zhang and Singh 2006, 2007a; Reddy and Ganguli 2012a). Also, statistical parameters of the marginal structure are employed to model joint association, which often desire for separate modelling of their marginal and joint structure (Schmidt 2007). Limited space is usually available to justify the joint structure under conventional multivariate functions, thus often revealing a tough challenge (Song and Singh 2010). Besides this, conventional models attribute for the heavy dependency of flood exceedance on the right tail and thus might result for complexity during the demonstrations of observed samples and thus could demand for the separate modelling of margins from their joint dependence structure for securing their joint

association significantly (Zhang and Singh 2006; Reddy and Ganguli 2013). Actually, separate modelling of univariate marginal and their joint structure could optimize the reliability of the modelling outcomes (Ane and Kharoubi 2003 and Reddy and Ganguli 2012a).

After encountering the above limitations it motivated firstly, De Michele and Salvadori (2003) and Favre et al. (2004) introduce the concept of copula function as a model risk for hydrological observations. After that, a series of literature incorporated copula function, i.e. for flood samples (Salvadori and De Michele 2004; De Michele et al. 2005; Grimaldi and Serinaldi 2006; Zhang and Singh 2006; Zhang and Singh 2007b; Renard and Lang 2007; Genest et al. 2007; Salvadori et al. 2011; Grimaldi et al. 2013; Graler et al. 2013; Sraj et al. 2014; Daneshkhan et al. 2015; Bedford et al. 2015; Fan and Zheng 2016 and references therein), for rainfall characteristics (Salvadori and De Michele 2006; Zhang and Singh 2007a; Kao and Govindaraju 2008; Vernieuwe et al. 2015) and for drought episodes (Shiau 2006; Shiau and Modarres 2009; Song and Singh 2010; Ma et al. 2013; Saghafian and Mehdikhani 2014; Rauf and Zeephongsekul 2014; Zhang et al. 2016). Besides their extended applicability in extreme event modelling, copulas were significantly applied in the ground water modelling (Reddy and Ganguli, 2012) and also modelling of hydroclimatic samples (Maity and Kumar 2008 and Cong and Brady 2011). Actually, copulas segregate modelling of individual univariate vectors and their joint structure separately into two distinct stages, which attribute higher flexibility in selecting most appropriate and justifiable marginal and their joint structure among the peer family members to capture a wider extent of dependency, along with preservation in their joint association (Saklar 1959; De Michele and Salvadori 2003; Salvadori and De Michele 2004; and Nelsen 2006). The essential mathematical terminologies and theorems associated with copula function reader are advised to follow Saklar (1959) and Nelsen (2006) and also 'International Association of Hydrological Sciences (or IAHS)' for extended details and lists of their applicability in the field of hydroclimatological observations.

Copula-based bivariate probability distributions

In extreme hydrological modelling, the copula-based methodology can be classified as parametric, semiparametric and non-parametric estimation procedures depending upon the way of estimating its univariate marginals and joint dependence structure (Choros et al. 2010; Santhosh and Srinivas 2013). Current copula attempt in the recent decades (i.e. Favre et al. 2004; Grimaldi and Serinaldi 2006; Zhang 2005; Sraj et al. 2014 and references therein) frequently incorporated parametric settings for establishing multivariate flood distributions analysis using standard parametric distribution approach. On the other side, few demonstrations (i.e.

Karmakar and Simonovic 2008, 2009; Reddy and Ganguli 2012a) incorporated semiparametric copulas, also called heterogeneous or mixed marginal environment, where flood marginals are approximated using non-parametric distribution approach (i.e. kernel density estimators or orthonormal series) but still, parametric copula functions are introduced to modelled their joint dependencies. Besides this, few attempts (i.e. Dupuis 2007) pointed few limitations of the copula function in the context of finding best-fitted copula among their peer classes which is not a simple and consistent procedure and also limitation in the context of different extents of dependence measuring capabilities of each copula functions (Nelsen 2006). Therefore, literature (i.e. Santhosh and Srinivas 2013) incorporated the non-parametric approach for multivariate flood frequency analysis using the diffusion kernel functions which is earlier motivated by Botev et al. (2010).

Among the interactive sets of frequently incorporated copulas such as the extreme value class (i.e. Gumbel-Hougaard, Galambos and Husler-Reiss), elliptical class (i.e. Gaussian family), unclassified Plackett and Farlie-Gumbel-Morgenstern (or FGM) parametric functions and three-parametric Twan family (i.e. belong to extreme value class), the Archimedean class (i.e. Ali-Mikhail or A-M-H family, Frank family, Clayton or Cook-Johnson (C-J) family and Gumbel-Hougaard family) copulas are frequently accepted due to large varieties of families and its capability to capture joint dependencies for a wider extent also, exhibiting several desirable properties which attributes much flexibility during joint probability simulations (De Michele and Salvadori 2003; Salvadori and De Michele 2004; Favre et al. 2004; Nelsen 2006; Grimaldi and Serinaldi 2006; Zhang and Singh 2006; Salvadori and De Michele 2007; Corbella and Stretch 2013; Madadgar and Moradkhani 2013; Chebana et al. 2013; Rauf and Zeephongsekul 2014; Bender et al. 2014; Jiang et al. 2015; Papaioannou et al. 2016; Galiatsatou and Prinos 2016; Requena et al. 2016). Mathematically, the copula function (i.e. $[C : [0, 1]^2 \longrightarrow [0, 1]]$) approximates the bivariate Archimedean class copula, if it justifies the representation (i.e. $[C(u, v) = \varnothing^{-1}(\varnothing(u) + \varnothing(v))$ for $u, v \in [0, 1]]$), where $\varnothing(.)$ and $\varnothing^{-1}$ signify the generator function of the specified Archimedean copulas and their inverse such that the generator $(\varphi : I \longrightarrow \Re^+)$ signifies for the positive, convex and decreasing function and could be approximated for $\varnothing(1) = 0$ and $\varnothing(1) = \infty$ (Nelsen 2006). Each family of Archimedean class is characterized by a specific extent of dependency capturing capability, which is constrained by the degree of intersection between random vectors and will be investigated based on the dependency measure. As such, AMH family could model for both positive and negative associations but the dependence parameter is restricted for Kendall's tau $\tau_\theta \in [-0.181, 0.333]$ and could be insignificant for outside this range, similarly for C-J and GH family the Kendall's tau $\tau_\theta \geq 0$ and only

significant to capture the positive dependency (Salvadori and De Michele 2004; Nelsen 2006; Xu et al. 2015). Frank family functions exhibited higher versatility due to its capability in accommodating and capturing the entire range of dependencies (i.e. $\tau_\theta \in [1, -1]$) and only member that justified radial symmetry as well (i.e. symmetric to $u + v = 1$) (De Michele and Salvadori 2003; Favre et al. 2004; Nelsen 2006; Grimaldi and Serinaldi 2006; Zhang and Singh 2007a). All the Archimedean copulas except the Frank family exhibited non-symmetrical behaviour with respect to secondary diagonals such as the GH copula that is much suitable to model the dependence structure between vectors with upper-tail dependence; similarly, Clayton copula exhibited strong capability to model with lower tail dependency while Frank has no tail dependency (Poulin et al. 2007). Besides the above families, the extreme-value (or EV) copula is also incorporated for establishing bivariate joint relation, that can be formulated as $[C(u, v) = uv^{A(\log(u)/\log(uv))}]$, for $u, v \in [0, 1]$ which can be uniquely defined through Pickands dependence function (i.e. $[A : [0, 1] \longrightarrow [1/2, 1])$ and having non-symmetrical behaviour over the secondary diagonals (Twan 1988; Papaioannou et al. 2016). Nelsen (2006) demonstrated the extended examples for the Archimedean class functions; also see Twan (1988) for the extreme value functions.

Trivariate joint dependency constructions via 3-dimensional copulas

Unless extended efforts are often motivated towards establishing the copula-based methodology for estimating bivariate design variable quantiles under different notations of return periods, but such attempts still might be insufficient for revealing justifiable and comprehensive studies of flood probability analysis. Actually, dealing with multiple design variables, i.e. flood peak, volume and durations, would limit the applicability of analysing only through bivariate joint concurrency for the flood episodes but, due to its triplet distribution behaviour, could be demanding for simultaneous accountability of its all intercorrelated vectors (Salvadori et al. 2011; Graler et al. 2013; Fan and Zheng 2016; Reddy and Ganguli 2013; Daneshkhan et al. 2016). Actually, potential damage could likely be a function of multiple relevant vectors of specified hydrological episodes such that ignorance of spatial dependency among these uncertain vectors might be attributed for the underestimation of uncertainty, which frequently encountered during risk evaluation (Renard and Lang 2007; Graler et al. 2013; Vernieuwe et al. 2015). Few literatures incorporated copula-based methodology for establishing trivariate joint distribution and defining the concept of trivariate return periods by introducing an interactive class of 3-dimensional copula functions, but still, such computational strategies are quite limited over the literature.

Grimaldi and Serinaldi (2006) performed flood probability analysis through adapting three distinct forms of trivariate functions such as the mono-parametric and asymmetric or fully nested structure of Frank functions along with the Gumbel logistic distributions and pointed the significance of Frank function under FNA structure. Similarly, Serinaldi and Grimaldi (2007) fitted the same fully nested structure for deriving trivariate flood structure. Genest et al. (2007) adopted meta-elliptical copulas for annual spring flood analysis over Romaine River in Canada and revealed that it could be an effective tool for multidimensional hydrological data by preserving the pairwise dependencies among the random vectors through the correlation matrix but exhibited some modelling limitation such as might be ineffective under the low probabilities, unless the asymptotic properties of data will be justified through the strong arguments. Reddy and Ganguli (2013) examined the significance of multidimensional designs events by comparing univariate, bivariate and trivariate return periods for the flood episodes via fully nested Archimedean or FNA class copula and Student's t copula (Elliptical class copula) and revealed that it could be an essential effort to demonstrate joint and conditional flood occurrence in the light of trivariate return periods. Fan and Zheng (2016) adopted entropy copula structure in conjunction with Gibbs sampling along with the Gaussian and the Archimedean copula for simulation of trivariate flood episodes and revealed that entropy copula could be easily projected into higher dimensional frame directly just like as the Gaussian copula.

Similarly, Kao and Govindaraju (2008) applied the non-Archimedean copula function for simulating the trivariate structure of extreme rainfall episodes. This demonstration pointed out the modelling flexibility of Plackett family of copulas which concluded for faithful preservation of lower-level dependencies among relevantly associated vectors, which often reveal crucial strategies under trivariate or higher dimension dependency simulations. Madadgar and Moradkhani (2013) captured the joint behaviour of drought episodes under the climate change scenario using trivariate copula structure. This study integrated the significant drought vectors like severity, duration and its intensity using trivariate Gumbel copula (i.e. Archimedean family function) and t copula (i.e. elliptical family) framework for capturing joint and conditional probabilities. Also, the stress of dynamic environmental arising over the occurrence of future drought risk was also addressed through integrating GCM output under A 1 B scenario. Few other methodological efforts are of Song and Singh (2010) (i.e. drought frequency analysis under meta-elliptical copula structure), Wong et al. (2010) (i.e. modelling of trivariate drought characteristics) and so on.

The n-dimensional Archimedean copula can be formulated by extending two-dimensional form into 'n' order series, which can be as expressed by $[C(x_1, x_2, \ldots x_n) = \emptyset^{-1}(\emptyset(x_1) + \emptyset(x_2) \ldots \emptyset(x_n))]$ where consistency of this

equation will be preserved as long as the generator function $\varnothing(.)$ is fully monotonic; otherwise, it might be inconsistent for hydrological samples due to hypothesis in terms of homogenous dependency across the variables; also, '$\varnothing$' tends to the strict generator if pseudo-inverse function $\varnothing^{-1}(.)$ becomes as ordinary inverse function when $\varnothing(0)$ tends to infinity (Grimaldi and Serinaldi 2006; Nelsen 2006; Reddy and Ganguli 2013). From the perspective of lower dimension, i.e. bivariate copulas modelling, the symmetric Archimedean copulas frequently motivated over the literature and often justified significant outcomes through the inferential testing (i.e. good-of-fit test) measures but would be sounded for inconsistency when projecting into higher dimension distributional frame (i.e. $n{\geq}3$). In actuality, it approximates the dependencies between multiple vectors pairs by employing single dependence parameters due to its mono-parametric behaviour, but it would be incapable to preserve all pairwise dependency at the lower stages (Renard and Lang 2007; Genest et al. 2007; Kao and Govindaraju 2008; Madadgar and Moradkhani 2013). Therefore, it could desire to approximate each random pair individually through multiple parametric joint asymmetric functions (Serinaldi and Grimaldi 2007; Savu and Trede 2010; Reddy and Ganguli 2013). Whelan (2004) pointed a flexible structure that permit for the heterogeneous dependency across vectors in the context of fully nested Archimedean or FNA copulas which solely based on the joint integration of two or multiple bivariate or any dimensional Archimedean copulas structure through another Archimedean structure and can be formulated by $[C(x_1, x_2, x_3) = \varnothing_2(\varnothing_2^{-1} \circ \varnothing_1[\varnothing_1^{-1}(x_1) + \varnothing_2^{-1}(x_2)] + \varnothing_2^{-1}(x_3)) = C_2[C_1(x_1, x_2), x_3]]$ where: $\varnothing_1$ and $\varnothing_2$ signify Laplace transformation for first derivatives $[\varnothing_2^{-1} \circ \varnothing_1]$ will be monotonic; the symbol '$\circ$' indicates the composite of functions. The formulated copula $C(x_1, x_2, x_3)$ signifies the joint simulation of two bivariate structure through trivariate asymmetric Archimedean functions, but its applicability could be significantly justified only, if the dependency strength among the two variables i.e. (x_1, x_2) will dominate over the correlation structure between these variables and the third variables i.e. (x_1, x_3) and (x_2, x_3) (Savu and Trede 2010; Reddy and Ganguli 2013). Some literatures such as Grimaldi and Serinaldi (2006), Serinaldi and Grimaldi (2007), Madadgar and Moradkhani (2013) and Reddy and Ganguli (2013) demonstrated the flexibility of FNA structure for the hydrological observations. But some literature still pointed the issue of faithful preservations of lower stages dependency via the FNA structure and their modelling limitation which is only limited for positive range and thus pointed the applicability of few other standard class of trivariate copulas (i.e. Renard and Lang 2007; Genest et al. 2007; Kao and Govindaraju 2008; Fan and Zheng 2016). Such that Renard and Lang (2007) pointed Gaussian function (Elliptical class copula) for hydrological observations which can be significantly projected into any higher dimensional

frame directly due to the symmetric and definite positive matrix which can demonstrate the dependence between various attributes pairs. Genest et al. (2007) pointed the meta-elliptical copulas which could preserve pairwise dependencies via correlation matrix but exhibited some modelling limitation under low probabilities unless the asymptotic properties of data will be justified through strong arguments. Similarly, Kao and Govindaraju (2008) pointed the non-Archimedean Plackett families, which are based on the principle of the constant cross product, are another alternative to justify the preservations issue at lower-level dependencies. Ma et al. (2013) modelled trivariate drought characteristics via Gaussian and Student's t copula structure. Fan and Zheng (2016) highlighted the significance of maximum entropy theory in conjunction with entropy copula as a dynamic modelling strategy for higher dimensional space without imposing any assumption of copula family, more especially in conjunction with Gibbs sampling technique which could justified much comprehensive demonstration but surrounded with computational complexity due to the lack of analytical based parameter estimation. Justifiable preservation of all the lower-level dependencies often seems a challenging effort in the higher dimensional copula-based methodology especially, if complex pattern of dependency exhibited over the multidimensional data structure (Joe 1997; Kurowicka and Cooke 2006; Aas et al. 2009).

Vine copulas or PCC framework for trivariate joint distributions

The previous section highlighted different efforts and motivation towards simultaneous accountability of multiple design vectors via higher dimensional (i. e. $n \geq 3$) copula-based joint probability simulations for the hydrological characteristics, but still, such incorporations are quite limited. In actuality, the above-undertaken copulas encountered several statistical issues or queries such as complexity during the approximation of justifiable parametric distributions for higher dimensional hydrological attributes (Aas et al. 2009) and also might be quite ineffective to capture and reflect all the possible mutual concurrency among multidimensional vectors (Daneshkhan et al. 2016). In actuality, due to the higher degree of uncertainty and complexity, resolving the dependence structure of multivariate extreme via conventional copula formulation is quite complex, which often demands a flexible methodology through precise estimation of tail dependence coefficient under various tail dependency (Aas et al. 2009; Daneshkhan et al. 2016). Therefore, literature such as Kurowicka and Cooke (2006), Joe (1997), Aas et al. (2009) and Bedford and Cooke (2001); Bedford and Cook (2002) was directed towards a comprehensive way of uncertainty characterization for higher dimensional hydrological entities using the vine or pair-copula constructions (or PCC). Applicability of PCC

simulations seems much popular in finance and risk management (i.e. Aas et al. 2009; Czado and Min 2010; Nikoloulopoulos et al. 2012; Zhang 2014), but in the past few years, such incorporations were significantly recognized for the hydroclimatic observations such as frequency analysis for drought episodes (i.e. Song and Singh 2010; Saghafian and Mehdikhani 2014), for flood characteristics (i.e. Song and Kang 2011; Graler et al. 2013; Daneshkhan et al. 2016) and for storm or rainfall modelling (i.e. Gyasi-Agyei and Melching 2012; Vernieuwe et al. 2015).

Actually, vine copula construction are solely based on the principle of decomposition of full multivariate density into a cascade or simple local building blocks via conditional independence or pair-copula (Aas and Berg 2009; Bedford and Cook 2002; Graler et al. 2013). Due to conditional mixing via the stage-wise hierarchical nesting procedure, the pair-copula concept exhibited much effective and flexible modelling environments. Such multivariate simulation ignited from the works which is earlier demonstrated by Joe (1997) and after their underlying structural theory extended by Bedford and Cooke (2001); Bedford and Cook (2002) and Aas et al. (2009) and also Hobaek et al. (2010) demonstrated the different aspects in their structural and computational framework. Such incorporation usually comprises through the interactive sets of multiple bivariate or 2-dimensional copulas to cascade in fitting a copula to the random vectors and their conditional and unconditional distribution functions instead of introducing a fixed multidimensional structure to all the characteristics which might be attributed for ineffective over the data exhibited complex dependence structure in the tail and which often a stringent challenges in hydrological modelling (Joe 1997; Bedford and Cooke 2001; Bedford and Cook 2002). Distinct varieties of pair-copula decomposition are attributed under the regular vine structure such as canonical or C-vine and D-vine distribution is the two special modes of parametric regular vine construction (Kurowicka and Cooke 2006; Czado and Min 2010; Czado et al. 2013). Applicability of the D-vine structure is frequently sounded from the existed literature due to their higher flexibility than the C-vine structure and that would be effective when the existence of any particular vectors that regulate the level of mutual interactions within distributed observations are predefined or known (Aas et al. 2009; Daneshkhan et al. 2016). In actuality, the degree of mutual concurrency among multiple targeted vectors comprises the basis to adopt a justifiable vine tree structure (Graler et al. 2013). Let us suppose, for trivariate flood characteristics, if stronger association exhibited between the flood peak (P) and volume (V) and volume (V) and durations (D), that means it will be point to select D-vine structure by placing 'V' in between peak and durations. Czado et al. (2013) explored the extended details over the selection procedure of the regular vine constructions. The approximation capability of vine copula for multidimensional structure depends upon the manner of their decomposition and which further reveals that the choice of conditioning is not fixed or unique in vine or PCC (Hobaek et al. 2010; Graler et al. 2013). For further details of the C- and D-vine structure, readers are advised to follow Kurowicka and Cooke (2006), Aas et al. (2009) and Aas and Berg (2009).

Figure 2 illustrates the general computational flow of vine copula framework (i.e. Bedford and Cook 2002; Aas and Berg 2009; Aas et al. 2009; Czado et al. 2013; Graler et al. 2013; Daneshkhan et al. 2016). Computational strategies are usually initiated by selecting a significant vine structure, which solely depends upon the degree of mutual concurrency and having the following stages as reviewed from the above-mentioned literature,

- First stage of modelling
- Capturing correlation structure or pairwise dependency by selecting a justifiable bivariate parametric copula function for each flood pairs.
- Estimating conditional cumulative functions or 'h-functions' through conditioning each of the joint structure through variables, which share with both the other flood vectors, e.g. flood volume or V (from the above Fig. 2).
- Mathematically, conditioning structure can be deriving through the partial differentiation of each bivariate structure as formulated from the Eq. (1):

$$F_{P|V}(p|v) = \frac{\partial C_{PV}(p, v)}{\partial V} \; and \; F_{D|V}(d|v) = \frac{\partial VD(v, d)}{\partial V} \qquad (1)$$

where $C_{P, V}$ and $C_{V, D}$ signifies for bivariate copula structure; $F_{P|V}$ and $F_{D|V}$ defines conditional cumulative functions

- Second stage of modelling
- Synthesizing full density structure of 3-dimensional copula function using conditional CDFs of Eq. (1), for investigating the conditional CDFs using Eqs. (2) and (3) as the following

$$C_{PVD}(p, v, d)$$
$$= C_{PD|V}\big(F_{P|V}(p|v), F_{D|V}(d|v)\big).C_{PV}(p, v).C_{VD}(v, d) \quad (2)$$

also,

$$f_{PVD}(p, v, d) = C_{PVD}(p, v, d).f_X(x).f_Y(y).f_Z(z)$$
$$= C_{PD|V}\big(F_{P|V}(p|v), F_{D|V}(d|v)\big).C_{PV}(p, v). \qquad (3)$$
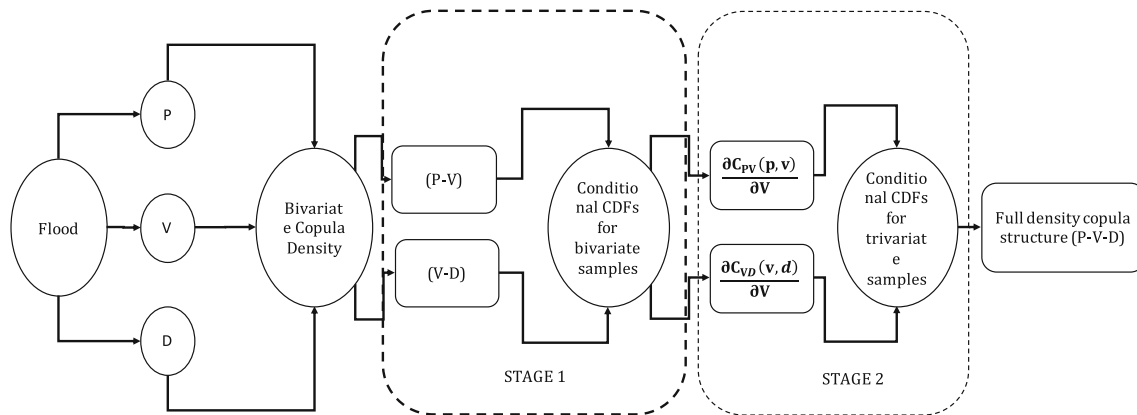$$C_{VD}(v, d).f_P(p).f_V(v).f_D(d)$$

Fig. 2 Stage-wise hierarchy of bivariate copulas density under 3-dimensional pair-copulas construction (or PCC)

An approach via minimum information PCC

Hydrological samples often surrounded with a higher degree of randomness and complexity in their multivariate dependence structure, which often attributes a stringent challenge to justify a precise approximation of multidimensional joint density structure. Also, justifiable accuracy in the estimated exceedance probability of river flow response often demands a longer duration of historical time series. The efficacy and modelling potential of vine copula construction for trivariate distributions are already reviewed from the above-cited literature but still have some modelling issues, i.e. complexity during the selection and synthesis of justifiable copulas structure under parametric density concept for vine constructions (Bedford et al. 2015). Therefore, a new methodological framework is pointed through introducing the concept of minimum information based vine framework. Such non-informative vine methodology facilitates a basis to further exaggerate the modelling potential of traditional PCC (i.e. which is defined via parametric copula framework) by approximating any undertaken copulas density to desire degree of approximations which is already demonstrated by Daneshkhan et al. 2016 for the trivariate flood distribution analysis. Minimum information PCC captures the complex multivariate structure for various tail dependencies through the precise estimation of tail coefficient for a given selected copula and also facilitates to model multivariate extreme in the presence of limited data length (Daneshkhan et al. 2016).

The fundamental concept of building the minimum information PCC for any bivariate joint densities structure say D_1 and D_2 can be demonstrated through establishing the relative information between undertaken densities which can be further minimized to 0 under identical bivariate densities (i.e. $D_1 = D_2$) as pointed from Eq. (4) (Bedford and Meenuwissen 1997 and Daneshkhan et al. 2016).

$$I(D_1, D_2) = \iint \ln \left(\frac{D_1 \left(x_1, x_2 \right)}{D_2 \left(x_1, x_2 \right)} \right) D_1(x_1, x_2) dx_1 dx_2 \qquad (4)$$

The generalized algorithmic explanation for establishing minimum information structure between any adjacent arbitrary pair of targeted extreme vectors, say between 'P' and 'V' under vine model, can be formulated by integrating the concept of moments constraint using Eq. (5) (i.e. Bedford et al. 2015; Daneshkhan et al. 2015, 2016)

$$\phi_i(P, V) = \phi_I' \left(F_1^{-1}(P), F_2^{-1}(V) \right), \quad \text{for } i = 1, 2, \ldots, k \qquad (5)$$

where $F_1^{-1}(P)$ and $F_2^{-1}(V)$ represent the univariate cumulative functions of the targeted vectors. Selection of appropriate basis functions (i.e. ϕ_i for $i = 1, 2 \ldots k$) controls the fitness level of copula structure for each random pair (Daneshkhan et al. 2016). Also, selecting appropriate number of grid size has also influence over their approximation level such that larger value would attribute for longer computational period (Bedford et al. 2015). Therefore, it often demands a perfect synchronization or balancing between analysis duration and the accuracy level (Daneshkhan et al. (2016). Readers are advised to follow Bedford et al. (2015) and Daneshkhan et al. (2015, 2016) for extended details of this non-informative copula framework.

Return periods under multivariate settings

This section overviewed the statistical significance of return periods under multidimensional design concept for tackling different hydrologic problems. In actuality, selection of return periods is depending upon the importance of undertaken structure as well as its consequences of failure where their appropriate selection often attributed an impact over the strength of

design variables quantiles (Brunner et al. 2016). Hydrology and hydraulic applications mostly interested in the evaluation of the mean inter-arrival period between two design events which usually defined in a year called the return period (Shiau 2003; Salvadori 2004). Especially, the design quantiles define a higher return period often seems a feasible practice in the hydraulic structure designs (Requena et al. 2016). In multidimensional risk framework, return periods can be derived from the exceedance probabilities of flood attributes pair, such that joint return period retrieves from the joint exceedance probabilities. Estimating multivariate design variable quantiles under different notations of return periods, i.e. based on joint and conditional probability distribution functions or via Kendall's distribution or survival functions, is often a justifiable and essential concern in the hydrologic risk assessments (Salvadori 2004; Graler et al. 2013; Salvadori et al. 2013; Brunner et al. 2016). Shiau (2003), Salvadori (2004), Salvadori and De Michele (2004, 2007), Salvadori et al. (2011) and Serinaldi (2015) pointed an extended mathematical framework towards the deriving of different notations of return periods under copula-based methodology.

Primary return periods

The return periods can be segregated into two distinct groups, i.e. primary return period comprise via the inclusive probability such as 'AND and 'OR' return period and the secondary or 'Kendall' return period, which can be define based on the Kendall's probability distribution or survival function (Salvadori 2004; Salvadori et al. 2011; Salvadori et al. 2013). Concurrence probability usually define the probability that any extreme happening, i.e. flood episodes, which characterize through either univariate (say flood peak discharge or 'X') or multivariate variable (say 'X', 'Y'…) exceeding certain a threshold level say 'x' (or 'x', 'y'… for the multivariate structure) (Yue and Rasmussen 2002; Shiau 2003; Salvadori 2004). Under the one-dimensional probability framework, the return periods of hydrological or flood events exceeding a threshold value say $\{X \geq x\}$ can be defined through fitting univariate cumulative distribution functions or cdfs using Eq. (6), as given below:

$$T = \frac{\mu}{\text{total no of flood per year}} = \frac{\mu}{P(X \geq x)}$$
$$= \frac{1}{1 - \text{univariate cdf or } F(x)} \qquad (6)$$

where μ = mean inter-arrival duration between two consecutive episodes = 1, for annual maxima based extreme modelling (Yue and Rasmussen 2002).

In actuality, notation of return period under univariate concept might be useful only if the concentration of single hydrological attribute will justify the requirements of the design

process and, in another way, it will also indicate the existence of no significant inter-association exhibited between multiple relevant vectors (Veronika and Halmova, 2014). Each separate approach of return periods has their own significance, and that will be solely based on the nature of the undertaking problem, which cannot be interchanged and also impossible to decide for the most consistent ways (Serinaldi 2015). Therefore, the return periods which demonstrate the undertaken assessments have requirements in a much better way, only the things that create a sharp distinct through selecting most consistency and justifiable return period (Tosunoglu and Kisi 2016). Reddy and Ganguli (2013) demonstration revealed that the assessments of both primary (i.e. 'OR and 'AND) and secondary (i.e. Kendall's) return periods could be an effective practice more especially from hydraulic or flood defence infrastructure designing prospects, such that concentrating over only the return period in either 'OR' case or 'AND' might reveal under-dimensioned or over-dimensioned. Actually, the joint return period facilitates different possible ways to capture joint relationship between the targeted vectors such as under the bivariate distribution between flood vectors say 'X' and 'Y'; some alternative probability relations are given below (Yue and Rasmussen 2002; Salvadori 2004; Brunner et al. 2016).

•when both targeted vectors say 'X' and 'Y' simultaneously exceeds certain value say 'x' and 'y', i.e. $\{X > x, Y > y\}$,

•when only vector 'Y' exceeds the threshold say 'y', i.e. $\{X \leq x, Y > y\}$,

•when neither 'X' or 'Y' vector exceeds threshold, i.e. $\{X \leq x, Y \leq y\}$,

•when only vector 'X' exceeds threshold say 'x' $\{X > x, Y \leq y\}$.

Let us suppose, if $'X \geq x'$ and $'Y \geq y'$ are two potential flood vector, representing peak and volume series exceeding certain a threshold value say, 'x' and 'y', then according to statistics of return periods for the joint probability under 'OR' and 'AND' case (i.e. Yue and Rasmussen 2002; Salvadori 2004; Salvadori and De Michele 2004; Zhang and Singh 2006, 2007a) can be formulated using Eqs. (7) and (8):

For 'OR' case

$$T_{XY} = \mu/P(X \geq x \text{ OR } T \geq y) = \mu/1 - C[F(x), F(y)] \qquad (7)$$

similarly, and for 'AND' case

$$T'_{XY} = \mu/P(X \geq x \text{ AND } T \geq y)$$
$$= \mu/1 - F(x) - F(y) + C[F(x), F(y)] \qquad (8)$$

where $C[F(x), F(y)]$ signifies for copula joint density of flood margins $F(x)$ and $F(y)$ of the undertaken vectors and μ=mean inter-arrival duration of two successive episodes = 1, for annual maxima based extreme generations.

In most of the hydrological design requirements, it could be demanding to define events through highlighting the

significance or priority of one design variables over another design vectors and thus literature pointed out the necessity of conditional distributional framework for defining the concept of conditional return periods, i.e. Salvadori and De Michele (2004), Shiau (2006), Zhang and Singh (2006, 2007a), Kao and Govindaraju (2008), Salvadori and De Michele (2010), Salvadori et al. (2011), Vandenberghe et al. (2011), Rauf and Zeephongsekul (2014), Veronika and Halmova (2014), Salvadori et al. (2014), Saghafian and Mehdikhani (2014), Zhang et al. (2016), Brunner et al. (2016) and Tosunoglu and Kisi (2016). For example, probability of flood peak conditional to volume or durations or either flood volume conditional to peak or durations or flood durations conditional to flood peak or volume information would be benefited from the hydraulic design prospects. Let us consider if 'X' and 'Y' are the flood vectors then the conditional distribution of 'X' given various percentile value of 'Y' or vice-versa can be formulated using Eqs. (9) and (10):

$$P_{X/Y} = 1 - \left(P(x) - H(x,y)/1 - P(y) \right) \tag{9}$$

$$P_{Y/X} = 1 - (P(y) - H(x,y)/1 - P(X)) \tag{10}$$

Formulation of conditional probability framework under bivariate distributions between any pair of targeted flood vectors say 'X' and 'Y' can be formulated using Eqs. (11) and (12), for the various possible combinations in accordance with suitability or nature of the undertaken problem, as given below (Shiau 2003; Reddy and Ganguli 2012a; Veronika and Halmova, 2013).

$$P(X \leq x \backslash Y \leq y) = P(X \leq x, Y \leq y)/P(Y \leq y)$$
$$= H(X,Y) \text{ or } C(X,Y)/F(Y) \tag{11}$$

$$P(X \leq x \backslash Y \geq y) = P(X \leq x, Y \geq y)/P(Y \geq y)$$
$$= F(X) - C(X,Y)/1 - F(Y) \tag{12}$$

Similarly, Eqs. (13) and (14) represented the conditional cumulative function of Y given $X \geq x$, which can be expressed as

$$P(Y \leq y \backslash X \leq x) = P(Y \leq y, X \leq x)/P(X \leq x)$$
$$= C(X,Y)/F(X) \tag{13}$$

$$P(Y \leq y \backslash X \geq x) = P(Y \leq y, X \geq x)/P(X \geq x)$$
$$= F(Y) - C(X,Y)/1 - F(X) \tag{14}$$

where $H(X, Y)$ and $C(X, Y)$ signify the joint cumulative distributions, estimated using conventional and copula density structure of the univariate margins $F(X)$ and $F(Y)$ of targeted vectors 'X' and 'Y'. Therefore, the cumulative structure $H(X, Y)$ can be expressed in the context of bivariate copula density structure, say $'C(X, Y)'$ for the representation of conditional return period, as expressed from the Eqs. (15) and (16):

$$T_{X \backslash Y \geq y} = 1/\left(1 - (F(y))\left(1 - F(x) - F(y)\right.\right.$$
$$\left.+ (H(X,Y) \text{ or } C(X,Y))\right) \tag{15}$$

$$T_{y \backslash X \leq x} = 1/\left(1 - (F(x))\left(1 - F(x) - F(y)\right.\right.$$
$$\left.+ H(X,Y) \text{ or } C(X,Y)\right) \tag{16}$$

Really, it is very difficult and tough challenge under the design process to point out that which definition of return contours perform better and consistence measure of design event for attributing a justifiable and significant risk expectation for the undertaken water related problems.

Demonstrating the risk of supercritical extreme via the Kendall's distribution and survival functions (or secondary return periods)

Actually, utilizing the standard definition of return period solely in the light of inclusion probability or primary returns might be problematic and attributed for underestimation of correct value (Salvadori and De Michele 2010). In actuality, hydrologic consequences, i.e. flood, drought or rainfall, exhibited either the critical, sub-critical or super-critical behaviour. Primary return periods (i.e. joint and conditional distributions) for the annual flood analysis often attributing towards capturing of mean forecasting and would not facilitate to demonstrate the risk of supercritical or dangerous scenario which are rare and could be outlined by investigating mean time lapse between the occurrence of supercritical episodes (Salvadori and De Michele 2010; Salvadori et al. 2011; Vandenberghe et al., 2011; Mirabbasi et al. 2012). Appropriate reliability of hydraulic design system often intended towards the definition of exceedance probabilities for rare episodes (Sarhadi et al. 2016). Actually, the supercritical scenario of hydrological extremes often reveals a serious potential threat for designing facilities due to its rare happening risk in comparison with given design return periods (Graler et al. 2013; Reddy and Ganguli 2013). Therefore, it often demands to make a sharp distinction through the segregation of probability distribution space into a non-critical and super-critical region based on critical cumulative probability level through Kendall distribution functions or $'K_C'$ (Graler et al. 2013; Brunner et al. 2016). Thus, literature, i.e. Salvadori (2004), Salvadori and De Michele (2004) and Salvadori and De Michele (2007), demonstrated the efforts towards recognizing the concept of return period under supercritical extreme scenario for defining design events from the Kendall distribution functions also called secondary return period. Kendall return is usually demonstrated through an appropriate discrimination between the non-critical and supercritical episodes using critical cumulative probability levels and that will be further extended into the multidimensional

frame in the context of Kendall distribution function $'K_{C_\theta}(.)'$ (Graler et al. 2013). Under copula framework, Kendall joint return period can be derived from Kendall probability function under two different computational way, i.e. via analytically efforts or ether numerically, based on the simulation algorithm (Salvadori and De Michele 2007; Vandenberghe et al. 2010; Salvadori et al. 2011). According to Salvadori and De Michele (2010) and Salvadori et al. (2011), algorithmic expression for Kendall distribution can be pointed using Eq. (17):

$$[K_C(t) = \Pr[W \le t] = \Pr[C(U, \ V) \le t]], \tag{17}$$

where $W = C(U, \ V)$ signifies for univariate random variables and $K_C(t)$ only depends on $C(U, \ V)$ or copula function in which, copula level curves called isolines make a separation of distribution space into super-critical and a non-critical segments. Also, for the given probability level 't', Kendall's quantile can be demonstrated through an inverse of Kendall distribution (i.e. $q_t = K_C^{-1}(t)$) (Brunner et al. 2016). Actually, the above Kendall's equation facilitates to investigate the chance that random point in the unit square exhibited either larger or smaller copula value than a given critical probability level through the representation of multidimensional information via univariate form based on the cumulative function of copula's level curve (Salvadori et al. 2011; Graler et al. 2013). Efforts over the statistical evaluation of analytical expression for Kendall function are motivated by Ghoudi et al. (1998) and Salvadori and De Michele (2007) for the extreme value and Archimedean based bivariate copula distributions. On the other side, Salvadori et al. (2011) focused over the tackling of simulation algorithmic efforts (via numerical analysis) for defining $'K_C'$, in the absence of analytical expression. Salvadori et al. (2013) tackle some critical issues over standard Kendall's return estimation, which are actually pointed by Graler et al. (2013), through introducing the concept of survival function in conjunction with Kendall's return periods. According to Graler et al. (2013), it might be possible that few non-critical events reveal for larger value over any undertaken design value, but the conventional definition of Kendall's function attributes for longer joint concurrence probabilities for all the super-critical scenarios over the design value. Therefore, such computational challenges can be undertaken in the light of survival Kendall's structure by replacing Kendall's function by survival Kendall's function under copulas structure as demonstrated mathematically by Eq. (18) (Salvadori et al., 2013).

$$T_{\text{Kendall's survival}} = \frac{\mu}{1-\overline{K}_C(t)'} \text{ and } \overline{K}_C(t)'$$

$$= \Pr[C(1-U, 1-V) \ge t] \tag{18}$$

where $C(1 - U, 1 - V)$ = survival function of bivariate random vectors; $1-\overline{K}_C(t)'$ signifies the chance of multivariate extreme occurrence in the super-critical region at a critical probability level 't' (Salvadori et al. 2014). On the other side, survival Kendall quantile can be derived by replacing inverse of Kendall function through survival Kendall's distribution as $q_t = \overline{K}_C^{-1}(t)$ (Salvadori et al. 2014). Volpin and Fiori (2014) demonstrated structure-based concurrence probability estimations which could be an essential concern in the hydraulic designing facility. Such efforts hypothetically establish an interlinking between hydrological variables with their design parameters via strictly monotonic structure function as a statistically formulated equation and which can be facilitated for structural failure return period as demonstrated using Eqs. (19) and (20) (i.e. Volpin and Fiori 2014).

$$Z = g(X, Y) \tag{19}$$

$$T_Z = \frac{\mu}{1-F_Z(z)'} \tag{20}$$

where F_Z is the distribution structure of a random variables say, Z.

The most interesting desire just after the recognition of return periods is the characterisation of the most appropriate design. Multivariate nature of design problem often demands to select multiple design events for a given estimated return periods and that will be further parameterizing during the hydraulic designing procedure (Salvadori et al. 2011; Graler et al. 2013). In actuality, an infinite number of possible combinations among targeted or flood vectors corresponding to each concurrence probabilities under multidimensional framework often reveal a tough challenge during selecting the most promising and effective design process. Salvadori et al. (2011) pointed an efforts to justify the above requirements under the two different perspectives such that one approach concentrated through 'component-wise excess design realization' while other one focused 'most-likely design realization'. Selection of design under the later practices can be justified through targeting a point with the largest probability distributions (Salvadori et al. 2011; Graler et al. 2013). Besides this approach, Salvadori et al. (2014) focused another alternative, i.e. design realization via H-conditional approach which can be defined in the presence of ruling variable. As pointed by Brunner et al. (2016), multivariate simulation often yielded large outcomes; thus, the selection of just one design realization often reveals for flexibility. On the other side, many practitioners desire for much design information through selecting one design event via opting a sub-set of design process, which can be tackled through splitting a return curve into two distinct parts called naive and a proper part as demonstrated by Chebana and Ouarda (2009) or via across sampling of return contour plot according to likelihood function, called ensembles of design events (Graler et al. 2013). Few statistical significances of ensemble-based strategies are also pointed in the literature, i.e. Vandenberghe et al. (2010) and Salvadori et al. (2011).

For 2-D joint probability structure, it can be formulated using Eq. (21);

$$(U_1, U_2) = \text{argmax}_{T_{X_1 X_2}} f_{XY}\big(f_1^{-1}(u_1),\ f_2^{-1}(u_2)\big) \qquad (21)$$

Similarly, for 3-D probability frame, using Eq. (22);

$$(U_1, U_2, U_3)$$
$$= \text{argmax}_{T_{X_1 X_2, X_3}} f_{XYZ}\big(f_1^{-1}(u_1),\ f_2^{-1}(u_2), f_3^{-1}(u_3)\big) \quad (22)$$

Development of few ML algorithms in the flood prediction

Flood is often considered as the most destructive natural disaster and thus, often, motivated hydrologist and water practioner towards discovering the more efficient and accurate flood forecasting model or machine learning algorithm (or MLA) for the appropriate assessment extreme hazards. This section pointed the distinguish varieties of MLA, which are frequently and widely accepted among researchers in the treatment of hydroclimatic samples, also listed in the Table 1. ANNs or artificial neural networks is considered as one of the most interactive and efficient MLA in terms of more accurate approximations, higher modelling speed and could be able in modelling of complex flood structure (i.e. Mosavi et al. 2018; Li et al. 2010; Wu and Chau 2010; Jain and Prasad Indurthy 2004). Frequently applied for the modelling of river flow characteristics is rainfall-runoff modelling or prediction or extrapolation of streamflow characteristics, as revealed from Table 1. But, besides their several advantages, ANNs exhibited some

modelling issues in flood modelling such as complexity in data handling and network architecture (Deo and Sahin 2015). Besides this algorithm, the SVM or support vector machine, a supervised machine learning algorithm which works on the principal of statistical learning theory as well as rule of structural risk minimizations, is recognized as the most efficient and robust approach, especially in solving non-linear regression issues in the flood predictions and modelling (i.e. Ortiz-García et al. 2014; Gizaw and Gan 2016; Gong et al. 2016; Jajarmizadeh et al. 2015; Tehrany et al. 2015). Based on training from the historical observations, SVM can extrapolate the data for the future time frame; also in few literature, it is incorporated as a regression tools called Support vector regression (or SVR) (i.e. Li et al. 2016; Tehrany et al. 2015). The WNN or wavelet neural network is another most interactive machine learning approach in the time series extrapolations of flood characteristics, which is based on the principal of decomposition of initial observation sets into individual resolution levels. It is widely applied in the modelling of daily streamflow characteristics, rainfall-runoff as well as reservoir inflow modelling (i.e. Supratid et al. 2017; and Ravansalar et al. 2017). On the other side, the ANFIS or adaptive neuro-fuzzy inference system algorithm often poses quick and easy implementation as well as accurate and higher abilities in the learning procedure and thus often poses a good choice in the forecasting of flood episodes (i.e. Choubin et al. 2014; Lafdani et al. 2013; Shu and Ouarda 2008). Besides this, the decision tree (or DT) algorithm, which is based on the technique of tree of decision-making, is widely applicable in the prediction of flood events (Tehrany et al., 2014; Liaw and Wiener 2002), which further classified as fast algorithm (Tehrany et al. 2013), classification and regression tree (or CART) (i.e. Dehghani et al. 2017), random forest

Table 1 Machine learning algorithm (MLA) in the treatment of hydrologic samples

Algorithm	Applicability in the treatment of hydroclimatic samples	References
ANNs (artificial neural networks)	1. Forecasting of streamflow characteristics	Kisi (2007)
	2. River flow modelling	Badrzadeh et al. (2013)
	3. Rainfall-runoff modelling	Smith and Eli (1995)
WNN (wavelet neural network)	1. Rain-fall runoff modelling	Ravansalar et al. (2017)
	2. Reservoir inflow modelling	Supratid et al. (2017)
	3. Modelling of daily basis streamflow observations	Guimarães Santos and da Silva (2014)
SVM (Support Vector Machine)	1. Streamflow modelling	Jajarmizadeh et al. (2015)
	2. Estimation of Flood quantiles	Gizaw and Gan (2016)
	3. Precipitation modelling	Ortiz-García et al. (2014)
ANFIS (adaptive neuro-fuzzy inference system)	1. Modelling of short-term rainfall characteristics	Lafdani et al. (2013)
DT (decision tree)	1. Fast algorithm—flood prediction and modelling	Tehrany et al. (2013)
	2. The classification and regression tree (or CART)—flood simulation procedure	Dehghani et al. (2017)
	3. Random forests (RF) method—flood prediction	Liaw and Wiener (2002)
	4. M5 decision-tree algorithm—flood modelling	Etemad-Shahidi and Mahjoobi (2009)

method (or RFM) (i.e. Liaw and Wiener 2002) and M5 decision tree algorithm (i.e. Etemad-Shahidi and Mahjoobi 2009). All the above-mentioned ML algorithms are classified into two groups depending upon the length of samples or prediction lead-time under considerations such as short term and long term and which are further categorized as single and hybrid method.

Research discussion

Statistical inferencing of extreme hydroclimatic samples, for retrieving flow exceedance probabilities or return period, is often revealing an insightful concern for assessing hydrologic risk in the basin perspective water resources planning, management, and designing facilities. Actually, the hydrometeorological stimulations either via the extension of historical rainfalls samples or through joint distribution framework over the variables of interest are the two distinct ways to address the risk assessments for an extreme flood scenario. Few attempts extracted flood frequency curve via integrating hydrological models in conjunction with probabilistic rainfall models, i.e. either via the conventional based lumped and distributed models or via continuous or event-based hydroclimatic simulations. But due to the longer computational analysis, as it is demanding high spatial and temporal resolutions, it would be attributed for an ineffective characterization of catchments behaviour. Due to the higher degree of uncertainty and complex flood characteristics, it often demanding to establish a probability distributions framework instead of any deterministic procedure (Sen 1999 and Hosking et al. 1985). Multidimensional behaviour often demands the necessity of multivariate constructions for retrieving design variable quantiles under the different notations of return periods through accounting its multiple design vectors instead of just examining the univariate frequency relationship or return periods. In actuality, univariate frequency analysis would be incapable to recognize the full screen of flood or inflow hydrograph and thus could be demanding to introduced multiple intercorrelated flood vectors, i.e. flood peak, volume and its duration to establish joint probability density functions or pdf and joint cumulative distribution functions cdf, especially from the prospects of hydraulic designing procedures where accountability of multivariate design parameters could be a feasible desire based on their multivariate exceedance probabilities. In other words, selection of return period depends upon the importance of undertaken structure as well as its consequences of failure where their appropriate selection often attributed an impact over the strength of design variables quantiles (Brunner et al. 2016). Therefore, unreliability and impractical consequences of the univariate flood modelling motivated numerous demonstrations towards the development of a multivariate distribution framework via introducing distinguished varieties of traditional probability functions for

establishing bivariate joint relations between flood peak-volume, volume-durations or peak-duration, i.e. Choulakian et al., 1990; Yue et al. 1999; Yue 1999, 2000, and references therein), but due to several statistical shortcomings which often limited the applicability of traditional multivariate functions and thus motivated extended demonstrations in the light of bivariate copulas simulations under the parametric or semiparametric distribution settings (De Michele and Salvadori 2003; Salvadori and De Michele 2004; Salvadori 2004; Nelsen 2006; Karmakar and Simonovic 2009 and references therein). In multivariate risk statistics, return period usually associated with certain exceedance probabilities and their selection is not an arbitrary process which solely based on the nature of work assessments that will further decide the importance of design vectors into considerations.

During copula constructions, the approximations of the marginal distribution of univariate random vectors via the parametric functions would be problematic due to unsymmetrical or skewed distribution behaviour of hydrologic samples. Also, the parametric functions often imposed an assumption that random samples are drawn from the population whose density structure is pre-defined, i.e. the marginal distribution of flood characteristics is assumed to follow some specific family of parametric density functions, but in actuality, no universally accepted models are fixed or assigned to any of hydrologic vector. Thus, many literature pointed the flexibility of non-parametric probability concept in the light of Kernel density estimations or kde which is recognized as a much stable data smoothing procedure in the field of hydrologic or flood frequency analysis and yielding a bonafide density (i.e. Adamowski 1996, 2000; Ghosh and Mujumdar 2007; Santhosh and Srinivas 2013 and references therein). Distinguished varieties of copulas are incorporated such as the extreme value class (i.e. Gumbel-Hougaard, Galambos and Husler-Reiss), elliptical class (i.e. Gaussian family), unclassified Plackett and Farlie-Gumbel-Morgenstern (or FGM) parametric functions and three-parametric Twan family (i.e. belong to extreme value class) for establishing bivariate dependencies of hydroclimatic samples and among which the Archimedean class (i.e. Ali-Mikhail or A-M-H family, Frank family, Clayton or Cook-Johnson (C-J) family and Gumbel-Hougaard family) copulas are frequently accepted due to large varieties of families and its capability to capture joint dependencies for a wider extent also, exhibiting several desirable properties which attribute much flexibility during joint probability simulations (i.e. De Michele and Salvadori 2003; Salvadori and De Michele 2004; Favre et al. 2004; Nelsen 2006; Grimaldi and Serinaldi 2006; Papaioannou et al. 2016; Galiatsatou and Prinos 2016; Requena et al. 2016 and references therein). Each family of Archimedean class is characterized by a specific extent of dependency capturing capability, which is constrained by the degree of intersection between random vectors and will be investigated based on the

dependency measure. Unless extended, efforts are motivated towards establishing copula-based bivariate design estimations but such attempts still might be insufficient for revealing comprehensive studies of flood probability analysis due to its triplet distribution behaviour and could be demanding for the simultaneous accountability of its all intercorrelated vectors. The potential damage could be depending upon the multiple relevant vectors of specified hydrological episodes such that ignorance of spatial dependency among these random vectors might be attributed for underestimation of uncertainty (Renard and Lang 2007; Graler et al. 2013; Vernieuwe et al. 2015). Thus, a limited number of literature appeared in the context of 3-dimensional copula distribution analysis for establishing the trivariate joint relationship and their associated return periods (i.e. Reddy and Ganguli 2013; Graler et al. 2013; Daneshkhan et al. 2016 and references therein). Distinguished varieties of standard trivariate copulas are incorporated, i.e. Grimaldi and Serinaldi (2006) (mono-parametric and asymmetric or FNA structure of frank function), Serinaldi and Grimaldi (2007) (FNA structure), Genest et al. (2007) (meta-elliptical copulas), Reddy and Ganguli (2013) (FNA and Student's t copulas which belong to elliptical class) and Fan and Zheng (2016) (entropy copulas). Genest et al. (2007) revealed that the meta-elliptical copulas could be effective incorporation for preserving the pairwise dependencies among the random vectors through the correlation matrix but might be ineffective under the low probabilities unless the asymptotic properties of data will be justified through the strong arguments. Similarly, the flexibility of Plackett family of copulas which concluded for faithful preservation of lower-level dependencies during higher dimension modelling is pointed by Kao and Govindaraju (2008) for rainfall samples. Madadgar and Moradkhani (2013) captured the joint behaviour of drought episodes using trivariate Gumbel copula (i.e. Archimedean family function) and t copula (i.e. elliptical family). Some literature still pointed the issue of faithful preservations of lower stages dependency via the FNA structure and their modelling limitation which is only limited for positive range and thus pointed the applicability of few other standard class of trivariate copulas. Actually, justifiable preservation of all the lower-level dependencies often seems a challenging effort in the higher dimensional copula-based methodology especially, if the complex pattern of dependency exhibited over the multidimensional data structure. Also, it often demands a flexible methodology through precise estimation of tail dependence coefficient under various tail dependency. Therefore, literature such as Kurowicka and Cooke (2006), Joe (1997), Aas et al. (2009) and Bedford and Cooke (2001); Bedford and Cook (2002) directed towards a comprehensive way of uncertainty characterization for higher dimensional hydrological entities using the vine or pair-copula construction (or PCC). Actually, vine copula construction is solely based on the principle of the decomposition of full multivariate density into a cascade or simple local building blocks via conditional independence or pair-copulae (Aas and Berg 2009; Bedford and Cook 2002). Due to conditional mixing via the stage-wise hierarchical nesting procedure, the pair-copula concept exhibited much effective and flexible modelling environments. In PCC construction, interactive sets of multiple bivariate copulas in cascade form are often employed in fitting a copula to random vectors and their conditional and unconditional distribution, instead of just introducing a fixed multidimensional structure to all the characteristics and which might be attributed for ineffective over the data exhibited complex dependence structure in the tail and which often a stringent challenges in hydrological modelling (Joe 1997; Bedford and Cooke 2001; Bedford and Cook 2002). Distinct varieties of pair-copula decomposition are attributed under the regular vine structure such as canonical or C-vine and D-vine distribution in which the applicability of D-vine structure is frequently sounded from the existed literature due to their higher flexibility than the C-vine structure. In actuality, the degree of mutual concurrency among multiple targeted vectors comprises the basis to adopt a justifiable vine tree structure (Graler et al. 2013). The approximation capability of vine copula for multidimensional structure depends upon the manner of their decomposition, unless the modelling efficacy of PCC structure is reviewed from the above-cited literature but still having some modelling issues, i.e. complexity during the selection and synthesis of justifiable copula structure under parametric density concept for vine constructions (Bedford et al. 2015). Therefore, the concept of minimal information based vine structure is discussed such that this non-informative vine concept could further exaggerate the flexibility of conventional PCC structure (Daneshkhan et al. 2016). Actually, minimum information PCC captures the complex multidimensional flood structure for various tail dependency by the precise estimation of their tail coefficient for given selected copulas and also facilitates to the model multivariate extreme in the presence of limited data length (Daneshkhan et al. 2016).

The statistical significance of return periods under multidimensional design concept for tackling different hydrologic problems is reviewed in the separate section. Estimating multivariate design quantiles under different notations of return periods, i.e. based on joint and conditional probability distribution functions or via Kendall's distribution or survival functions, is often an essential concern in the hydrologic risk assessments (Salvadori 2004; Graler et al. 2013; Salvadori et al. 2013). Brunner et al. (2016), Shiau (2003), Salvadori (2004), Salvadori and De Michele (2004, 2007), Salvadori et al. (2011) and Serinaldi (2015) pointed the extended mathematical formulation of defining the different notations of return periods using copula-based methodology. In actuality, univariate return period might be useful only, if the concentration of single hydrological vector will justify the requirements of the

design process where each of the separate return period approach has their own significance and will be solely based on the nature of the undertaking problem, and also, it is much difficult task to decide for the most consistent ways (Veronika and Halmova, 2013; Serinaldi 2015). Therefore, the return periods demonstrate the undertaken assessment requirements in a much better way, only the things that create a sharp distinct through selecting most consistency and justifiable return period. According to Reddy and Ganguli (2013), considering both the primary as well secondary return period could be an effective practice more especially from the prospect of flood defence infrastructure designs. Such that concentrating over only the return period in either 'OR' case or 'AND' might reveal for under-dimensioned or over-dimensioned. Actually, the joint return period facilitates different possible ways to capture joint relationship for various possible combinations among the multiple flood vectors. In other words, for a given return period, various possible design combinations can be possible or vice versa. Besides the importance of joint return contour, most of the hydrological design requirements often demand to define events through highlighting the significance or priority of one design variables over another design vectors, i.e. conditional distribution or conditional return periods (i.e. Salvadori and De Michele 2004, Shiau 2006, Zhang and Singh 2006, 2007a, Kao and Govindaraju 2008, Salvadori and De Michele 2010, Salvadori et al. 2011). Such that probability of flood peak conditional to volume or durations or either the flood volume conditional to peak or durations or flood durations conditional to flood peak or volume information would be benefited from the hydraulic design prospects.

Hydrological consequences, i.e. flood, drought or rainfall, exhibited either the critical, sub-critical or super-critical behaviour; thus, in such circumstances, only the accountability of primary return period might be problematic or attributed for underestimation of correct value (Salvadori and De Michele 2010). Capturing of only mean forecasting would not facilitate to demonstrate the risk of supercritical or dangerous episodes which are rare. Appropriate reliability in hydraulic design facilities often intended towards defining the exceedance probabilities of rare episodes (Sarhadi et al. 2016). Therefore, it could be demanding to make a sharp distinction through segregating probability distribution space into a non-critical and super-critical region based on the critical cumulative probability level and will be further extended into the multidimensional frame in the context of Kendall distribution function $'K_{C_\theta}$ $(.)'$ (i.e. Salvadori 2004; Salvadori and De Michele 2004; Salvadori and De Michele 2007 and Graler et al. 2013). The analytical efforts or ether through numerical approach based on simulation algorithm are the two different computational ways to estimate the Kendall joint return period, derived from Kendall probability function under the copula distribution framework (Salvadori et al., 2007; Vandenberghe et al.

2010; Salvadori et al. 2011). The analytical expression for Kendall function are motivated by Ghoudi et al. (1998) and Salvadori and De Michele (2007) for the bivariate extreme value and Archimedean copula distributions while, Salvadori et al. (2011) focused, tackling of simulation algorithmic efforts (via numerical analysis) for defining $'K_C'$ in the absence of analytical expression. It might be possible that few non-critical events reveal for larger value over any undertaken design value, where the Kendall's function attributes for longer joint concurrence probabilities for all the super-critical scenario over the design value thus can be undertaken in the light of survival Kendall's function (Salvadori et al. 2013). This structure-based concurrence probability estimations establishing an inter-association between hydrological characteristics and design parameters via the strictly monotonic structure function as a statistically formulated equation and which can be facilitated for structural failure return period (i.e. Volpin and Fiori 2014). Due to multivariate nature of the design problems, it often demands the characterization of most justifiable design for a given estimated return period requirement under the two different perspectives, i.e. one approaches concentrated through 'component-wise excess design realization' while the other one focused 'the most-likely design realization' (Salvadori et al. 2011). The design realization via H-conditional approach is another alternative which can be defined in the presence of ruling variable (Salvadori et al. 2014).

Research conclusions

Basin perspective water resources operational planning, managements or the hydraulic structural designs often demand an accurate estimation of flow exceedance probability for assessing the flood risk. Due to higher degree of uncertainty and complex flood dependence structure, it often demands a probabilistic approach for the treatment of historical streamflow observations within the catchments region based on several mathematical and statistical frameworks. The flood frequency analysis or FFA statistically defines an inter-association between flood design quantiles and their recursion interval by fitting a univariate or multivariate probability distribution functions or pdfs. Multivariate flood distribution analysis often provides a comprehensive understanding in the flood generating probability which usually comprises a combination of the joint probability density functions or pdfs and joint cumulative distribution functions or cdfs. In actuality, flood is a multivariate complex and stochastic hydrologic consequence usually characterized completely through its multiple intercorrelated random vectors, i.e. flood peak discharge, volume and duration of flood hydrograph. Therefore, the reliability of univariate flood frequency analysis or return periods often stands for several queries which would be

attributed for underestimation and overestimations of hydrologic risk and thus often demanding the establishment of multivariate joint distribution of flood characteristics by accounting its multiple intercorrelated flood vectors. Actually, univariate flood probability constructions would be incapable to recognize the full screen of flood or inflow hydrograph and reduce uncertainty in the estimated design quantiles.

In this literature, the efficacy of copula-based methodologies is reviewed for establishing multivariate distributions of flood episodes, which is recognized as highly flexible tool for establishing multivariate joint dependency and their associated return periods in comparison with traditional multivariate functions that are discussed in the context of a theoretical and mathematical simulation for the flood characteristics. In this study, different methodological attempts in the light of bivariate and trivariate copula distribution analysis are pointed for tackling multivariate design problems and estimating design variable quantiles under different notations of return periods. The section 'Flood frequency analysis via one-dimensional probability distribution framework or approximation of marginal distributions' pointed a distinguish variety of one-dimensional mono-parametric, bi-parametric or tri-parametric based parametric distribution family functions, which are often employed for establishing univariate marginal distributions and which often a mandatory prerequisite desires before introducing individual hydrologic or flood vectors into multivariate or copula framework. It is also revealed that the different density structures attribute different estimations of design quantiles, especially in the tail of the distributions; also, flexibility of available univariate functions exhibited a control to justify an appropriate fit with the given samples such that it depends upon its associated vectors of unknown statistical or model parameters. But, simulations via parametric functions often imposed an assumption that random samples are drawn from the population whose density structure is pre-defined. In actuality, no specific models are categorized and opted universally for any specific hydrologic variables, which would follow different distributions. Therefore, flexibility of non-parametric based kernel density estimator is recognized as a much stable data smoothing procedure in the field of hydrologic or flood modelling and which yielding a bonafide density as revealed from section 'An approach via non-parametric distribution framework'. The non-parametric framework does not require any prior distribution assumptions and will be directly derived from distribution series with higher extent of flexibility as compared with parametric density estimators. Unless the univariate frequency analysis defines the general concept of return period via the estimated cumulative distribution function, it might be unsatisfactory when the requirement demands the consideration of

multivariate design parameters, which often reveals an essential concern in the water-related queries.

Multivariate practices via the traditional probability distribution functions often attribute for the several statistical shortcomings and limitations, as revealed from section 'Limitation of traditional multivariate distribution framework'. Actually, the classical statistical approaches of estimating degree of association would be incapable for characterizing the comovements tendencies of hydrologic or flood vectors. In such aspects, the copula function appeared as a most effective multivariate tool which segregates modelling of individual univariate vectors and their joint structure separately into two distinct stages, which thus attribute higher flexibility in selecting most appropriate and justifiable marginal distributions and their joint structure to capture a wider extent of dependency along with preservation in their joint structure, as revealed from section 'Copula-based bivariate probability distributions'. The copula-based methodology can be classified as parametric, semiparametric and non-parametric estimation procedures depending upon the way of estimating its univariate marginals and joint dependence structure. An interactive set of copula family function such as the extreme value class (i.e. Gumbel-Hougaard, Galambos and Husler-Reiss), elliptical class (i.e. Gaussian family), unclassified Plackett and Farlie-Gumbel-Morgenstern (or FGM) parametric functions and three-parametric Twan family (i.e. belong to extreme value class) is often incorporated for establishing bivariate joint dependence structure, in which the Archimedean class (i.e. Ali-Mikhail or A-M-H family, Frank family, Clayton or Cook-Johnson (C-J) family and Gumbel-Hougaard family) copulas are frequently accepted due to large varieties of families and its capability to capture joint dependencies for a wider extent also, exhibiting several desirable properties which attributes much flexibility during joint probability simulations, as revealed from same section. Unless an extended efforts are often motivated towards establishing copula-based bivariate simulations and estimation of bivariate design variable quantiles under the different notations of return periods but, such attempts still might be insufficient for revealing a justifiable and comprehensive studies of flood probability analysis, due to its trivariate behaviour. Actually, potential damage could likely be a function of multiple relevant vectors of specified hydrological episodes such that an ignorance of spatial dependency among these uncertain vectors might be attributed for the underestimation of uncertainty, which frequently encountered during risk evaluation. Therefore, section 'Trivariate joint dependency constructions via 3-dimensional copulas' discussed the applicability of 3-dimensional copula function for establishing trivariate joint simulation of flood characteristics and their associated return periods but whose computational strategies are quite limited over the existing literature. The above conventional trivariate copula simulation often encountered some statistical issues such as complexity during

the approximation of justifiable parametric distributions for higher dimensional hydrological attributes and also might be quite ineffective to capture and reflect all the possible mutual concurrency among multidimensional flood vectors, as revealed from section 'Vine copulas or PCC framework for trivariate joint distributions'. Due to higher degree of uncertainty and complex flood dependence structure, resolving the dependence structure of multivariate extreme via conventional copula formulation is quite complex and often demands a flexible methodology through precise estimation of tail dependence coefficient under various tail dependency. For solving such issues, the vine or pair-copula construction (PCC) provides a comprehensive way of uncertainty characterization for the higher dimensional hydrological entities which is solely based on the principle of the decomposition of full multivariate density into a cascade or simple local building blocks via conditional independence or pair-copulas. In actuality, due to conditional mixing via the stage-wise hierarchical nesting procedure, the pair-copula concept exhibited much effective and flexible modelling environments. The PCC structure also exhibited some modelling issues as revealed from 'Vine copulas or PCC framework for trivariate joint distributions' and thus motivated towards the minimum information PCC which capture complex multidimensional flood structure for various tail dependencies by the precise estimation of their tail coefficient for given selected copulas and also facilitate to the model multivariate extreme in the presence of limited data length.

The statistical significance of return periods under multidimensional design concept for tackling different hydrologic problems is discussed in the section 'Return periods under multivariate settings'. Return periods can be derived from the exceedance probabilities of flood attributes pair in the multidimensional risk framework, such that joint return period retrieves from the joint exceedance probabilities and segregated into two distinct groups, i.e. primary return period comprise via the inclusive probability such as 'AND and 'OR' return period and the secondary or 'Kendall' return period, which can be define based on the Kendall's probability distribution or survival function. Utilizing the standard definition of return period based on inclusion probability or primary returns might be attributed for underestimation of correct value. Actually, primary return periods captured the mean forecasting which would be incapable to demonstrate the risk of supercritical or dangerous scenario. Appropriate reliability of hydraulic design system often intended towards the definition of exceedance probabilities for rare episodes and thus pointed the mathematical significance and derivation of the secondary return periods derived from the Kendall's distribution and survival function called the Kendall's return period, as discussed in section 'Demonstrating the risk of supercritical extreme via the Kendall's distribution and survival functions (or secondary return periods)'.

Some ideas to strengthen the current attempts of multivariate practices via incorporating time-varying copula framework

Unless extended efforts are motivated via the multivariate stochastic generations of flood characteristics in the context of bivariate or trivariate copulas simulations for retrieving the flood exceedance probabilities or design quantiles under different notations of return periods, how one could appropriately justify the desire of defensive task without addressing dynamic environmental arising (Climate change and/or LULCC) often poses an isolation behaviour or independence with such phenomenon (Kartz et al. 2002; Strupczewski and Kaczmarski 2001; Khaliq et al. 2006; El Adlouni et al. 2007; Villarini et al. 2009; Wigley 2009; Lopez-Paz et al. 2013; Condon et al. 2015). Consistencies and accuracy in the estimated design quantiles under stationary risk framework might be doubted due to the ignorance of the accountability of changing phenomenon either over univariate structure (i.e. temporal variability in their mean and variance) of individuals flood vectors or in their joint correlation structure (Zhang 2005; Bender et al. 2014; Galiatsatou and Prinos 2016; Sarhadi et al. 2016). The existence of the non-stationarity, due to external controlling factors, usually tries to interrupt the hydrological behaviour within catchment region and which might be further altered the expectation of such extremity happening under time-invariant hydrologic risk assessment efforts. Actually, traditional flood modelling is often designed with the hypothesis of independent and identically distributed (or i.i.d) behaviour of hydrologic samples and such assumption often adapted as a standard design procedure for tackling water-related issues but, due to the time-varying consequences, it would interrupt the statistical characteristics of hydrological samples and might lead to non-stationarity (Gilroy and Macuen 2012; Lima et al. 2015). Time-varying controlling covariates would pose their stress over the hydrological characteristics differently from the future prospects as compared to what their impact surrounded in the past or present scenario (Khaliq et al. 2006; Chebana et al. 2013; Jiang et al. 2015), such that the chance of the future occurrence of flood episodes would be likely changed over the time functions such as like goodness of 100-year flood in a given year would change (Gilroy and Macuen 2012; Du et al. 2015). Thus, stress of the time-varying consequences over flood exceedance probabilities might resulted that the actual associated 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 strategies (Lima et al. 2015). More especially from the prospect of engineering based hydraulic structural designing procedure where, it could be an essential desire to outline the expectation of potential for future changing over their design value to justify the anticipated structural design life

appropriately both from the present and from the future prospects (Bender et al. 2014; Sarhadi et al. 2016). Numerous efforts often motivated to addressed the dynamic consequences over univariate hydrological characteristics through the implementation of univariate extreme value modelling frame via covariate analysis (i.e. Strupczewski and Kaczmarski 2001; Coles 2001; Kartz et al. 2002; Zhang 2005; Wong et al. 2006; Clarke 2007; El Adlouni et al. 2007; Villarini et al. 2010; Gilroy and Macuen 2012; Lopez and Frances 2013; Lima et al. 2015). But the multidimensional behaviour of flood episodes often demands a multivariate stochastic framework in conjunction with addressing of dynamic consequences over their design quantile framework. In the traditional copula-based methodology, both the marginal distribution and the copula-based joint dependence parameters will not allow varying over time in order to adjoin the stress of covariates over the flood characteristics (Corbella and Stretch 2013; Jiang et al. 2015; Galiatsatou and Prinos 2016). With my best knowledge, very limited attempts adapted the multivariate modelling strategies for hydrological characteristics in the light of dynamic copula frameworks, i.e. Corbella and Stretch (2013), Bender et al. (2014), Jiang et al. (2015), Sarhadi et al. (2016), and Galiatsatou and Prinos (2016). These computational strategies are usually segregated into two distinct stages such as modelling the non-stationarity consequences over the univariate flood attributes through time-varying marginal distribution structure while modelling the dynamic scenario over the joint probability structure of multiple random vectors in conjunction with copulas structure with dynamic parameters under time-varying joint frame structure. As concluded from above literature that dynamic copula simulation often incorporated via demonstrating the temporal variation within bivariate joint relationship and their return periods, also the importance of trivariate joint distributions and their return periods is already pointed in the above section (i.e. Graler et al. 2013; Reddy and Ganguli 2013 and references therein). On another side, the flexibility of PCC and minimum information PCC structure for constructing the higher dimensional copulas are already pointed from the above cited literature (i.e. Daneshkhan et al. 2015, 2016). Thus, the overall conclusion could point the ideas towards an attempt via integrating dynamic concept for capturing the effect of temporal influence in the trivariate joint distribution or design quantiles.

References

Aas K, Berg D (2009) Models for construction of multivariate dependence- a comparison study. Eur J Financ 15:639–659. https://doi.org/10.1080/13518470802588767

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

Adamowski K (1989) A Monte Carlo comparison of parametric and nonparametric estimation of flood frequencies. J Hydrol 108:295–308. https://doi.org/10.1016/0022-1694(89)90290-4

Adamowski K (1985) Nonparametric kernel estimation of flood frequencies. Water Resour Res 21(11):1885–1890. https://doi.org/10.1029/WR021i011p01585

Adamowski K (1996) Nonparametric estimations of low-flow frequencies. J Hydraul Eng 122(1):46–49. https://doi.org/10.1061/(ASCE)0733-9429(1996)122:1(46)

Adamowski K (2000) Regional analysis of annual maximum and partial duration flood data by nonparametric and L-moment methods. J Hydrol 229(3):219–231. https://doi.org/10.1016/S0022-1694(00)00156-6

Adamowski K, Feluch W (1990) Nonparametric flood-frequency analysis with historical information. J Hydraul Engg ASCE 116(8):1035–1047. https://doi.org/10.1061/(ASCE)0733-9429(1990)116:8(1035)

Adamowski K, Labatiuk C (1987) Estimation of flood frequencies by a nonparametric density procedure. Hydrol Frequency Modeling:97–106. https://doi.org/10.1007/978-94-009-3953-0_5

Ane T, Kharoubi C (2003) Dependence structure and risk measure. J Bus Res 76(3):411–438. https://doi.org/10.1086/375253

Bacchi B, Becciu G, Kottegoda NT (1994) Bivariate exponential model applied to intensities and durations of extreme rainfall. J Hydrol 155(1/2):255–236. https://doi.org/10.1016/0022-1694(94)90166-X

Badrzadeh H, Sarukkalige R, Jayawardena A (2013) Impact of multiresolution analysis of artificial intelligence models inputs on multistep ahead river flow forecasting. J Hydrol 507:75–85. https://doi.org/10.1016/j.jhydrol.2013.10.017

Bardsley WE (1988) Towards a general procedure for analysis of extreme random events in the earth sciences. Math Geol 20(5):513–528. https://doi.org/10.1007/BF00890334

Bardsley WE, Manly BFJ (1987) Transformations for improved convergence of distributions of flood maxima to a Gumbel limit. J Hydrol 91:137–152. https://doi.org/10.1016/0022-1694(87)90133-8

Bartlett MS (1963) Statistical estimation of density functions, Sankhya. Indian Journal of Statistics Series A 25(3):245–254

Bedford T, Cook RM (2002) Vines-a new graphical model for dependent random variables. Ann. Stat 30(4):1031–1068. https://doi.org/10.1214/aos/1031689016

Bedford T, Cooke R (2001) Probability density decomposition for conditional dependent random variables modelled by vines. Annal of Mathematics and Artificial Intelligence 32:245–268. https://doi.org/10.1023/A:1016725902970

Bedford T, Daneshkhan A, Wilson KJ (2015) Approximate uncertainty modelling in risk analysis with vine copulas. Risk Anal. https://doi.org/10.1111/risa.12471

Bedford T, Meeuwissen A (1997) Minimally informative distributions with given rank correlation for use in uncertainty analysis. J Stat Comput Simulat 57, 143–174.

Bender J, Wahl T, Jensen J (2014) Multivariate design in the presence of non-stationary. J Hydrol 514:123–130. https://doi.org/10.1016/j.jhydrol.2014.04.017

Blazkova S, Beven K (2004) Flood frequency estimation by continuous simulation of subcatchments rainfalls and discharges with the aim of improving dam safety assessments in a large basin in the Czech Republic. J Hydrol 292:153–172. https://doi.org/10.1016/j.jhydrol.2003.12.025

Bobee B (1975) The log Pearson type 3 distribution and its application in hydrology. Water Resour Res 11(5):681–689. https://doi.org/10.1029/WR011i005p00681

Bobee B, Ashkar F (1989) Log-logistic flood frequency analysis-comment. J Hydrol 107:367–372. https://doi.org/10.1016/0022-1694(89)90067-X

Bobee B, Rasmussen PF (1994) Statistical analysis of annual flood series. In: Menon J (ed) Trend in hydrology, 1. Council of Scientific Research Integration, Anusandhan Bhawan, pp 117–135

Botev ZI, Grotowski JF, Kroese DP (2010) Kernel density estimation via diffusion. Ann Stat 38(5):2916–2957. https://doi.org/10.1214/10-AOS799

Boughton W, Srikanthan S, Weinmann E (2002) Benchmarking a new design flood estimation system. Aust I Water Resour 6(1):45–52. https://doi.org/10.1080/13241583.2002.11465209

Bowman AW (1984) An alternative method of cross-validation for the smoothing of density estimates. Biometrika 71:353–360. https://doi.org/10.2307/2336252

Bowman A, Azzalini A (1997) Applied smoothing techniques for data analysis: the kernel approach with S-plus illustrations. Oxford University Press, New York

Bras RL (1990) Hydrology: an introduction to hydrologic science. Addison-Wesley ISBN 0201059223:9780201059229

Brunner MI, Favre A, Seibert J (2016) Bivariate return periods and their importance for flood peak and volume estimations. Wiley Interdisciplinary Reviews: Water 3(6):819–833. https://doi.org/10.1002/wat2.1173

Burn DH (1990) Evaluation of regional flood frequency analysis with a region of influence approach. Water Resour Res 26(10):2257–2265. https://doi.org/10.1029/WR026i010p02257

Calver A, Lamb R (1995) Flood frequency estimation using continuous rainfall-runoff modelling. Physic and Chemistry of the Earth 20: 479–483. https://doi.org/10.1016/S0079-1946(96)00010-9

Chebana F, Ouarda TBMJ (2009) Multivariate quantiles in hydrological frequency analysis. Environmetrics 22:63–78. https://doi.org/10.1002/env.1027

Chebana F, Ouarda TBMJ, Duong TC (2013) Testing for multivariate trends in hydrological frequency analysis. J Hydrol 486:519–530. https://doi.org/10.1016/j.jhydrol.2013.01.007

Choros B, Ibragimov R, Permiakova E (2010) Copula estimation. In: Jaworski P et al (eds) Copula theory and its applications-Proceeding of the Workshop Held in Warsaw 25-26 September 2009. Springer, NewYork, pp 77–91

Choubin B, Khalighi-Sigaroodi S, Malekian A, Ahmad S, Attarod P (2014) Drought forecasting in a semi-arid watershed using climate signals: a neuro-fuzzy modeling approach. J Mt Sci 11:1593–1605. https://doi.org/10.1007/s11629-014-3020-6

Choulakian V, Jabi EIN, Issa M (1990) On the distribution of flood volume in partial duration series analyses of flood phenomenon. Stoch Hydrol and Hydraul 4(3):217–226. https://doi.org/10.1007/BF01543085

Chow VT, Maidment DR, Mays LW (1988) Applied hydrology. McGraw Hill, New York

Clarke R (2007) Hydrological prediction in a non-stationary world. Hydrol Earth Syst Sci 11:408–414. https://doi.org/10.5194/hess-11-408-2007

Coles S (2001) An introduction statistical modelling of extreme values. Springer, London

Condon LE, Gangopadhyay S, Pruitt T (2015) Climate change and non-stationary flood risk for the upper Truckee River basin. Hydrol Earth Syst Sci 19:159–175. https://doi.org/10.5194/hess-19-159-2015

Cong RG, Brady M (2011) The interdependence between rainfall and temperature: copula analyses. The Scientific World Journal: 405675. https://doi.org/10.1100/2012/405675

Corbella A, Stretch DD (2013) Simulating a multivariate sea storm using Archimedean copulas. Coast Eng 97:37–52. https://doi.org/10.1016/j.coastaleng.2013.01.011

Cuadras CM (1992) Probability distributions with given multivariate marginal and given dependence structure. J Multivar Anal 42:51–66. https://doi.org/10.1016/0047-259X(92)90078-T

Cunnane C (1987) Review of statistical models for flood frequency estimation. In: Singh VP (ed) Hydrological frequency modelling. Reidal, Dordrecht, pp 49–95

Cunnane C (1988) Methods and merits of regional flood frequency analysis. J Hydrol 100:269–290. https://doi.org/10.1016/0022-1694(88)90188-6

Cunnane C (1989) Statistical distributions for flood frequency analysis. Operational Hydrology Report no. 33, WMO no. 718. World Meteorological Organization, Geneva https://library.wmo.int/doc_num.php?explnum_id=1695

Czado C, Min A (2010) Bayesian inference for D-vines: estimation and model selection. In: Dependence Modelling. World Scientific, Singapore. https://doi.org/10.1142/9789814299886_0012

Czado C, Jeske S, Hofmann M (2013) Selection strategies for regular vine copulae. Journal of the French Society of Statistics 154(1): 174–190

Daneshkhan A, Parham GA, Chatrabgoun O, Jokar M (2015) Approximation multivariate distribution with pair copula using the orthonormal polynomial and Legendre multiwavelets basis function. Communications in Statistics – Simulation and Computation. https://doi.org/10.1080/03610918.2013.804557

Daneshkhan A, Remesan R, Omid C, Holman IP (2016) Probabilistic modelling of flood characteristics with parametric and minimum information pair-copula model. J Hydrol 540:469–487. https://doi.org/10.1016/j.jhydrol.2016.06.044

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

De Michele C, Salvadori G, Canossi M, Petaccia A, Rosso R (2005) Bivariate statistical approach to check the adequacy of dam spillway. J Hydrol Eng 10(1):50–57. https://doi.org/10.1061/(ASCE)1084-0699(2005)10:1(50)

Dehghani M, Saghafian B, Rivaz F, Khodadadi A (2017) Evaluation of dynamic regression and artificial neural networks models for real-time hydrological drought forecasting. Arab J Geosci 10:266. https://doi.org/10.1007/s12517-017-2990-4

Dooge JCE (1986) Looking for hydrologic laws. Water Resour Res 22(9):465–485. https://doi.org/10.1029/WR022i09Sp0046S

Du T, Xiong L, Xu CY, Gippel CJ, Gou S, Liu P (2015) Return period and risk analysis of non-stationary low-flow series under climate change. J Hydrol 527:234–250. https://doi.org/10.1016/2Fj.jhydrol.2015.04.041

Duins RPW (1976) On the choice of smoothing parameters of Parzen estimator of probability density functions. IEEE Trans Comput C-25:1175–1179. https://doi.org/10.1109/TC.1976.1674577

Duong T, Hazelton ML (2003) Plug-in bandwidth selectors for bivariate kernel density estimations. J Nonparametr Stat 15:17–30. https://doi.org/10.1080/10485250306039

Dupuis DJ (2007) Using copulas in hydrology: benefits, cautions, and issues. J Hydrol Eng 12(4):381–393. https://doi.org/10.1061/(ASCE)1084-0699(2007)12:4(381)

Durrans SR (1992) Parameter estimation for the Pearson type 3 distribution using order statistics. J Hydrol 133(3–4):215–232. https://doi.org/10.1016/0022-1694(92)90256-U

Durrans SR, Eiffe MA, Thomas WO, Goranflo HM (2003) Joint Seasonal/Annual Flood Frequency Analysis. Journal of Hydrologic Engineering 8 (4):181–189

Efromovich S (1999) Nonparametric curve estimation: methods, theory and applications. Springer-Verlag, New York

El Adlouni S, Ouarda TBMJ, Zhang X, Roy R, Bobee B (2007) Generalized maximum likehood estimators for the nonstationary generalized extreme model. Water Resour Res 43:W03410. https://doi.org/10.1029/2005WR004545

Escalante C (2007) Application of bivariate extreme value distribution to flood frequency analysis: a case study of North-Western Mexico.

Nat Hazards 42(1):37–46. https://doi.org/10.1007/s11069-006-9044-7

Etemad-Shahidi A, Mahjoobi J (2009) Comparison between m50 model tree and neural networks for prediction of significant wave height in lake superior. Ocean Eng 36:1175–1181. https://doi.org/10.1016/j.oceaneng.2009.08.008

Fan L, Zheng Q (2016) Probabilistic modelling of flood events using the entropy copula. Adv Water Resour 97:233–240. https://doi.org/10.1016/2Fj.advwatres.2016.09.016

Fan YR, Huang WW, Huang GH, Huang K, Li YP, Kong XM (2015) Bivariate hydrological risk analysis based on coupled entropy-copula method for the Xiang Xi River in the Three Gorges Reservoir area. Theoretical and Applied Climatology, China. https://doi.org/10.1007/s00704-015-1505-z

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

Galiatsatou P, Prinos P (2016) Joint probability analysis of extreme wave heights and storm surges in the Aegean Sea in a changing climate. FLOOD risk 2016- 3rd European conference on Flood Risk Management, E3S Web of Conferences 7, 02002

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

Ghanbarpour MR, Abbaspour KC, Jalalvand G, Ashtiani Moghaddam G (2010) Stochastic modeling of surface stream flow at different time scales: Sangsoorakh Karst Basin, Iran. J Cave Karst Stud 72(1):1–10 http://caves.org/pub/journal/PDF/v72/cave-72-01-1.pdf

Ghosh S, Mujumdar PP (2007) Nonparametric methods for modeling GCM and scenario uncertainty in drought assessments. Water Resour Res 43:W07405. https://doi.org/10.1029/2006WR005351

Ghoudi K, Khoudraji A, Rivest LP (1998) Proprietes statistiques des copules de valeurs extremes bidimensionelles. Canadian Journal of Statistics 26:187–197 https://pdfs.semanticscholar.org/a700/c67b8ca24e13130912afde9fcb56ea79d952.pdf

Gilroy KL, Macuen RH (2012) A non-stationary flood frequency analysis method to adjust for future climate change and urbanization. J Hydrol 414-415:40–48. https://doi.org/10.1016/j.jhydrol.2011.10.009

Gizaw, M.S. Gan, T.Y (2016) Regional flood frequency analysis using support vector regression under historical and future climate. J Hydrol 538, 387–398. DOI: https://doi.org/10.1080/02626667.2018.1432056

Goel NK, Seth SM, Chandra S (1998) Multivariate modelling of flood flows. J Hydraul Eng 124(2):146–155. https://doi.org/10.1061/(ASCE)0733-9429(1998)124:2(146)

Gong Y, Zhang Y, Lan S, Wang H (2016) A comparative study of artificial neural networks, support vector machines and adaptive neuro fuzzy inference system for forecasting groundwater levels near Lake Okeechobee, Florida. Water Resour Manag 30:375–391. https://doi.org/10.1007/s11269-015-1167-8

Graler B, Berg MJV, Vandenberg S, Petroselli A, Grimaldi S, Baets BD, Verhost NEC (2013) Multivariate return periods in hydrology: a critical and practical review focusing on synthetic design hydrograph estimation. Hydrol Earth Sys Sci 17:1281–1296. https://doi.org/10.5194/hess-17-1281-2013

Grimaldi S, Serinaldi F (2006) Asymmetric copula in multivariate flood frequency analysis. Adv Water Resour 29:1155–1167. https://doi.org/10.1016/j.advwatres.2005.09.005

Grimaldi S, Baets BD, Verhost NEC (2013) Multivariate return periods in hydrology: a critical and practical review focusing on synthetic design hydrograph estimation. Hydrol Earth Sys Sci

Guimarães Santos CA, da Silva GBL (2014) Daily streamflow forecasting using a wavelet transform and artificial neural network hybrid

models. Hydrol Sci J 59:312–324. https://doi.org/10.1080/02626667.2013.800944

Gyasi-Agyei Y, Melching C (2012) Modelling the dependence and internal storm structure of storm evens for continuous rainfall simulation. J Hydrol 464:249–261. https://doi.org/10.1016/j.jhydrol.2012.07.014

Hobaek HI, Asa K, Frigessi A (2010) On the simplified pair-copula construction- simply useful or too simplistic? J Multivar Anal 101:1296–1310. https://doi.org/10.1016/j.jmva.2009.12.001

Hosking JMR, Wallis JR (1997) Regional frequency analysis. Cambridge University Press, Cambridge

Hosking JRM, Wallis JR, Wood EF (1985) Estimation of the general extreme value distribution by the method of probability weighted moments. Technometrics 27(3):251–261. https://doi.org/10.1080/00401706.1985.10488049

Huang YF, Mirzaei M, Yap WK (2016) Flood analysis in Langat River Basin using stochastic model. International Journal of GEOMATE 11(27):2796–2803. https://doi.org/10.21660/2016.27.1143

Jain A, Prasad Indurthy S (2004) Closure to "comparative analysis of event-based rainfall-runoff modeling techniques—deterministic, statistical, and artificial neural networks" by ASHU JAIN and SKV prasad indurthy. J Hydrol Eng 9:551–553. https://doi.org/10.1061/(ASCE)1084-0699(2004)9:6(551)

Jajarmizadeh M, Lafdani EK, Harun S, Ahmadi A (2015) Application of SVM and swat models for monthly streamflow prediction, a case study in South of Iran. KSCE J Civ Eng 19:345–357

Jenkinson AF (1955) The frequency distribution of the annual maximum (or minimum) values of meteorological elements. Q J Roy Meteor Soc 87:158–171. https://doi.org/10.1002/qj.49708134804

Jiang C, Xiong L, Xu CY, Guo S (2015) Bivariate frequency analysis of non-stationary low-flow series based on the time-varying copula. Hydrol Process 29:1521–1534. https://doi.org/10.1002/hyp.10288

Joe H (1997) Multivariate models and dependence concept. CRC Press, Boca Raton, Fla

Johnson NL (1994) Continuous univariate distribution, Wiley New York, Vol 1

Jones MC, PJ Foster (1996) A simple nonnegative boundary correction method for kernel density estimation, stat. Sin 6: 1005–1013

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

Karmakar S, Simonovic SP (2008) Bivariate flood frequency analysis. Part-1: determination of marginal by parametric and non-parametric techniques. J Flood Risk Manag 1:190–200. https://doi.org/10.1111/j.1753-318X.2008.00022.x

Karmakar S, Simonovic SP (2009) Bivariate flood frequency analysis. Part-2: a copula-based approach with mixed marginal distributions. J Flood Risk Manag 2(1):1–13. https://doi.org/10.1111/j.1753-318X.2009.01020.x

Kartz RW, Parlang MB, Naveau P (2002) Statistics of extremes in hydrology. Adv Water Resour 25(8):1287–1304. https://doi.org/10.1016/S0309-1708(02)00056-8

Khaliq M, Ouarda T, Ondo J-C, Gachon P, Bobee B (2006) Frequency analysis of a sequence of dependent and/or non-stationary hydro-meteorological observations: a review. J Hydrol 329(3–4):534–552. https://doi.org/10.1016/j.jhydrol.2006.03.004

Kim KD, Heo JH (2002) Comparative study of flood quantiles estimation by nonparametric models. J Hydrol 260:176–193. https://doi.org/10.1016/S0022-1694(01)00613-8

Kim TW, Valdes JB, Yoo C (2003) Nonparametric approach for estimating return periods of droughts in arid regions. J Hydrol Engg ASCE 8(5):237–246. https://doi.org/10.1061/(ASCE)1084-0699(2003)8:5(237)

Kisi O (2007) Streamflow forecasting using different artificial neural network algorithms. J Hydrol Eng 12:532–539 https://doi.org/10.1061/(ASCE)1084-0699(2007)12:5(532)

Krstanovic PF, Singh VP (1987) A multivariate stochastic flood analysis using entropy. In: Singh VP (ed) Hydrologic frequency modelling. Reidel, Dordrecht, pp 515–539

Kurowicka D, Cooke R (2006) Uncertainty analysis with high dimensional dependence modelling. John Wiley

Kyselý J, Gaál L, Picek J (2011) Comparison of regional and at site approaches to modelling probabilities of heavy precipitation. Int J Climatol 31:1457–1472. https://doi.org/10.1002/joc.2182

Lafdani EK, Nia AM, Pahlavanravi A, Ahmadi A, Jajarmizadeh M (2013) Research article daily rainfall-runoff prediction and simulation using ANN, ANFIS and conceptual hydrological MIKE11/NAM models. Int J Eng Technol 1:32–50

Lall U (1995) Recent advances in nonparametric function estimation: hydrological applications. Rev Geophys 33(S1):1093–1102. https://doi.org/10.1029/95RG00343

Lall U, Moon Y-I, Khalil AF (1993) Kernel flood frequency estimators: bandwidth selection and kernel choice. Water Resour Res 29(4):1003–1015. https://doi.org/10.1029/92WR02466

Lall U, Rajagopalan B, Tarboton DG (1996) A nonparametric wet/dry spell model for resampling daily precipitation. Water Resour Res 32(9):2803–2823. https://doi.org/10.1029/96WR00565

Lawrence D, Paquet E, Gailhard J, Fleig AK (2014) Stochastic semi-continuous simulations for extreme flood estimations in catchments with combined rainfall-snowmelt flood regimes. Natural Hazard and Earth System Sciences 14:1283–1298. https://doi.org/10.5194/nhess-14-1283-2014

Li L, Xu H, Chen X, Simonovic S (2010) Streamflow forecast and reservoir operation performance assessment under climate change. Water Resour Manag 24:83. https://doi.org/10.1007/s11269-009-9438-x

Li S, Ma K, Jin Z, Zhu Y (2016) A new flood forecasting model based on SVM and boosting learning algorithms. In Proceedings of the 2016 IEEE Congress on Evolutionary Computation (CEC), Vancouver, BC, Canada, 24–29 July 2016; pp. 1343–1348

Liaw A, Wiener M (2002) Classification and regression by random forest. R News, 2, 18–22. ISSN 115-3631

Lima CHR, Upmanu L, Troy TJ, Denieni N (2015) A climate informed model for nonstationary flood risk prediction: application to Negro River at Manaus, Amazonia. J Hydrol 522:594–602. https://doi.org/10.1016/2Fj.jhydrol.2015.01.009

Lopez J, Frances F (2013) Non-stationary flood frequency analysis in continental Spanish River, using climate and reservoir indices as an external covariate. Hydrol Earth Sys Sci 17:3189–3203. https://doi.org/10.5194/hess-17-3189-2013

Lopez-Paz D, Hernandez-Lobarto JM, Ghahramani Z (2013) Gaussian process vine copulas for multivariate dependence. Proceedings of the 30th International Conferences on Machine Learning, Vol 28, 10–18

Ma M, Song S, Ren L, Jiang S, Song J (2013) Multivariate drought characteristics using trivariate Gaussian and student copula. Hydrol Process 27:1175–1190. https://doi.org/10.1002/hyp.8432

Machekposhti KM, Sedghi H, Telvari A, Babazadeh H (2017) Flood analysis in Karkheh River basin using stochastic model. Civil Eng J 3(9):794–808. https://doi.org/10.21859/cej-030915

Madadgar S, Moradkhani H (2013) Drought analysis under climate change using copula. J Hydrol Eng 18:746–759. https://doi.org/10.1061/(ASCE)HE.1943-5584.0000532

Mahyat Shafapour Tehrany, Moung-Jin Lee, Biswajeet Pradhan, Mustafa Neamah Jebur, Saro Lee (2014) Flood susceptibility mapping using integrated bivariate and multivariate statistical models. Environmental Earth Sciences 72 (10):4001–4015

Maity R, Kumar DN (2008) Probabilistic prediction of hydroclimate variables with non-parametric quantification of uncertainty. J Geophys Res 113:D14105. https://doi.org/10.1029/2008JD009856

Mirabbasi R, Kakheri-Fard A, Dinpashoh Y (2012) Bivariate drought frequency analysis using the copula method. Theor Appl Climatol 108:191–206. https://doi.org/10.1007/s00704-011-0524-7

Moon Y-I, Lall U (1994) Kernel quantiles functions estimators for flood frequency analysis. Water Resour Res 30(11):3095–3103 https://digitalcommons.usu.edu/water_rep/194

Mosavi A, Ozturk P, Chau K-W (2018) Flood prediction using machine learning models: literature review. Water 10:1536. https://doi.org/10.3390/w10111536

Nadarajah S, Shiau J (2005) Analysis of extreme flood events for the Pachang River, Taiwan. Water Resour Manag 19:363–375. https://doi.org/10.1007/s11269-005-2073-2

Nelsen R.B. (2006). An introduction to copulas, Springer, New York

Nikouloulopoulos A, Joe H, Li H (2012) Vine copula with asymmetric tail dependence and applications to financial return data. Computational Statistics and Data Analysis 56:3659–3673. https://doi.org/10.1016/j.csda.2010.07.016

O'Connel PE (1977) ARIMA model in synthetic hydrology: mathematical model for surface hydrology. John Wiley, NY

Ortiz-García E, Salcedo-Sanz S, Casanova-Mateo C (2014) Accurate precipitation prediction with support vector classifiers: a study including novel predictive variables and observational data. Atmos Res 139:128–136

Ouarda TBMJ, Girard C, Cavadias G, Bobee B (2001) Regional flood frequency estimation with canonical correlation analysis. J Hydrol 254:157–173. https://doi.org/10.1016/S0022-1694(01)00488-7

Papaioannou G, Kohnova S, Bacigal T, Szolgay J, Hlavcova K, Loukas A (2016) Joint modelling of flood peaks and volumes: a copula application for the Danube River. J Hydrol Hydromech 64(4):382–392

Parzen E (1962) On the estimation of a probability density function and mode. The Annals of Mathematical Statistics 33:1065–1076. https://doi.org/10.1214/aoms/1177704472

Poulin A, Huard D, Favre AC, Pugin S (2007) Importance of tail dependence in bivariate frequency analysis. J Hydrol Eng 12(4):394–403. https://doi.org/10.1061/(ASCE)1084-0699(2007)12:4(394)

Rao DV (1980) Log Pearson type 3 distribution: a generalized evaluation. J Hydraul Division ASCE 106(5):853–872

Rao AR, Hameed KH (2000) Flood frequency analysis. CRC Press, Boca Raton, Fla

Rauf UFA, Zeephongsekul P (2014) Copula based analysis of rainfall severity and duration: a case study. Theor Appl Climatol 115(1–2):153–166. https://doi.org/10.1007/s00704-013-0877-1

Ravansalar M, Rajaee T, Kisi O (2017) Wavelet-linear genetic programming: a new approach for modeling monthly streamflow. J Hydrol 549:461–475. https://doi.org/10.1016/j.jhydrol.2017.04.018

Ravinesh C. Deo, Mehmet Şahin, (2015) Application of the Artificial Neural Network model for prediction of monthly Standardized Precipitation and Evapotranspiration Index using hydrometeorological parameters and climate indices in eastern Australia. Atmospheric Research 161-162:65–81

Raynal-Villasenor JA, Salas JD (1987) Multivariate extreme value distributions in hydrological analyses. In: Rodda JC, Matalas NC (eds) Water for the future: hydrology in perspective. IAHS Publication No. 164. IAHS, Wallingford, pp 111–119

Reddy MJ, Ganguli P (2012a) Bivariate flood frequency analysis of Upper Godavari River flows using Archimedean copulas. Water Resour Manage: DOI. https://doi.org/10.1007/s11269-012-0124-z

Reddy MJ, Ganguli P (2012b) Risk assessments of hydro-climatic variability on ground water levels in the Manjra basin aquifer in India using Archimedean copulas. J Hydrol Engg. https://doi.org/10.1061/(ASCE)HE.1943-5584.0000564

Reddy MJ, Ganguli P (2013) Probabilistic assessments of flood risks using trivariate copulas. Theor Appl Climatol 111:341–360. https://doi.org/10.1007/s00704-012-0664-4

Renard B, Lang M (2007) Use of a Gaussian copula for multivariate extreme value analysis: some case studies in hydrology. Adv

Water Resour 30:897–912. https://doi.org/10.1016/j.advwatres.2006.08.001

Requena A, Flores I, Mediero L, Garrote L (2016) Extension of observed flood series by combining a distributed hydro-meteorological model and a copula based model. Stoch Environ Res Risk Assess 30(5). https://doi.org/10.1007/s00477-015-1138-x

Rosenblatt M (1956) Remarks on some nonparametric estimates of a density function. Ann Math Stat 27(3):832–837. https://doi.org/10.1214/aoms/1177728190

Sackl B, Bergmann H (1987) A bivariate flood model and its application. In: Singh VP (ed) Hydrologic frequency modelling, 571-582, D. Reidel Publishing Company

Saghafian B, Mehdikhani H (2014) Drought characteristics using new copula-based trivariate approach. Nat Hazards 72:1391–1407. https://doi.org/10.1007/s11069-013-0921-6

Saklar A (1959) Functions de repartition n dimensions et leurs marges. Publications de l'Institut de Statistique de l'Université de Paris 8: 229–231

Salvadori G (2004) Bivariate return periods via-2 copulas. J Royal Stat Soc Series B 1:129–144. https://doi.org/10.1016/j.stamet.2004.07.002

Salvadori G, De Michele C (2004) Frequency analysis via copulas: theoretical aspects and applications to hydrological events. Water Resour Res 40:W12511. https://doi.org/10.1029/2004WR003133

Salvadori G, De Michele C (2006) Statistical characterization of temporal structure of storms. Adv Water Resour 29:827–842. https://doi.org/10.1016/j.advwatres.2005.07.013

Salvadori G, De Michele C (2007) On the use of copulas in hydrology: theory and practices. J Hydrol Eng 12:369–380. https://doi.org/10.1061/(ASCE)1084-0699(2007)12:4(369)

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

Salvadori G, De Michele C, Durante F (2011) Multivariate design via copulas. Hydrol Earth Sys Sci Discuss 8(3):5523–5558. https://doi.org/10.5194/hessd-8-5523-2011

Salvadori G, Durante F, Michele CD (2013) Multivariate return period calculation via survival functions. Water Resour Res 49:2308–2311. https://doi.org/10.1002/wrcr.20204

Salvadori G, Tomasicchio GR, D'Alessandro F (2014) Practical guidelines for multivariate analysis and design in coastal and off-shore engineering. Coast Engg 88:1–14. https://doi.org/10.1016/j.coastaleng.2014.01.011

Salvadori G, Durante F, Tomasicchio GR, D'Alessandro F (2015) Practical guidelines for the multivariate assessments of the structural risk in coastal and off-shore engineering. Coast Engg 95:77–83. https://doi.org/10.1016/j.coastaleng.2014.09.007

Santhosh D, Srinivas VV (2013) Bivariate frequency analysis of flood using a diffusion kernel density estimators. Water Resour Res 49: 8328–8343. https://doi.org/10.1002/2011WR0100777

Sarhadi A, Ausin MC, Wiper MP (2016) A new time-varying concept of risk in a changing climate, SCIENTIFIC REPORTS, NATURE|6: 35755|: https://doi.org/10.1038/srep35755

Savu C, Trede M (2010) Hierarchies of Archimedean copulas. Quant Finance 10(3):295–304. https://doi.org/10.1080/14697680902821733

Schmidt T (2007) Coping with copulas. In: Rank J (ed) chap. 1Copulas: from theory to application to finance. Risk Books London, U.K., pp 3–34

Schwartz SC (1967) Estimations of probability density by an orthogonal series. The Annals Mathematical Statistics 38:1261–1265. https://doi.org/10.1214/aoms/1177698795

Scott DW (1992) Multivariate density estimation: theory, Practice and visualization. John Wiley, New York

Sen Z (1999) Simple risk calculations in dependent hydrological series. Hydrol Sci J 44(6):871–878 http://hydrologie.org/hsj/440/hysj_44_06_0871.pdf

Serinaldi F (2015) Dismissing return periods! Stoch Environ Res Risk A 29(4):1179–1189. https://doi.org/10.1007/s00477-014-0916-1

Serinaldi F, Grimaldi S (2007) Fully nested 3-copula procedure and application on hydrological data. J Hydrol Eng 12(4):420–430. https://doi.org/10.1061/(ASCE)1084-0699(2007)12:4(420)

Shakeel AM, Idrees AM, Naeem HM, Sarwar BM (1993) Time series modelling of annual maximum flow of river Indus at Sukkur. Pakistan J Agric Sci 30(1)

Sherring A, Hafizishtiyaq A, Mishra AK, Mohd AA (2009) Stochastic time series model for prediction of annual rainfall and runoff for Lidder catchment of South Kashmir. J Soil Water Conserv 8(4): 11–15

Shiau JT (2003) Return period of bivariate distributed hydrological events. Stoch Environ Res Risk Assess 17(1–2):42–57. https://doi.org/10.1007/s00477-003-0125-9

Shiau JT (2006) Fitting drought duration and severity with two dimensional copulas. Water Resour Manag 20(5):795–815. https://doi.org/10.1007/s11269-005-9008-9

Shiau JT, Modarres R (2009) Copula-based drought severity-duration-frequency analysis in Iran. Meteorol Appl 16(4):481–489. https://doi.org/10.1002/met.145

Shu C, Ouarda T (2008) Regional flood frequency analysis at ungauged sites using the adaptive neuro-fuzzy inference system. J Hydrol 349: 31–43. https://doi.org/10.1016/j.jhydrol.2007.10.050

Silverman BW (1986) Density estimation for statistics and data analysis, 1st edn. Chapman and Hall, London

Singh RS (1977) Applications of estimators of a density and its derivatives. J Royal Stat Soc Series B 39(3):357–363. https://doi.org/10.1111/j.2517-6161.1977.tb01635.x

Singh VP, Singh K (1988) Parameter estimation for log-Pearson type III distribution by POME. J Hydraul Engg ASCE 114(1):112–122. https://doi.org/10.1061/(ASCE)0733-9429(1988)114:1(112)

Singh K, Singh VP (1991) Derivation of bivariate probability density functions with exponential marginals. Stoch Hydrol Hydraul 5:55–68. https://doi.org/10.1007/BF01544178

Smith J, Eli RN (1995) Neural-network models of rainfall-runoff process. J Water Resour Plan Manag 121:499–508. https://doi.org/10.1061/(ASCE)0733-9496(1995)121:6(499)

Song SB, Kang Y (2011) Pair-copula decomposition constructions for multivariate hydrologic drought frequency analysis. Proc 2011 International Symposium on Water Resource and Environmental Protection (ISWREP) 4:2635–2638

Song S, Singh VP (2010) Metaelliptical copulas for drought frequency analysis of periodic hydrologic data. Environmental Research Hazard Assessments 24(3):425–444. https://doi.org/10.1007/s00477-009-0331-1

Sraj M, Bezak N, Brilly M (2014) Bivariate flood frequency analysis using the copula function: a case study of the Litija station on the Sava River. Hydrol Process. https://doi.org/10.1002/hyp.10145

Strupczewski WG, Kaczmarski Z (2001) Non-stationarities approach to at-site flood frequency modelling II. Weighted least square estimation. J Hydrol 248:143–151. https://doi.org/10.1016/S0022-1694(01)00398-5

Supratid S, Aribarg T, Supharatid S (2017) An integration of stationary wavelet transform and nonlinear autoregressive neural network with exogenous input for baseline and future forecasting of reservoir inflow. Water Resour Manag 31:4023–4043. https://doi.org/10.1007/s11269-017-1726-2

Kim TW, Valdes JB, Yoo C (2006) Nonparametric approach for bivariate drought characterisation using Palmer drought index. J Hydrol Eng 11(2):134–143

Tehrany MS, Pradhan B, Jebur MN (2013) Spatial prediction of flood susceptible areas using rule based decision tree (DT) and a novel

ensemble bivariate and multivariate statistical models in GIS. J Hydrol 504:69–79. https://doi.org/10.1016/j.jhydrol.2013.09.034

Tehrany MS, Pradhan B, Mansor S, Ahmad N (2015) Flood susceptibility assessment using GIS-based support vector machine model with different kernel types. Catena 125:91–101. https://doi.org/10.1016/j.catena.2014.10.017

Tian P, Zhao GJ, Li J, Tian K (2011) Extreme value analysis of stream flow time series in Poyang Lake Basin, China. Water Sci Eng 4(2):121–132

Tosunoglu F, Kisi O (2016) Joint modelling of annual maximum drought severity and corresponding duration. J Hydrol (In Press). https://doi.org/10.1016/j.jhydrol.2016.10.018

Twan JA (1988) Extreme value theory: model and estimation. Biometrika 75:397–415

Vandenberghe S, Verhoest N, Baets BD (2010) Fitting bivariate copulas to the dependence structure between storm characteristics: a detailed analysis based on 105 year 10 min rainfall. Water Resour Res 46:W01512. https://doi.org/10.1029/2009WR007857

Vandenberghe S, Verhoest NEC, Onof C, De Baets B (2011) A comparative copula-based bivariate frequency analysis of observed and simulated storm events: a case study on Barlett Lewis Modeled rainfall. Water Resour Res. https://doi.org/10.1029/2009WR008388

Vernieuwe H, Vandenberghe S, Baets BD, Verhost NEC (2015) A continuous rainfall model based on vine copulas. Hydrol Earth Syst Sci 19:2685–2699. https://doi.org/10.5194/hess-19-2685-2015

Veronika BM, Halmova D (2014) Joint modelling of flood peak discharges, volume and duration: a case study of the Danube River in Bratislava. J Hydrol Hydromech 62(3):186–196. https://doi.org/10.2478/johh-2014-0026

Viglione A, Laio F, Claps P (2007) A comparison of homogeneity tests for regional frequency analysis. Water Resour Res 43:W03428. https://doi.org/10.1029/2006WR005095

Vijayakumar N, Vennila S (2016) A comparative analysis of forecasting reservoir inflow using ARMA model and Holt winters exponential smoothening technique. International Jour. of Innovation in Science and Mathematics, 2016, 4(2):. 85–90. ISSN (Online): 2347–9051

Villarini G, Smith JA, Serinaldi F, Bales J, Bates PD, Krajewski WF (2009) Flood frequency analysis for non-stationary annual peak records in an urban drainage basin. Adv Water Resour 32:1255–1266. https://doi.org/10.1016/j.advwatres.2009.05.003

Villarini G, Smith JA, Napolitano F (2010) Nonstationary modelling of a long record of rainfall and temperature over Rome. Adv Water Resour 33:1256–1267. https://doi.org/10.1016/j.advwatres.2010.03.013

Volpin E, Fiori A (2014) Hydraulic structure subject to bivariate hydrological loads: return period, design and risk assessments. Water Resour Res 50:885–897. https://doi.org/10.1002/2013WR014214

Wand MP, Jones MC (1995) Kernel smoothing. Chapman and Hall, London. https://doi.org/10.1007/978-1-4899-4493-1

Whelan N (2004) Sampling from Archimedean copulas. Quant Finance 4(3):339–352

Wigley TML (2009) The effect of changing climate on the frequency of absolute extreme events. Clim Chang 97(1–2):67–76. https://doi.org/10.1007/s10584-009-9654-y

Wong H, Hu BQ, Ip WC, Xia J (2006) Change-point analysis of hydrological time series using grey relational method. J Hydrol 324:323–338. https://doi.org/10.1016/j.jhydrol.2005.10.007

Wong G, Lambert MF, Leonard M, Metcalfe AV (2010) Drought analysis using trivariate copulas conditional on climatic states. J Hydrol Eng 15(2):129–141. https://doi.org/10.1061/(ASCE)HE.1943-5584.0000169

Wu C, Chau K-W (2010) Data-driven models for monthly streamflow time series prediction. Eng Appl Artif Intell 23:1350–1367. https://doi.org/10.1016/j.engappai.2010.04.003

Xu Y, Huang G, Fan Y (2015) Multivariate flood risk analysis for Wei River. Stoch Environ Res Risk Assess. https://doi.org/10.1007/s00477-015-1196-0

Yevjevich V (1972) Probability and statistics in hydrology. Water Resources Publications, Fort Colins

Yue S (1999) Applying the bivariate normal distribution to flood frequency analysis. Water Int 24(3):248–252. https://doi.org/10.1080/02508069908692168

Yue S (2000) The bivariate lognormal distribution to model a multivariate flood episode. Hydrol Process 14:2575–2588. https://doi.org/10.1002/1099-1085(20001015)14:14%3C2575::AID-HYP115%3E3.0.CO;2-L

Yue S (2001a) A bivariate gamma distribution for use in multivariate flood frequency analysis. Hydrol Process 15:1033–1045. https://doi.org/10.1002/hyp.259

Yue S (2001b) A bivariate extreme value distribution applied to flood frequency analysis. Hydrol Res 32(1):49–64. https://doi.org/10.2166/nh.2001.0004

Yue S, Rasmussen P (2002) Bivariate frequency analysis: discussion of some useful concepts in hydrological applications. Hydrol Process 16:2881–2898. https://doi.org/10.1002/hyp.1185

Yue S, Wang CY (2004) A comparison of two bivariate extreme value distribution. Stoch Environ Res Risk Assess 18:61–66. https://doi.org/10.1007/s00477-003-0124-x

Yue S, Ouarda TMBJ, Bobee B, Legendre P, Bruneau P (1999) The Gumbel mixed model for flood frequency analysis. J Hydrol 226(1–2):88–100. https://doi.org/10.1016/S0022-1694(99)00168-7

Yue S, Ouarda TBMJ, Bobee B (2001) A review of bivariate gamma distributions in hydrological application. J Hydrol 246:1–18. https://doi.org/10.1016/S0022-1694(01)00374-2

Zhang L (2005) Multivariate hydrological frequency analysis and risk mapping. Doctoral dissertation, Beijing Normal University

Zhang D (2014) Vine copulas and applications to the European Union sovereign debt analysis. Int Rev Financ Anal 36:46–56. https://doi.org/10.1016/j.irfa.2014.02.011

Zhang L, Singh VP (2006) Bivariate flood frequency analysis using copula method. J Hydrol Eng 11(2):150. https://doi.org/10.1061/(ASCE)1084-0699(2006)11:2(150)

Zhang L, Singh VP (2007a) Bivariate rainfall frequency distribution using Archimedean copulas. J Hydrol 332:93–109. https://doi.org/10.1016/j.jhydrol.2006.06.033

Zhang L, Singh VP (2007b) Trivariate flood frequency analysis using the Gumbel-Hougaard copula. J Hydrol Eng 12(4):431–439. https://doi.org/10.1061/(ASCE)1084-0699(2007)12:4(431)

Zhang R, Xi C, Cheng Q, Zhang Z, Shi P (2016) Joint probability of precipitation and reservoir storage for drought estimation in the headwater basin of the Huaihe River, China. Stoch Env Res Risk A 30:1641–1657. https://doi.org/10.1007/s00477-016-1249-z