Ocean Systems Engineering, *Vol. 9, No. 3 (2019) 287-328*
DOI: https:// doi.org/10.12989/ose.2019.9.3.287

Multivariate design estimations under copulas constructions. Stage-1: Parametrical density constructions for defining flood marginals for the Kelantan River basin, Malaysia

Shahid Latif * and Firuza Mustafa

Department of Geography, University of Malaya, 50603, Kuala Lumpur, Malaysia

(Received February 9, 2019, Revised July 10, 2019, Accepted July 22, 2019)

Abstract. Comprehensive understanding of the flood risk assessments via frequency analysis often demands multivariate designs under the different notations of return periods. Flood is a tri-variate random consequence, which often pointing the unreliability of univariate return period and demands for the joint dependency construction by accounting its multiple intercorrelated flood vectors i.e., flood peak, volume & durations. Selecting the most parsimonious probability functions for demonstrating univariate flood marginals distributions is often a mandatory pre-processing desire before the establishment of joint dependency. Especially under copulas methodology, which often allows the practitioner to model univariate marginals separately from their joint constructions. Parametric density approximations often hypothesized that the random samples must follow some specific or predefine probability density functions, which usually defines different estimates especially in the tail of distributions. Concentrations of the upper tail often seem interesting during flood modelling also, no evidence exhibited in favours of any fixed distributions, which often characterized through the trial and error procedure based on goodness-of-fit measures. On another side, model performance evaluations and selections of best-fitted distributions often demand precise investigations via comparing the relative sample reproducing capabilities otherwise, inconsistencies might reveal uncertainty. Also, the strength & weakness of different fitness statistics usually vary and having different extent during demonstrating gaps and dispensary among fitted distributions. In this literature, selections efforts of marginal distributions of flood variables are incorporated by employing an interactive set of parametric functions for event-based (or Block annual maxima) samples over the 50-years continuously-distributed streamflow characteristics for the Kelantan River basin at Gulliemard Bridge, Malaysia. Model fitness criteria are examined based on the degree of agreements between cumulative empirical and theoretical probabilities. Both the analytical as well as graphically visual inspections are undertaken to strengthen much decisive evidence in favour of best-fitted probability density.

Keywords: flood; block (annual) maxima; parametric functions; marginal distribution; goodness-of-fit

1. Introduction

Basin perspective water resources operational planning, the management or either flood-related infrastructure designing often demands accurate estimations of the flow exceedance probability or

*Corresponding author, Ph.D. Research Scholar, E-mail: macet.shahid@gmail.com

http://www.techno-press.org/?journal=ose&subpage=7 ISSN: 2093-6702 (Print), 2093-677X (Online)

design quantiles under the different notations of return periods in the form of joint probability distributions, conditional distributions or Kendall's distributions or survival functions (Cunnane 1988, Choulakian *et al.* 1990, Stedinger *et al.* 1992, 1993, Bobee and Rasmussen 1994, Calver and Lamb 1995, Yue *et al.* 1999, Yue 2000, Martins and Stedinger 2000, Rao and Hameed 2000, Morrison and Smith 2002, Coles 2001, Kartz *et al*. 2002, Shiau 2003, Salvadori 2004, Zhang and Singh 2006, Genest *et al*. 2007, Salvadori *et al*. 2011, Reddy and Ganguli 2012a, b, Grimaldi *et al*. 2013, Xu *et al*. 2016, Brunner *et al.* 2016). Flood is a multidimensional extreme hydro-climatic consequence, which can't be predicted accurately or precisely through any deterministic or physical procedures but, it could be possible to derive their design episodes through inferring the probability distributions of the historical catchment streamflow characteristics for assessing the hydrologic risk (Bras 1990, Singh and Singh 1991, Goel *et al*. 1998, Sen 1999, Yue and Rasmussen 2002, De Michele *et al*. 2005, Poulin *et al*. 2007, Sraj *et al*. 2014, Fan *et al*. 2015, Xu *et al*. 2015). Flood frequency analysis or FFA usually comprises an inter-association between extreme events quantiles and their non-exceedance probabilities or return periods by fitting univariate or multivariate probability distribution functions or PDF (Cunnane 1989, Yue 1999, Salvadori and De Michele 2004, Zhang and Singh 2007, Karmakar and Simonovic 2009). Trivariate characteristics of the flood hydrograph such as flood peak discharge flow, volume and duration could limit the reliability of univariate frequency analysis and their associated return periods which would be incapable for providing a full screen of the flood hydrograph (Grimaldi and Serinaldi 2006, Serinaldi and Grimaldi 2007, Genest *et al.* 2007, Fan and Zheng 2016). For example, flood events with a peak flow of 100-yr recursion interval could be less intensive and damaging than the same events described based on the joint occurrence between multiple flood vectors such as between the peak-volume /or peak-duration /or volume-duration. Therefore, univariate flood probability constructions would be revealed for the underestimation or overestimations of hydrologic risk and thus could be demanding for the accountability of multiple intercorrelated flood design variables for deriving the return period contours (Veronika and Halmova 2013, Graler *et al*. 2013, Grimaldi *et al.* 2013, Reddy and Ganguli 2012a, b). Frequency analysis through introducing the hydrological models in conjunctions with stochastic rainfall generators, for recognizing the catchment's rainfall-runoff profile through lumped or distributed models (i.e., Claver and Lamb 1995, Boughton *et al*. 2002, Blazkova and Bevan 2004, Lawrence *et al.* 2014), usually demands quite longer computational time and also high spatial and temporal resolutions to achieve a justifiable demonstration of flood extreme (Raquenna *et al.* 2016). Therefore, the statistical treatments of the flood samples under multivariate distributions framework would be an effective & flexible approach.

In general, multivariate constructions usually comprises a combination of the joint probability density functions or PDF and joint cumulative distribution functions or CDF where, 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). In the recent decades, copulas function widely accepted for dealing several hydro-climatic issues, which continuously emphasizing the desire of hydrologist & water practitioner for achieving much comprehensive and higher flexibility in the uncertainty analysis through capturing a wider extent of mutual dependencies (for both the linear and non-linear) in comparison with the traditional multivariate functions (Saklar 1959, Genest and Rivest 1993, De Michele and Salvadori 2003, Salvadori and De Michele 2004, Favre *et al*. 2004). Copulas function frequently incorporated for establishing bivariate joint dependency (i.e., Salvadori 2004, De Michele *et al.* 2005, Nelsen 2006, Zhang and Singh 2006). Similarly, few other demonstrations switched from bivariate into tri-variate probability distribution framework for investigating the importance of tri-variate return periods (Bedford and Cook 2002, Grimaldi and Serinaldi 2006, Genest *et al.* 2007, Kao and Govindaraju 2008, Madadgar and Moradkhani 2013, Graler *et al*. 2013, Daneshkhan *et al*. 2016, Fan and Zheng 2016). One of the insightful flexibilities of copulas methodology, it allows statistician or water experts for separate modelling of their intercorrelated univariate random vectors (i.e., marginal density approximations independently from their joint dependence structure) (Nelsen 2006, Dupins 2007, Mirabbasi *et al.* 2012, Papaioannou *et al*. 2016).

Selecting the most justifiable and parsimonious distributions for defining univariate flood marginals is often a mandatory pre-processing desire before introducing the random vectors into multivariate distribution framework (Reddy and Ganguli 2012, Tosunglou and Kisi 2016). Parametric family functions are frequently targeted to inference about the populations of extreme samples based on finite data, which already mentioned in the above-cited literature. Parametric distribution-based modelling often based on the assumption that the undertaken samples must be following some specific distributions or predefined PDF. In hydrologic data modelling, no universally accepted distributions are assigned from any literature or in favour of any probability distribution functions to model any extreme series (Adamowaski 1985, 1989, Silverman 1986, Dooge 1986, Yue *et al.* 1999, Santhosh and Srinivas 2013). Several models often would fit the data equally well but, each would give different estimates of a given quantile especially, in the tails of the distribution which is solely based on the goodness-of-fit procedure to visualize the compatibility of the fitted distributions (Karmakar and Simonovic 2008). Also, increasing the number of statistical parameters of the fitted distributions usually allows for better flexibility in the context of model complexity (Graler *et al.* 2013). In the most real case study, the best probability distribution for defining flood marginals need not be from the same family. On another side, the performance evaluation of the fitted distributions often demands much precise investigation by comparing their relative fitness measures in the context of the degree of agreement between the cumulative empirical. (i.e., based on plotting-positioning approach) and the theoretical observations (i.e., targeted functions) to assign the best fitted marginal density (Zhang and Singh 2006, Veronika and Halmova 2014) for each flood or hydrologic vectors. Adaption of both the quantitative as well as graphically inspection simultaneously could provide an effective way for revealing much decisive evidence in favour of the most justifiable flood probability density.

Different quantitative based fitness measures based on EDF criteria such as Kolmogorov-Smirnov & Anderson-Darling (A-D) statistics (i.e., Conover 1999, Fan *et al*. 2015, Anderson and Darling 1954), Information criteria statistics such as Akaike information criteria (AIC), Bayesian information criteria (BIC) & Hannan-Quinn Information criteria (HQC) (Akaike 1974, Schwarz 1978, Hannan and Quinn 1979) and either based on error index statistics (i.e., Singh *et al.* 2004, Moriasi *et al*. 2007, Bennett *et al.* 2013) often defines different extent of their strength and limitations to investigate the model suitability. Similarly, in the context of graphically visual inspections i.e., based on Probability-probability (p-p) plot (Chamber *et al.* 1983), Quantiles-quantiles (q-q) plot (Beirlant *et al.* 1996, Willems 1998) or either the Probability difference plot (Mathwave), which usually define different extent to demonstrates the fitness level of fitted distributions. Such as P-P plot defines the theoretical probabilities against specifies fitted distributions while q-q plot estimates the quantiles of data distribution with the standardized theoretical distribution from a specifies parametric family to interpret the shape of the tail of fitted distributions. Selecting an appropriate procedure for estimating the distribution parameter of the fitted distributions is also an essential concern in the fitting process. Maximum likelihood estimation or MLE (i.e., Cohn *et al.* 1997, Owen 2008) often consider as a standard parametric estimator which poses minimum sampling variance during the parameter estimations (Tosunoglu and Kisi 2016) and more effective during the large sample size due to its asymptotic optimal properties (Shao *et al*. 2008). Besides this, L-statistics based L-moment estimators (Hosking and Wallis 1987) and the Method of moment (MOM) (Bain and Engelhardt 1991, Rao and Hameed 2000) also seems quite popular in the fitting procedure where the L-moment estimators frequently undertaken for meteorological data analysis also, their performance seems competitive for small sample size and heavy tail distributions in relative with their peer estimators i.e., MLE and MOM.

As, copulas multivariate constructions eliminated the restrictions, in the context of approximating univariate flood marginals not necessary from the same parametric families, which would be following different distributions and need to model separately. On the other side, few demonstrations such as Schwarz (1967), Duin (1976), Singh (1977), Bowman (1984), Silverman (1986), Adamowski *et al*. (1989), Scott (1992), Lall (1995), Wand and Jones (1995); Jones *et al*. (1996), Lall *et al*. (1996), Adamowski (1996), Bowman and Azzalini (1997), Efromowhich (1999), Duong and Hazelton (2003), Kim *et al*. (2003, 2006), Ghosh and Mujumdar (2007) and Srinivas and Santhosh (2013) pointed the limitations of parametric distributions in case of unsymmetrical or multimodal distributions type and pointed towards the flexibility of nonparametric probability framework. The applicability of nonparametric estimations based on kernel density functions are beyond the scope of this literature and will be tackled in the separate paper. This paper only focuses via parametric estimation procedures for demonstrating the adequacy of an interactive set of probability distribution functions for characterizing the univariate flood marginals for at-site event-based methodology.

From the past few decades, the Kelantan River basin often subjecting to the most severe monsoonal flooding in Malaysia and perceiving for increasing in term of their frequency and magnitude (DID 2000, 2003, 2004, MMD 2007, Adnan and Atkinson 2011). According to Chan (1995) investigation and DID, (2000-2006) reports, the expectation of the occurrence of catastrophic flooding has increased from once in every 50 years to 15 years from 2004 in Kelantan. For example, intense and prolonged precipitation in the year 2002 caused flooding of a total area of 1640 km2 and affected the population of 714 287. Similarly, in the early month of December 2014, much heavy precipitation occurred for many of days triggered the flood event in the several parts of the east coast of the Kelantan river basin. It was the worst flood ever recorded in history and affected more than 20,0000 people. The Kelantan River basin is one of the largest basin Malaysia, which known to be flood prone. The maximum length and breadth of this catchments area are 150 km and 140 km, respectively. The river is about 248 km long and drains an area of 13100 km^2, occupying more than 85% of the State of Kelantan. The basin has an annual rainfall of about 2500 mm much of which occur during the North-East Monsoon (or wet season) between mid-October and mid-January. According to the study performed by Adnan and Atkinson (2011), it is clearly indicating the existence of a significant trend in streamflow samples for both the upstream (River Galas) and downstream (River Kelantan) sub-catchments such that in the downstream area streamflow increased in the wet season. Similarly, Hussain and Ismail (2013) investigation pointed that Gulliemard Bridge, Lebir and Galas stations have highest in flood frequency rather than Nenggiri station and also the value of damaged property got increased according to the frequency of flood happening. Abdulkareem and Sulaiman (2015) investigated the variability of precipitation in flood source area of Kelantan river basin through the trend analysis using based on annual maximum series of 24-hour precipitation data and annual maximum flood data and revealed that no statistically significant trend was detected in the annual maximum series of 24-hr precipitation between for the period 1984-2014 while the AMF series were significant at 5%

level at the targeted locations. Alamgir *et al.* (2018) performed multivariate analysis of flood episodes and established the joint distribution relation among multiple intercorrelated flood vectors for the different gauge station of this river basin. Result revealed that for most of the station, Normal or Log-normal fitted best for duration samples, Gamma or Gumbel for flood peak series and Exponential or Log-normal for volume series. Nashwan *et al.* (2018) performed flood susceptibility assessments at the different gauge station locations of Kelantan river basin and based on the multivariate joint return period of the flood variables, it concluded that the downstream area is the highest risk of devastating flood events. In the recent years, a lot of attention has been pointed out towards the impact of land use changes may affect the catchments response such as Hassan (2004) report revealed the substantial land use changes in this region and which might influence the rate of evapotranspiration and infiltrations (Wooldridge 2001). Wan (1996), Jamaliah (2007) literature also pointed the existence of rapid land use changes from the year 1970s to 2000s, mainly due to deforestation and conversion from natural land into agriculture for oil palm and rubber.

The cause of frequent failure of hydrologic or flood defence infrastructure in Malaysia due to the impact of moderately severe of flood episodes might be attributed due to the lack of complete flood hydrograph or in other words, where only flood peak discharge samples often targeted in deriving flood frequency curve during the structural development. Importance of the accountability of all three flood characteristics i.e., flood peak discharge flow, hydrograph volume and duration for practical applications in hydrology or hydrologic risk assessments are already discussed in the above paragraph. For example, according to Gaal *et al.* (2015) the designing of retention basins and spillways of reservoirs or any other flood defence hydraulic structures where the storage is involved, in such circumstances the estimation of hydrograph volume must be required along with peak discharge, in order to calculate the impact of inflow on the storage. Similarly, estimating the joint behaviour of flood peak-volume and volume-duration would be effective for flood diversion practices and flood control pressures practical (Fan *et al*. 2015, Xu *et al.* 2015). Therefore, multivariate designs and their associated return periods could be a comprehensive way of tackling such extreme issues through making a defensive risk-based decision making in this river basin.

The entire investigation process is divided into the two study portions, where each part covered a full paper. The objective of this study portion is to perform a better selection procedure and practices in the model evaluation criteria for pointing the most parsimonious marginal distribution of flood characteristics i.e., flood peak discharge flow, hydrograph volume and duration, by introducing a variety of mono-parametric (1-parameter), bi-parametric (2-parameters) and tri-parametric (3-parameters) and also tetra-parametric(4-parameters) probability distribution functions. At-site event-based or block (annual) Maxima-based methodology is adopted over the 50-years (1961-2016) continuously-distributed streamflow characteristics for the Kelantan River basin at Gulliemard Bridge gauge station, Malaysia for selecting an appropriate marginal density for flood distributions series. The flood characteristics such as flood peak, volume and duration are derived from the daily basis streamflow records at an annual scale, where hydrograph volume and duration series are derived corresponding to each flood peak samples, details are provided in section 3.2. The demonstrations over how these multiple inter-associated flood marginals will be introduced further for establishing copula-based multivariate dependency simulations will appear in the next paper separately which are beyond the scope of this study. As a prerequisite, an overview of distinct varieties of univariate parametric family functions as well as the briefing of different inferential measuring or goodness-of-fit test statistics for identifying most parsimonious and best-fitted models are given subsequently in the sub-section of the second section. Details of the case study and extraction of multivariate flood characteristics via event- based flood sampling procedure are in the third section. Fourth section provides result and discussion. Research conclusion are pointed in the fifth section.

2. Theoretical framework

2.1 Univariate parametric density function

Every joint distribution practice implicitly defines both the descriptive information about marginal behaviour of each targeted random vectors and their joint dependence structure for capturing mutual concurrency or correlation structure (Singh and Singh 1991, Rao and Hameed 2000, Kartz *et al.* 2002). If $x_1,\ x_2,\ x_3,\ \ldots\ldots\ldots x_n$ defines the nth set of historical random observations with independent and identically distributions or i.i.d i.e., time-independency behaviour or no serial correlations then the procedure for the univariate density approximation usually comprises (1) choosing distinguish class of probability functions as a candidate structure, (2) selecting a parameter estimations algorithm for determining vectors of unknown statistical or distribution parameters of fitted distributions, (3) outlining of uncertainty between the distributions parameters and fitted probabilities based on the models compatibility or goodness of fit criteria (Benth and Saltyte- Benth 2005, Cong and Brady 2012). In other words, marginal distribution construction of the flood vectors is the procedure to make an inference about populations based on finite random sample size or to extrapolate flood design quantiles beyond the existing data range (Sen 1999, Coles 2001, Griffis and Stedinger 2007).

The traditional approach of flood frequency analysis often incorporated via time-invariant or static probability distribution framework and which often a mandatory desire before introducing flood vectors into the univariate or multivariate framework (Khaliq *et al*. 2006, Tosunoglu and Kisi 2016, Daneshkhan *et al*. 2016). Therefore, samples will be undergoing for some statistical treatments such as, time-trend analysis (e.g., monotonic trend investigations) (Mann 1945, Kendall 1975, Zhang 2005, Liu and Cui 2008, Hameed 2008, Hamid *et al*. 2014) or either via the Q-statistics, for identifying serial correlations (Ljung and Box 1978, Cong and Brady 2012). Table 1 listed a distinguish variates of some frequently incorporated 1-dimensional probability functions, where each function exhibited a different modelling extent to capture flood distribution samples.

Earlier demonstrations such as McMohan and Srikanthan (1981), Rossi *et al*. (1984), Wallis (1988), Vogel *et al*. (1993) explored much closer attention towards flood probability analysis in the light of most favourable distribution framework. Besides this, Cunnane (1989) undertaken an extended review through investigating the issue of flood frequency for a region and country and revealed that the empirical suitability plays a much larger role in the selection of distributions. Adamowski (1989), Hall (1984) pointed towards the GEV distributions as most consistence parametric functions for the UK region. Similarly, the Log-Pearson Type-III are recommended as the most justifiable distribution among their peer parametric functions for the US river basin according to US Water Resources Council (Adamowski 1989). The Gumbel functions are categorized as a standard and most effective distribution over the Spain and Finland region (Salinas *et al.* 2014). Actually, the Gumbel function is highly flexible to make an accurate representation of any extreme maxima samples (i.e., peak flood discharge or maximum rainfall samples) (Alam *et al*. 2018). Mathematically, Gumbel and Frechet distributions are the special categories which can be derived from the GEV distribution. Such that, by letting the shape parameter of GEV functions to zero i.e., '$k \to 0$ (tends to zero)' could result for Gumbel functions which, usually characterized by the thin upper or light tail and exhibited unbounded behavior (Coles 2001, Khaliq *et al*. 2006). Similarly, by approximating the positive value to the GEV shape parameter i.e., '$k > 0$ (greater than zero)', it will take the form of Frechet distribution, which usually characterized via heavy or polynomially decreasing tail structure with unbounded behavior (Graler *et al*. 2013, Reddy and Ganguli 2012b). The Weibull distribution is also categorized in the GEV distribution family, which often signifies for bounded upper tail behaviour (Stedinger *et al*. 1993). Ekanayake and Cruise (1993) performed a comparative analysis between the Weibull distribution and Exponential distribution for partial duration-based flood characteristics. On another side, the Exponential function is a special class of Gamma distribution family (Hosking and Wallis 1997). Chen *et al*. (2017) demonstrated the importance of Generalized Gamma distribution in the flood frequency analysis where the principle of maximum entropy (POME) theory is adopted to estimate the distribution parameter for the flood samples. It is a generalized version of 2-parameter Gamma distribution.

In probability theory, the Inverse Gaussian distribution (also known as Wald distribution) is a two-parameter family of continuous distribution with support on $(0\ \infty)$. According to Table 6, if the shape parameter of the Inverse Gaussian tends to infinity, it becomes more like a Gaussian distribution. Markiewicz *et al.* (2015) modelled flood samples via Inverse Gaussian and Generalized exponential distributions. Daneshkhan *et al.* (2016) introduced Inverse Gaussian distribution in modelling of flood peak and volume samples. Haktanir (1992) pointed out the prediction capability of bi-parametric (two-parameters) and tri-parametric (three parameters) Log-normal and Gumbel functions for inferring the right-tail extreme events for annual basis flood peak series. Similarly, Log-logistic function also exhibited a lot of attention in extreme value analysis (Singh *et al*. 1993, Shoukri *et al*. 1988, Ashkar and Mahdi 2003). The 3-parameter Log-Gamma function also exhibited considerably level flexibility and importance in hydrological data stimulations (Veronika and Halmova 2014). Besides this, a special class of four parameters Johnson family functions called Johnson SB distribution are rarely incorporated in hydrologic or flood modelling (Johnson (1994) & Cugerone and De Michele (2005). Few demonstrations introduced Johnson SB function such as a dynamic stochastically hydrological simulation of the probable maximum discharge series by Kushment and Gelfan (2011) or either, investigating the rainfall integral parameters estimation by comparing Johnson SB distribution with 3-parameter Gamma function by D'Adderio *et al.* (2016). Johnson SB functions would be a better choice for modelling extreme samples, as their PDF are thinner and decreasing more exponentially rather than algebraically for a larger value of the given observations, which also suitable for investigating survival functions of a given sample. Table 6 listed the probability density function or PDF, cumulative function or CDF and parameters of marginals distribution going to incorporate in this study.

2.2 Model compatibility investigation based on goodness-of-fit measures

Selecting a justifiable probability distribution based on the degree of agreement between empirical and theoretical observations often demands higher accuracy and precision by adopting a series of quantitative as well as graphical inspections. Empirical non-exceedance probabilities i.e., $\mathrm{P}(\mathrm{K} \leq \mathrm{k})$ are frequently derived from the Gringorten based plotting-positioning formula (Gringorten 1963, Cunnane 1978) and which usually compared against the CDF of the fitted distributions for pointing the gaps or deviations between empirical and fitted samples and can be mathematically formulated as

$$\text{Empirical Cumulative frequency} = P(K \leq k) = (k - 0.44)/(N + 0.12) \quad (1)$$

where, N signifies the length of the sample (i.e., the total number of observations) and k is the kth smallest observations in the data set arranged in ascending order. Statistical approaches of the several analytical fitness measures i.e., based on empirical distribution functions (or EDF), information criteria statistics and error indices statistics are reviewed in the next paragraph which is often considered as the most effective ways for outlining the relative gaps and discrepancy (Dufour *et al*. 1998, Seier 2002, Arshad *et al*. 2003).

The Kolmogorov-Smirnov or K-S is an empirical distribution function (or EDF) for investigating the largest vertic

$$D_n = \sup_x |F^*(x) - F_n(x)| \quad (2)$$

$$P(D_n \leq D_n^{\alpha}) = 1 - \alpha \quad (3)$$

where, $F^*(x)$ and $F_n(x)$ pointed out the theoretical & empirical probabilities and D_n is the critical value at level of significance 'α'. Statistically, the individual flood characteristics which following some specific or pre-defined distribution are usually taken under the null hypothesis 'H_0' against their alternative hypothesis 'H_1' as given below:

Null hypothesis (H_0) = data follow the specified distribution

Alternative hypothesis (H_1) = data didn't follow any specific distribution

Table 1 Families of univariate probability distribution functions

Parametric family functions	References
Log-Pearson Type III	Bobee 1974, Rao 1980, McMohan and Srikanthan (1981), Arora and Singh 1988, Singh and Singh 1988,Adamowaski 1989, Haktanir and Horlacher 1993, Singh 1998, Griffis and Stedinger 2007, Vogel *et al*. 1993, Haddad and Rahman 2008
Log-Logistic	Haktanir and Horlacher 1993, Singh *et al*. 1993
Gamma	Reddy and Ganguli 2012, Xu *et al.* 2015
Generalized Gamma	Keshtkaran *et al*. 2011, Chen *et al.* 2017
Johnson SB	Keshtkaran *et al.* 2011, Cugerone and De Michele 2015
Lognormal	Yue 2000, Yue 2001, Karmakar and Simonovic 2008
Generalized extreme value (or GEV)	Karim and Choudhary 1995, Nadarajah and Shiau 2005
Exponential	Correia 1987, Krstanovic and Singh 1987, Chouklian *et al*. 1990
Gumbel	Jain and Singh 1987, Veronika and Halmova 2014
Inverse Gaussian	Daneshkhan *et al*. 2016 and Markiewicz *et al.* (2015)
Log Gamma	Veronika and Halmova 2014
Frechet	Reddy and Ganguli 2016
Weibull	Ekanayake and Cruise 1993, Graler *et al*. 2013, Rauf and Zeephongsekul 2014

The acceptance of null hypothesis ‘H_0’ will be based on the p-value of the fitted distributions and that must be greater than the critical value i.e., $\mathrm{p}_{\mathrm{critical}}(0.05) < p_{model}$ for considerably better performance (Xu *et al*. 2015). Also, the estimated K-S value must be below their critical value i.e., $d_{model} < d_{critical}$ at a pre-defined significance level i.e., ‘$\alpha = 0.05$’ that could be revealed for the positive decision in favor of good agreement between fitted and empirical distributions otherwise performance could be inferior and liable for the rejection (Yue and Rasmussen 2002, O’Connor and Kleyner 2012, Reddy and Ganguli 2012, Madadgar and Moradkhani 2013, Zhang *et al*. 2016).

Farrel and Stewart (2006) study revealed that the K-S statistics are characterized by relatively flat in their tail distributions for both the empirical and fitted probabilities thus, would be complicating for handling maxima samples. Also, K-S is quite less sensitive near the distribution tails in relative with the center of distributions (ITL NIST e-Handbook). Therefore, the quadratic class of EDF, also called the Anderson-Darling (A-D) statistics widely accepted for evaluating the model performance through examining whether the random samples come from some specified distributions (Anderson and Darling 1954, Scholz and Stephens 1987, Farrel and Stewart 2006, Fan *et al.* 2015, Alam *et al*. 2018) can be mathematically expressed as

$$\mathrm{A}^2 = -\mathrm{n} - \frac{1}{\mathrm{n}}\sum_{\mathrm{i}=1}^{\mathrm{n}}(2\mathrm{i}-1)\{\log \mathrm{F_X}(\mathrm{x_i}) + \log(1-\mathrm{F_X}(\mathrm{x_{n+1-i}}) \tag{4}$$

where, $F_X(x)$ representing the theoretical cdf for i^{th} sets of random samples ‘$\mathrm{x_i}$’. In compare with K-S statistics, A-D statistics facilitates an extra weight towards the tail of probability distributions and quite be more consistent in case of large extreme modelling (Farrel and Stewart 2006, Alam *et al.* 2018). The estimated A-D values are below their critical level at the significance level 'α', could indicate better model performance with the observed samples.

On another side, the Kullback- Leibler information measures (i.e., Kullback-Leibler (1951)) often comprises a key to derive the information criteria statistics which investigate the fitness compatibility of the distribution models. Information criteria statistics such as Akaike Information criteria (or AIC) (Akaike 1974, Cong and Brady 2012), Schwartz’s Bayesian Information criteria (or BIC) (Schwarz 1978) and Hannan-Quinn Information criteria (HQC) (Hannan and Quinn 1979, Burnham and Anderson 2002) usually highlights the trade-off relationship between model bias or uncertainty with the number of fitted parameters. The AIC statistics included the lack of the fit of model at one hand and unreliability of the model due to the number of model parameters on the other hand (Zhang and Singh 2007) and mathematically expressed as

$$A\mathrm{IC} = -2\log(\text{Maximized Likehood for Fitted Model}) + 2(\text{Number of Fitted Model Parameters}) \tag{5}$$

It can be express in the context of Mean Square Error (or MSE) (Karmakar and Siminovic 2008)

$$A\mathrm{IC} = -2\log(\mathrm{MSE}) + 2(\text{Number of Fitted Model Parameters}) \tag{6}$$

Minimum the AIC value indicates for best-fitted distribution among their peer candidates. Similarly, the Bayesian based information criteria statistics or BIC can be formulated as

$$B\mathrm{IC} = -(\text{Sample size})\log(\text{Maximized likehood for fitted distirbutions}) + [\text{No. of fiited model parameters Log(Sample size)} \tag{7}$$

or

$$BIC = -(\text{Sample size})\log(\text{MSE}) + [\text{No. of fiited model parameters Log(Sample size)} \quad (8)$$

Minimum the BIC value indicates for the best-fitted distribution among their peer candidates.

On the other side, Hannan-Quinn Information criteria or HQC is another alternative to AIC and BIC statistics which can be formulated as (Hannan and Quinn 1979, Burnham and Anderson 2002)

$$\text{HQC} = -2\text{L}_{\text{max}} + 2\text{klog}(\log(\text{N})) \quad (9)$$

where, L_{max} signifies the model log-likelihood of the total number of fitted parameters 'k' for the 'N' sample size. According to Burnham and Anderson (2002), HQC is not the estimator of Kullback-Leibler divergence and also not exhibited asymptotically efficient criteria (Claesken and Hjort 2008, Haggag 2014). Such characteristics are identical with the BIC criteria statistics. As pointed from the same literature that HQC exhibited a higher level of consistency in compare to AIC or BIC. Minimum the value of HOQ statistics usually indicates for justifiable model performance with the observational samples.

Table 2 Error indices statistics and their analytical expression

Measuring statistics	Analytical expressions	References
Mean Square Error (MSE) & Root Mean Square Error (RMSE)	$MSE = \sum_{i=1}^{N}\left(x_i^{Model} - x_i^{Empirical}\right)^2/N$ & $RMSE = \sqrt{MSE} = \sqrt{\sum_{i=1}^{N}\left(x_i^{Model} - x_i^{Empirical}\right)^2/N}$, where x_i indicating the ith series of sample size 'N also, $RMSE = \sqrt{\sum_{i=1}^{N}\left(x_i^{Model} - x_i^{Empirical}\right)^2/N - k}$, in term of accounting the number of fitted model parameters 'k'.	Singh *et al.* 2004, Moriasi *et al.* 2007, Gupta *et al.* 2009, Chai *et al.* 2014
Mean Absolute Error (MAE)	$\frac{1}{N}\sum_{i=1}^{N}\lvert x_i - \widehat{x_i}\rvert$	Singh *et al.* 2004, Moriasi *et al.* 2007, Bennett *et al.* 2013
RMSE-observations standard deviations ratio (RSR)	$\sqrt{\sum_{i=1}^{N}\left(x_i^{Observed} - x_i^{Simulated}\right)^2} \Big/ \sum_{i=1}^{N}\left(x_i^{Observed} - x^{Average}\right)^2$, where $x_i^{Observed}$ & $x_i^{Simulated}$ demonstrating the empirical and fitted probabilities of a random series 'x_i' and their mean value $x^{Average}$.	Moriasi *et al.*, 2007, Bennett *et al.* 2013
Nash-Sutcliffe Efficiency (NSE)	$1 - \sqrt{\sum_{i=1}^{N}\left(x_i^{Observed} - x_i^{Simulated}\right)^2} \Big/ \sum_{i=1}^{N}\left(x_i^{Observed} - x^{Average}\right)^2$	Nash and Sutcliffe 1970, Sevat and Dezetter 1991, Legates and McCabe 1999, Gupta *et al.* 2009

Table 2 summarizing some frequently undertaken error indices statistics in the field of hydro-climatic model validations. Statistical metrics such as mean square error or MSE, root mean square error or RMSE and mean absolute error or MAE, usually defines error statistics in the units of constituents of interest such that, a value closer or either zero often indicates for consistency or optimum model performance (Singh *et al*. 2004, Moriasi *et al*. 2007, Chai and Draxler 2014). Both MAE & RMSE metrics are frequently incorporated over the literature where the MAE statistic is solely based on the absolute value to minimize bias towards the large events in relative with the RMSE (Bennett *et al*. 2013). Willmott and Matsuura (2005) study revealed that MAE would be a better approach than RMSE while, Chai and Draxler (2014) pointed that RMSE performance could be dominating over the MAE statistics only for Gaussian distribution samples. The RMSE-observations standard deviations ratio or RSR statistics often facilitates additional flexibility to achieve simultaneous significance of error index statistics and normalizations factors (Singh *et al*. 2004, Bennett *et al*. 2013). In actual, RSR statistics are obtained through standardizing the RMSE statistics by integrating the influence of samples standard deviations. Numerically, the RSR statistics often defined within a range of 0 (which revealing for justifiable performance) to any larger positive value which usually indicates for model inconsistencies. Similarly, Nash-Sutcliffe Efficiency or NSE based model evaluation criteria compared the data and residual variance structure and numerically defined within a range of $-\infty$ (i.e., inferior performance) to 1(i.e., ideal fitness level) (Nash and Sutcliffe 1970). Numerically, the value of NSE within the range of [0.0, 1.0] must signify for good agreements where the zero value often reveals that the model performance is no better than simply using their average value (Bennett *et al*. 2013). Literature such as Sevat and Dezetter (1991), Legates and McCabe (1999) demonstrated the fitness measuring strength of NSE statistics, especially during the hydro-climatic model simulations.

3. Details of the case study

3.1 Study Area

Monsoonal flood happening seems to be increased in the Kelantan River basin, Malaysia from the last few decades in term of frequency as well as magnitude, according to the report of Drainage and Irrigation Department, Malaysia (DID 2000, 2003, 2004) and Malaysia Meteorological Department (MMD 2007). Few flood related studies pointed that such extreme hydrologic consequences are mainly due to rapid human intervention from natural to land use activities in the form of deforestations or land clearance either for promoting the agricultural activities i.e., palm oil and rubber plantations or either due to logging activities (Chan 1995, Jamaliah 2007, Adnan and Atkinson 2011). The Kelantan River is the longest river of Kelantan state originating from the Tahan mountain range to the South China Sea in the north-eastern part of Peninsular Malaysia between the geographical location of $\text{Lat}\ 4^\circ\ 30'\ \text{N to}\ 6^\circ\ 15'\ \text{N and Long}\ 101^\circ\text{E to}\ 102^\circ\ 45'\ \text{E}$. River Galas and River Lebir are the two major tributaries of Kelantan River. The river is about 248 km long and drains an area of 13100 km^2, occupying more than 85% of the state of Kelantan. According to DID flood report i.e., DID (2000), the estimated runoff is about 500 $m3sec{-1}$ and the variations of annual precipitations for this region in between 0 mm (dry period)-1750 mm (wet or north-eastern monsoonal period). The major land use of this area is agriculture (i.e., paddy, rubber and oil palm) for midstream and downstream and forest for upstream (i.e. near to Gua Musang). The study performed by Adnan and Atkinson (2011), clearly indicating the existence of a significant trend in streamflow samples for both the upstream (River Galas) and downstream (River Kelantan) sub-catchments such that in the downstream area streamflow increased in the wet season. Also, precipitation trends were also increasing in the wet season or winter monsoon circulation in the downstream region. Also, Hussain and Ismail (2013) investigation pointed that Gulliemard Bridge, Lebir and Galas stations have highest in flood frequency rather than Nenggiri station and also the value of damaged property got increased according to the frequency of flood happening. Therefore, this literature targeted 56 years (1961-2016) daily basis streamflow discharge records which are collected and provided by the Drainage and Irrigation Department (DID) Malaysia for the Gulliemard Bridge gauge stations, which is located at the downstream of Kelantan river near the Kuala Kari region. At-site event-based methodology is adopted for constructing univariate marginal density of flood characteristics (i.e., flood peak flow, volume and duration).

3.2. Sampling of flood episodes

Flood probability construction via the partial data series only focuses the extreme hydrograph portion i.e., either high flow (for flood episodes) or low flow (for drought events) instead of visualizing the entire hydrograph (Kite and Stuart 1977, Hosking *et al.* 1985, Chow *et al*. 1988, Bras 1990, Rao and Hameed 2000, Brunner *et al.* 2016). Annual (Maximum) series or AM also called block (annual) maxima and Peak over Threshold (or POT) are the two frequently modelling techniques widely accepted in the extreme probability simulations (Hosking 1987, Bras 1990, Madsen 1997). Event-based or AM series usually defining the random distributions at an annual scale for each targeted or study site by pointing the maximum streamflow discharge value i.e., flood peak. On the other side, POT based flood sampling includes all the peak values above the pre-defined threshold values not just by considering a single peak like the AM series. Therefore, such restrictions in the AM based flood sampling sometime might be problematic, if the second largest peak in the same years is larger than the other year samples events. On another side, POT based event sampling could demand for time independency between the two-consecutive peak within the same observation year, which usually a complicated effort during the selection of flood peak (Wilems 2005).

Therefore, the present study incorporated event-based procedure, which often revealing an intuitive procedure for handling complex hydrological practical and also facilitates to establishing conditional distribution relations among the multiple design vectors in case of multivariate joint distribution analysis (Reddy and Ganguli 2012a, Sraj *et al.* 2014, Papaiannou *et al*. 2016, Tosunglou and Kisi 2016). The flood characteristics which are used in this study such as flood peak discharge flow (P), hydrograph volume (V) and duration (D) are obtained from daily basis stream flow observations. The characterizations of flood peak flow are based on their maximum streamflow discharge records at an annual scale which means for each year there will be only one flood episodes at the targeted site (Yue *et al*. 2000, Yue and Rasmussen 2002, Xu *et al*. 2015, Fan *et al.* 2015). On the other side, its other inter-associated vectors i.e., hydrograph volume & duration are retrieved corresponding to each annually basis peak flow samples by separating the base flow (i.e., low frequency components) from high frequency or direct runoff components (i.e., volume extraction), and based on time differencing between the rising (SDi) and recession (EDi) limb of targeted peak curve for exacting the durations series (Yue and Rassumsen 2002, Yue *et al.* 2002, Sraj *et al.* 2014). Recursive digital filtering procedure in the form of either one parameter digital filtering (i.e., Lyne and Hollick 1979, Nathan and McMahon 1990, Arnold and Allen 1999, Eckhardt 2004, Ladson *et al*. 2013) or either based on Eckhardt (2005) recursive filtering algorithm are the two different ways for extracting the low- frequency components or base flow separation. Eckhardt (2005) algorithm usually provided an effective way for discriminating low frequency from direct-surface runoff portions and would be effective for the wider verity of catchments to reveal much consistent measure (Lim *et al*. 2005, Zhang 2013). Besides this, readers are advised to follow Gonzales *et al.* (2009) for the extended comparison among several baseflow separation algorithms.

Flood peak discharge often attains their maximum value but not mandatory for hydrograph volume & duration series (Sraj *et al.* 2014, Xu *et al.* 2015). Mathematically, we are required to derive three random vectors for each of the ith years can be formulated as

$$V_i = V_i^{total} - V_i^{Baseflow} = \sum_{j=SD_i}^{ED} Q_{ij} - \frac{(1+D_i)(Q_{is}+Q_{ie})}{2} = \text{ hydrograph volume series} \quad (10)$$

$$D_i = ED_i - SD_i = \text{Hydrograph durations for ith year} \quad (11)$$

and

$$P_i = \max\{Q_{ij}, j = SD_i + SD_i + 1, \dots\dots.., ED_i\} = \text{Annual flood peak series for the ith year} \quad (12)$$

where $'Q_{ij}'$ signifies for j^{th} days streamflow magnitude for the i^{th} year; $'Q_{is}'$ & $'Q_{ie}'$ reveals for streamflow magnitude for the start date ‘SD_i' and end date ‘ED_i' of the flood runoff.

4. Results and discussions

4.1 Test for independency and identical distribution

Each univariate flood vectors exhibited positively skewed distributions in which, the duration series exhibited a quite a higher degree of unsymmetrical behaviour as observed from Table 3 and Fig. 2. Thus, the preliminary investigation pointed towards the sharp with right-tailed distributions would be suitable for each individual series (i.e., Xu *et al.* 2015). The individual flood characteristics need to be stationary or time-independency behaviour before introducing into the univariate distribution framework. Statistically, the independent series usually comprising for no inter-connection between any two or more random events while, the identical distribution could indicate for zero trend/or fluctuations which also reveals that all samples are drawn from the same probability distributions (Khaliq *et al*. 2006, Daneshkhan *et al*. 2016, Tosunglou *et al.* 2016).

For this, Ljung and box (1978) based hypothesis testing also called Q-statistics, are undertaken for investigating whether the individual series are time-independent (i.e., stationary) with no serial correlation or autocorrelation (Benth and Saltyte- Benth 2005, Cong and Brady 2012). Actually, the presence of autocorrelations increases the variance of residual and could be responsible for minimizing model efficiency (Cong and Brady 2012). Statistically, under the null hypothesis H0, Q-statistics usually follows a chi-square distribution with ‘h’ degree of freedom (Ljung and Box 1978 and Daneshkhan *et al.* 2012). Thus, Q-statistics for the sample size ‘n’ with total no of lag being tested i.e., ‘t’ with sample autocorrelations at the specific lag i.e., $\widehat{\rho_t}$ are given below

$$Q = n(n+2)\sum_{t=1}^{h} \hat{\rho}_t^2 / n - t \quad (13)$$

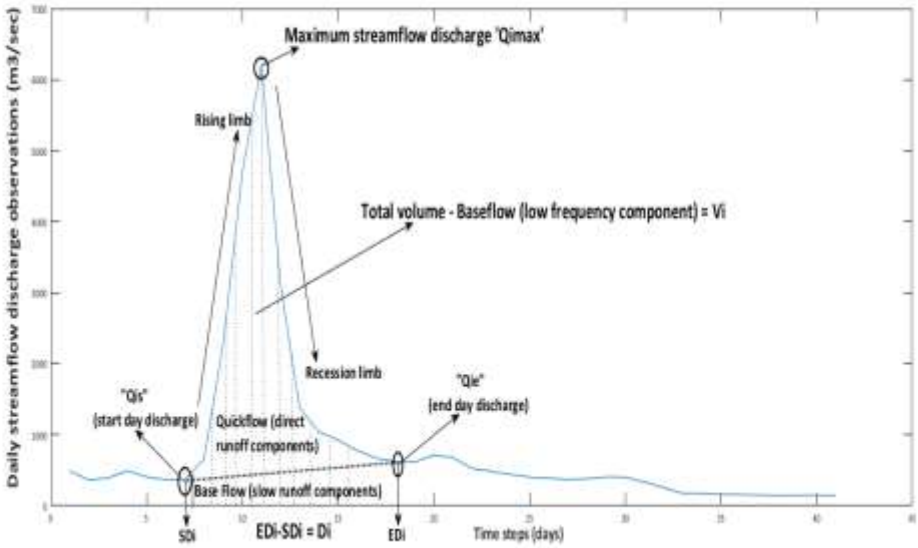


Fig. 1 A typical hydrograph characteristic for the ith flood episodes

where, the null and alternative hypothesis is
Null hypothesis (H_0) = zero autocorrelation or independent distributions
Alternative hypothesis (H_1) = existence of serial correlation (or autocorrelation)

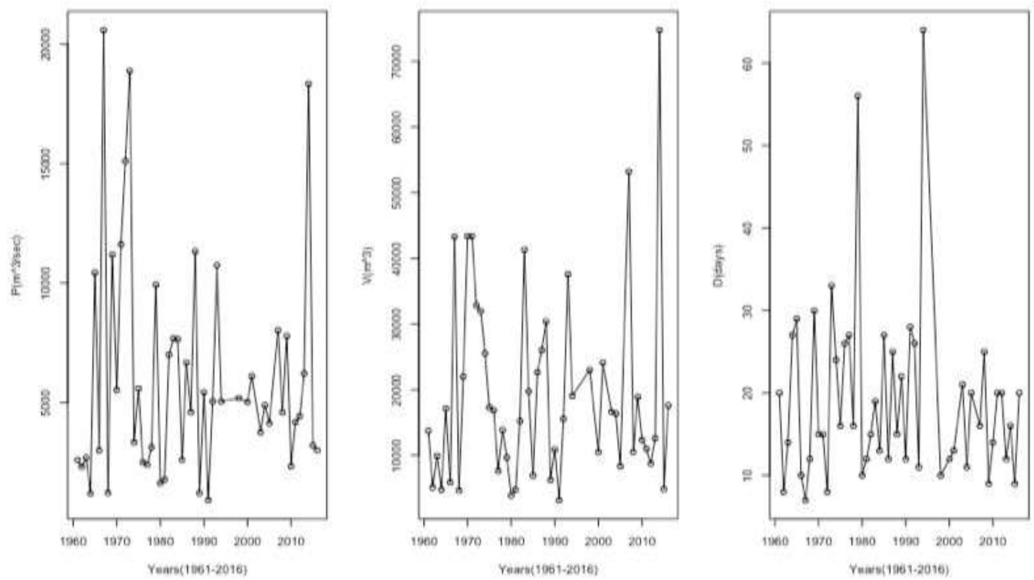
Table 4a summarizing the estimated Q-statistics and their associated p-value for different lag size (i.e., 30, 20, 15, 10 & 5) and which pointing for almost negligible or zero first-order autocorrelations as their estimated statistics are below their critical value for each of the univariate series by accepting the null hypothesis (H_0) at 5% or 0.05 significance level against their alternative hypothesis (H_a). Besides this, Fig. 3 illustrating the sample autocorrelation as well as partial autocorrelation graphs and which also indicating the existence of no serial correlation.

Also, the non-parametric rank-based Mann-Kendall or M-K test is also incorporated for visualizing the monotonic trend within the historical series under the null hypothesis (H0) against their alternative hypothesis (H1) (Mann 1945, Kendall 1975, Hameed *et al.* 2008). Such time-series manipulations exhibited a lot of flexibility such as tackling of skewed or unsymmetrical distributions, also justifiable performance for the missing distributions along with the higher degree of resistance with outliers (Kahya and Kalayci 2004, Xu *et al.* 2005, Modarres and da Silva 2007).

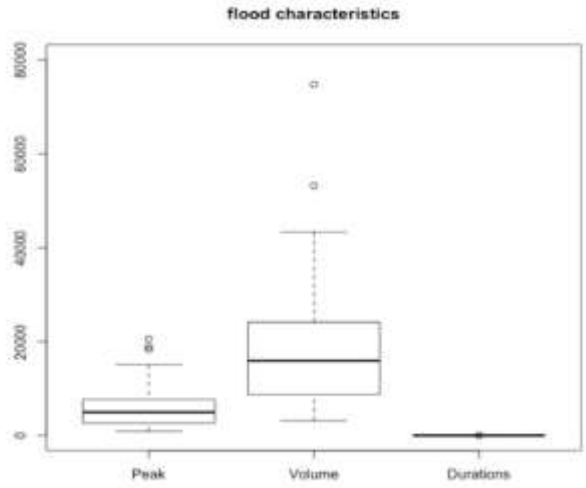
$$S = \sum_{i=1}^{n-1} \sum_{j=1+1}^{n} \text{sgn}(T_j - T_i) \tag{14}$$

$$\text{sgn}(T_j - T_i) = \begin{cases} 1 & , \text{if } T_j > T_i \\ 0 & , \text{if } T_j = T_i \\ -1 & , \text{if } T_j < T_i \end{cases} \tag{15}$$

where, T_j & T_i representing the annual value in year 'j' and 'i', respectively.

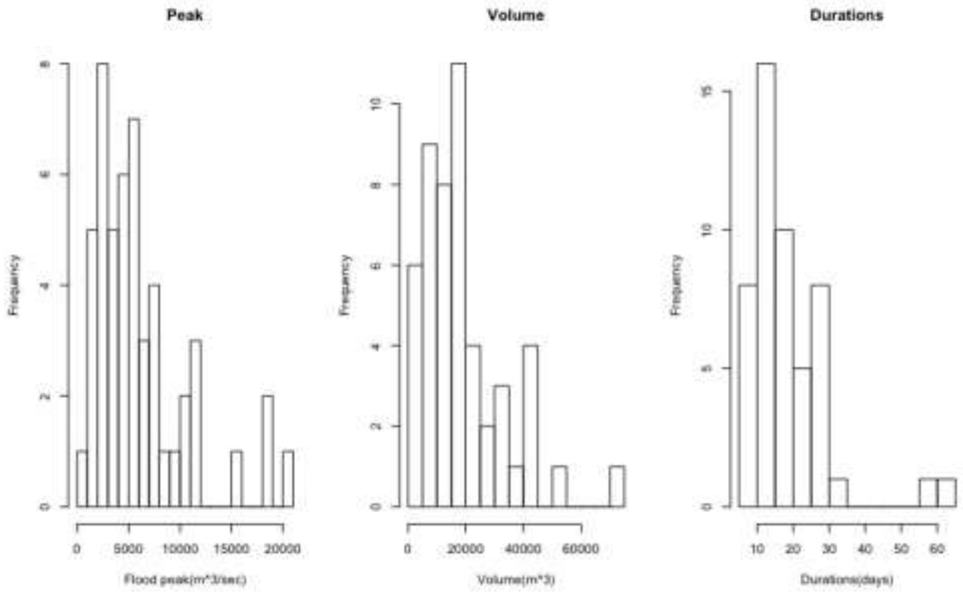


(a)

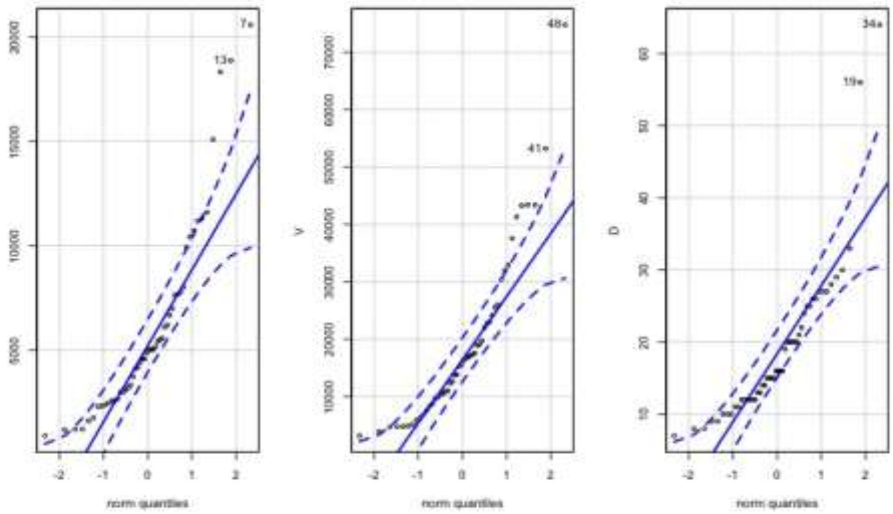


(b)

Continued-



(c)



(d)

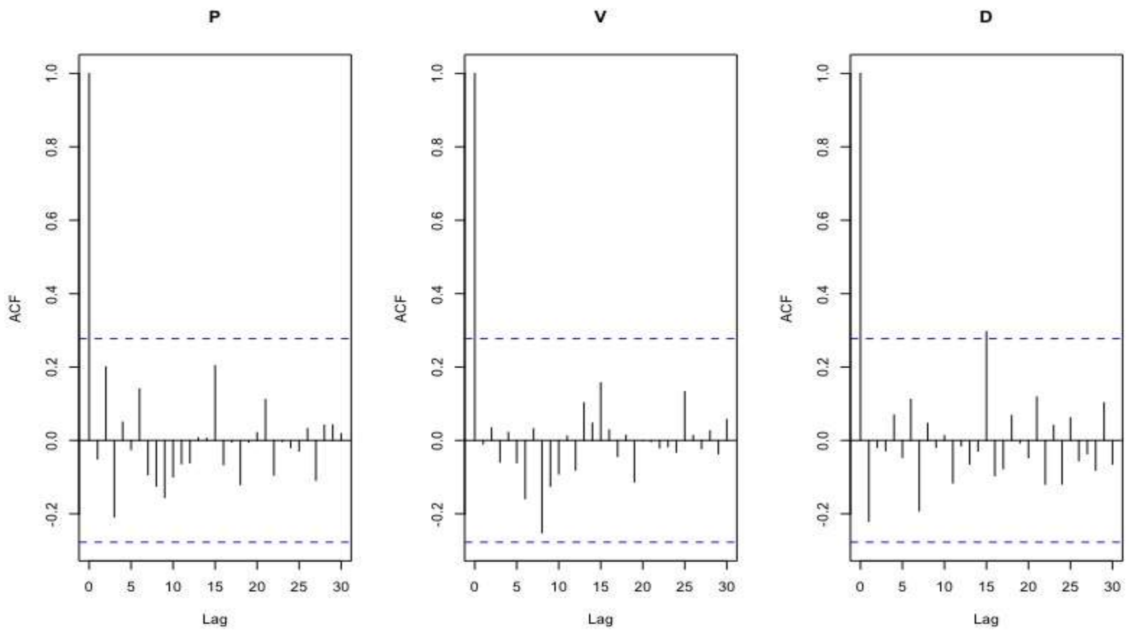
Fig. 2 Characterization of flood characteristics in the context of (a)Time series representations based on the daily streamflow discharge between 1961-2016, (b) box whisker plot, (c) Histogram plot and (d) Normal q-q plot

Under the two-tailed hypothesis attempts which usually define as
Null hypothesis (H_0) = zero trend within flood series
Alternative Hypothesis (H_1) = existence of monotonic trend within series

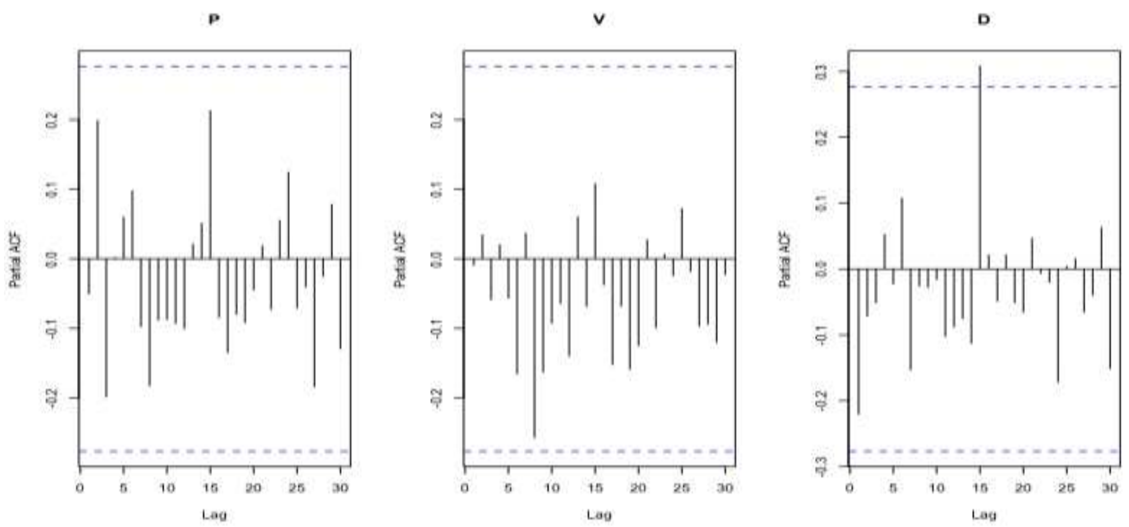
Thus, the estimated statistics from Table 4b, revealing the acceptance of null hypothesis (H0) which further pointing the existence of zero monotonic trends at the 5% or 0.05 level of significance level within each flood series. Similarly, an existence of homogenous environment between any two-given time point are also investigated for each flood characteristics through incorporating a series of test i.e., Pettit test (Pettitt 1979, Kang and Yusof 2012), Buishand test (Buishand 1982), von Neumann's test (Toreti 2011, Jaiswal 2015) and Alexanderson's SNHT based hypothesis testing (Alexandersson 1986, Toreti 2011). The p-value for each undertaken statistic is computed based on Monte-Carlo simulations at the confidence interval of 99% or 0.09. From Table 6, and Fig. 4, it could be revealing that each computed statistic is in favour of the null hypothesis (H0), which further revealing for the existence of homogeneity within the flood characteristics. In conclusion, no significant trends are detected for the flood characteristics, therefore detrend or pre-whitening procedure is not adopted (i.e., Razawi and Vogel 2018) before introducing the flood samples into univariate probability distributions framework.

Table 3 Basic descriptive summary of the flood characteristics

Descriptive measure	P(m3/sec)	V(m3)	D(days)	Percentile	P(m3/sec)	V(m3)	D(days)
Sample Size	50	50	50	Min	916.3	3182.3	7
Range	19670	71558	57	5%	1209.1	4334.7	8
Mean	6078	19122	19.04	10%	1647.1	4811.7	9.1
Variance	2.15E+07	2.14E+08	117.75	25% (Q1)	2671.8	8668.5	12
Std. Deviation	4639	14623	10.851	50% (Median)	4961	15959	16
Coef. of Variation	0.76324	0.76473	0.56993	75% (Q3)	7711.7	24476	25
Std. Error	656.05	2068.1	1.5346	90%	11584	43077	28.9
Skewness (Fisher)	1.5532	1.6392	2.2793	95%	18581	47790	43.35
Skewness (Pearson)	1.506	1.590	2.210	Max	20586	74740	64
Kurtosis (Pearson)	1.883	2.864	6.252				
Excess Kurtosis (Fisher)	2.2158	3.3029	7.0557				
Standard error of the mean	656.050	2068.071	1.535				
Lower bond on mean (95%)	4759.628	14966.495	15.956				
Upper bound on mean (95%)	7396.392	23278.381	22.124				
Standard error of the variance	4347713.616	43203375.975	23.790				



(a)



(b)

Fig. 3 Testing the existence of serial correlations within time series of flood characteristics (a) sample autocorrelation and (b) partial autocorrelation functions

4.2 Parametric estimations

In the hydro-climatic data smoothing practices, no fix distributions from any parametric family functions are assigned or in support to model any of the extreme random vectors (Karmakar and Siminovic 2008) where each probability function usually have different capability to estimates for the pre-define quantiles, especially in the tail of distribution (Lall 1995, Adamowski 1996, Yue *et al*. 1999, Sharma 2000, Kim *et al*. 2006).

Table 4 (a) Q statistics and their corresponding p-value (b) Mann Kendall (or M-K) test for identifying monotonic time trend existence within time series of flood episodes

(a)

Flood vectors	Box-Ljung test	Lag size (30)	Lag size (20)	Lag size (15)	Lag size (10)	Lag size (5)
Peak-flow	X-squared (Q-statistics)	18.869	14.88	13.338	9.7188	4.89
	df	30	20	15	10	5
	P-value	0.9427	0.7828	0.5762	0.4655	0.4295
Volume	X-squared (q-statistics)	14.724	11.981	10.683	7.4968	0.49837
	df	30	20	15	10	5
	P-value	0.9912	0.9167	0.7747	0.6779	0.9922
Duration	X-squared(Q-statistics)	23.343	15.702	13.958	6.1707	3.035
	df	30	20	15	10	5
	P-value	0.8009	0.7349	0.5287	0.8007	0.6946

(b)

M-K TEST	Peak	volume	duration
z	0.066919	0.058556	-0.40259
P-value	0.9466	0.9533	0.6872
S	9	8	-4.9000e+01
varS	1.429167e+04	1.429067e+04	1.42150e+04

The cumulative distribution function or CDFs for flood characteristics estimated by parametric procedures are fitted to the data series of peak flow, volume and duration and are compared with their empirical nonexceedance probabilities, $P[X \leq x] = F(x)$. Empirical probabilities for each flood series are estimated from the Gringorten unbiased position-plotting formulae using Eq. (1) (Gringorten 1963, Cunnane 1978, Guo 1990). An interactive set of 1-dimensional functions with varying numbers of the unknown statistical parameter (i.e., 1-parameter, 2-parameters, 3-parameters & 4-parameters) are introduced as a candidate functions in modelling univariate marginals of the flood characteristics. Table 6 listed the PDFs and CDFs of candidate parametric functions. The vector of the unknown statistical parameter of the fitted distributions for each flood series are estimated based on maximum likelihood estimation or MLE, method of moments or MOM, least square method or LS, and method of L-moments, as listed in Table 7 and estimated parameters are listed in Table 8.

Table 5 Homogeneity test statistics of flood characteristics

Test	Statistics	P	V	D	Overall Conclusion
Pettitt	K	138.000	140	128	HOMOGENOUS
	T	4	8	34	
	p-value (two-tailed)	0.715	0.744	0.555	
	Confidence interval@99% on p-value	]0.704, 0.727 [	] 0.591, 0.616	] 0.542, 0.568	
SNHT	T0	3.614	2.992	2.504	
	T	13	6	34	
	p-value (Two-tailed	0.501	0.603	0.697	HOMOGENOUS
	Confidence interval@99% on p-value	]0.488, 0.513 [	4.051	] 0.685, 0.708	
Buishand's	Q	5.956	4.015	5.273	
	T	13	6	34	
	p-value (Two-tailed	0.363	0.817	0.519	HOMOGENOUS
	Confidence interval@99% on p-value	] 0.351, 0.376 [	] 807, 0.827	] 0.506, 532 [	
Von Nuemann's	N	2.080	2.015	2.441	
	p-value (Two-tailed	0.592	0.501	0.970	
	Confidence interval@99% on p-value	] 0.580, 0.605 [	] 0.488, 0.513	] 0.965, 974 [	HOMOGENOUS

Note: p-value are computed using 10,000 Monte Carlo simulations

Table 6 Families of 1-dimensional parametric probability distribution functions

Parametric marginal distribution functions	PDF	Remarks
Exponential (1P) & 2(P)	$f(x) = \lambda e^{-\lambda x}$ & $f(x) = \lambda e^{-\lambda(x-\gamma)}$	$\lambda > 0$ - continuous inverse scale parameter; γ- continuous location parameter Domain: $y < x < +\infty$
Frechet (2P) & (3P)	$f(x) = \frac{\alpha}{\beta}\left(\frac{\beta}{x-\gamma}\right)^{\alpha+1} e^{(\beta/x-\gamma)^{\alpha}}$ & $f(x) = \frac{\alpha}{\beta}\left(\frac{\beta}{x}\right)^{\alpha+1} e^{-\left(\frac{\beta}{x}\right)^{\alpha}}$	$\alpha > 0$ (shape), $\beta > 0$(scale) , $\gamma > 0$ (location), such that, $\gamma \equiv 0$ yield 2-parameter Frechet functions Domain: $y < x < +\infty$
Gamma (2P) & (3P)	$f(x) = \frac{(x-\gamma)^{\alpha-1}}{\beta^{\alpha}\Gamma(\alpha)} e^{\frac{-(x-\gamma)}{\beta}}$ & $f(x) = \frac{x^{\alpha-1}}{\beta^{\alpha}\Gamma(\alpha)} e^{\frac{-x}{\beta}}$	$\alpha > 0, \beta > 0, \gamma > 0$ - shape, scale and locations parameter such that $\gamma \equiv 0$ yield 2-parameter gamma structure
GEV(3P)	$f(x) =$ $\frac{1}{\sigma} e^{-(1+kz)^{-1/k}(1+kz)^{-1-1/k}}$ for $k \neq 0$ $\frac{1}{\sigma} e^{(-1-e^{(-z)})}$ for $k = 0$	k, σ, μ signifies for shape, scale & their location parameter, such that, $\sigma > 0$ & $z \equiv \frac{(x-\mu)}{\sigma}$ Domain: $1 + k(x-\mu)/\sigma$ for $k \neq 0$ & $-\infty < x < +\infty$ for $k = 0$
Gen. Gamma (3P)	$f(x) = \frac{k(x)^{k\alpha-1}}{\beta^{k\alpha}\Gamma(\alpha)} e^{-(x/\beta)^{k}}$	Domain: $y \leq x < +\infty$; $k > 0$ & $\alpha > 0$ (shape), $\beta > 0$(scale), $\gamma > 0$(location)
Gumble max(2P)	$f(x) = \frac{1}{\sigma} e^{(-z-e^{-z})}$ $\{z = \frac{x-\mu}{\sigma}\}$	Domain: $-\infty < x < +\infty$ μ & $\sigma > 0$ be the scale and location parameter
Inv. Gaussian (2P)	$f(x) = \sqrt{\frac{\lambda}{2\pi x^{3}}} e^{-\frac{\lambda(x-\mu)^{2}}{2\mu^{2}(x)}}$	$\lambda > 0, \mu > 0$ (continuous parameter, for $\gamma < x < +\infty$
Johnson SB(4P)	$f(x) = \frac{\delta}{\lambda\sqrt{2\pi}z(1-z)} e^{-0.5\left(\gamma+\delta \ln\frac{z}{1-z}\right)^{2}}$	Domain: $\xi \leq x \leq \xi + \lambda$ $\gamma, \delta > 0$ (shape); $\lambda > 0$ (scale); ξ location parameter)
Log-Gamma (2P)	$f(x) = \frac{(\ln x)^{\alpha-1}}{x\beta^{\alpha}\Gamma(\alpha)} e^{-\left(\frac{\ln x}{\beta}\right)}$	Domain: $0 < x < +\infty$ $\alpha > 0$, $\beta > 0$ (shape parameter)
Log-Logistic (3P) & (2P)	$f(x) = \frac{\alpha}{\beta}\left(\frac{x-\gamma}{\beta}\right)^{\alpha-1}\left(1 + \left(\frac{x-\gamma}{\beta}\right)^{\alpha}\right)^{-2}$ & $f(x) = \frac{\alpha}{\beta}\left(\frac{x}{\beta}\right)^{\alpha-1}\left(1 + \left(\frac{x}{\beta}\right)^{\alpha}\right)^{-2}$	Domain: $\gamma < x < +\infty$ $\alpha > 0$ (shape); $\beta > 0$(scale)& $\gamma > 0$ (location)
Lognormal (3P) & (2P)	$f(x) = \frac{e^{-0.5\left(\frac{\ln(x-\gamma)-\mu}{\sigma}\right)^{2}}}{(x-\gamma)\sigma\sqrt{2\pi}}$ & $f(x) = \frac{e^{-0.5\left(\frac{\ln(x)-\mu}{\sigma}\right)^{2}}}{(x)\sigma\sqrt{2\pi}}$	$\gamma < x < +\infty$; $\sigma > 0$ (shape parameter); γ (location parameter); μ (scale parameter
Log-Pearson (3P)	$f(x) = \frac{1}{x\lvert\beta\rvert\Gamma(\alpha)}\left(\frac{\ln(x)-\gamma}{\beta}\right)^{\alpha-1} e^{-\left(\frac{\ln(x)-\gamma}{\beta}\right)}$	$0 < x < e^{\gamma}$ for $\beta < 0$ & $e^{\gamma} \leq x < +\infty$ for $\beta > 0$ $\alpha > 0, \beta \neq 0, \gamma$ (continuous parameter)
Weibull (2P) & (3p)	$f(x) = \frac{\alpha}{\beta}\left(\frac{x-\gamma}{\beta}\right)^{\alpha-1} e^{-\left(\frac{x-\gamma}{\beta}\right)^{\alpha}}$ & $f(x) = \frac{\alpha}{\beta}\left(\frac{x}{\beta}\right)^{\alpha-1} e^{-\left(\frac{x}{\beta}\right)^{\alpha}}$	Domain: $\gamma \leq x < +\infty$ $\alpha > 0$ (shape), $\beta > 0$ (scale) & γ (location parameter)

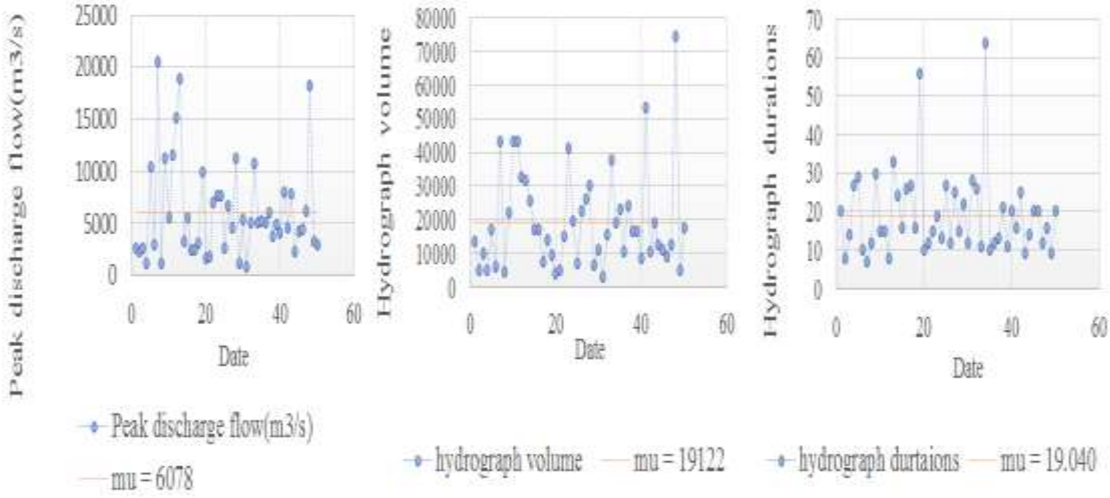


Fig. 4 Homogeneity level within flood distributions series

Table 7 Parameter estimation procedure for marginal distributions

Functions	Parameters estimation method
EXPONENTIAL(1P)	Method of moments
EXPONENTIAL(2P)	Maximum likelihood method
FRECHET(2P)	Least square method
FRECHET(3P)	Maximum likelihood method
GAMMA(2P)	Method of moments
GAMMA(3P)	Maximum likelihood method
GEV(3P)	Method of L-moments
LOG-GAMMA(2P)	Method of moments
LOG-LOGISTIC(2P)	Least squares method
LOG-LOGISTIC(3P)	Maximum likelihood method
LOG-PEARSON(3P)	Method of moments
GUMBEL MAX (2P)	Method of moments
LOG-NORMAL(2P)	Maximum likelihood method
LOG-NORMAL(3P)	Maximum likelihood method
WEIBULL(2P)	Least squares method
WEIBULL(3P)	Maximum likelihood method
INV. GAUSSIAN (2P)	Method of moments
JOHNSON SB (4P)	Method of moments
GEN. GAMMA (3P)	Maximum likelihood method

Table 8 Estimated parameters of marginal distribution of flood characteristics

Parametric Functions	Flood Peak (P)	Flood Volume (V)	Flood Durations (D)
EXPONENTIAL(1P)	l=1.6453E-4	l=5.2295E-5	l=0.05252
EXPONENTIAL(2P)	l=1.9373E-4 g=916.3	l=6.2735E-5 g=3182.3	l=0.08306 g=7.0
FRECHET(2P)	a=1.576 b=3207.5	a=1.5703 b=10017.0	a=2.6001 b=13.304
FRECHET(3P)	a=3.1238 b=7764.6 g=-4076.2	a=2.8923 b=22571.0 g=-11129.0	a=3.6283 b=20.616 g=-6.7647
GAMMA(2P)	a=1.7166 b=3540.6	a=1.71 b=11183.0	a=3.0786 b=6.1845
GAMMA(3P)	a=1.2106 b=4290.0 g=884.47	a=1.0848 b=14723.0 g=3150.8	a=1.4696 b=8.3319 g=6.7958
GEV(3P)	k=0.22596 s=2683.6 m=3765.6	k=0.20446 s=8736.0 m=11890.0	k=0.20682 s=6.0766 m=13.98
LOG-GAMMA(2P)	a=129.15 b=0.06544	a=164.32 b=0.05839	a=35.165 b=0.08037
LOG-LOGISTIC(2P)	a=2.2801 b=4541.7	a=2.2731 b=14202.0	a=3.6928 b=16.426
LOG-LOGISTIC(3P)	a=2.0775 b=4217.5 g=423.18	a=1.8662 b=12305.0 g=2091.6	a=2.3027 b=10.393 g=5.654
LOG-PEARSON(3P)	a=663.54 b=-0.02887 g=27.608	a=1787.0 b=-0.01771 g=41.234	a=14.523 b=0.12506 g=1.0099
GUMBEL MAX (2P)	s=3617.0 m=3990.2	s=11402.0 m=12541.0	s=8.4608 m=14.156
LOG-NORMAL(2P)	s=0.7362 m=8.4513	s=0.74093 m=9.5943	s=0.47178 m=2.826
LOG-NORMAL(3P)	s=0.75437 m=8.4267 g=85.951	s=0.8237 m=9.4858 g=1115.2	s=0.69194 m=2.413 g=4.8982
WEIBULL(2P)	a=1.599 b=6398.7	a=1.5993 b=20008.0	a=2.5437 b=20.375
WEIBULL(3P)	a=1.1175 b=5389.8 g=899.42	a=1.0689 b=16369.0 g=3155.6	a=1.1951 b=12.878 g=6.9279
INV. GAUSSIAN (2P)	l=10434.0 m=6078.0	l=32699.0 m=19122.0	l=58.617 m=19.04
JOHNSON SB (4P)	g=1.5161 d=0.74495 l=27319.0 x=1304.2	g=2.2027 d=1.0357 l=1.3052E+5 x=961.8	g=2.5314 d=0.92215 l=118.81 x=8.2791
GEN. GAMMA (3P)	k=1.054 a=1.8127 b=3540.6	k=1.0521 a=1.8019 b=11183.0	k=1.0877 a=3.4664 b=6.1845

4.3 Models compatibility testing based on Goodness-of-fit (or GOF) statistics

The theoretical cumulative density for each flood distribution series of P, V and D, are estimated by parametric procedure and compared against the empirical non-exceedance probabilities for the outlining data reproducing and fitness consistency with observational samples. The estimators such as maximum likelihood, method of moments, least squares method and method of L-moment are employed for estimating the vectors of model parameters. Model compatibility investigations are conducted based on both the analytical as well as graphical visual inspections.

In the first stage for investigating model compatibility with observational samples, EDF based distance statistics i.e., Anderson-Darling (AD_n) and Kolmogorov- Smirnov ($K\text{-}S_n$) test are estimated for each model fitted to flood characteristics where K-S statistics investigating the largest vertical gaps between cumulative empirical and theoretical probabilities based on Eqs. (2) and (3). On another side, the quadratic class of EDF or A-D statistics examines whether the random samples come from some specified distributions based on Eq. (4). Table 9(a) listed the estimated K-S & A-D values and revealed for the satisfactory performance for most of the candidate functions for each flood characteristics.

From section 2.2, it is already mentioned that if the estimated p-value of the fitted distribution in the K-S test is above 0.05 (i.e., p-value $>\mathrm{p_{critical}}(0.05)$), which often indicated for considerably better performance, otherwise liable to reject. Also, if the D-statistics of the estimated K-S value is below the critical value (i.e., $d_{model} < d_{critical}$) at a pre-defined significance level then it often indicating for the positive decision in favour of good agreement between fitted and empirical distributions otherwise performance liable for the rejection (Yue and Rasmussen 2002, O'Connor and Kleyner 2012, Reddy and Ganguli 2012). Similarly, if the estimated A-D values are below their critical level at the significance level $'\alpha'$ then it often indicated for better model performance with the observed samples. Also revealed from section 2.2 that the K-S statistics exhibited relatively flat in their tail distributions for both the empirical and fitted probabilities and quite less sensitive near the distribution tails in relative with the center of distributions in compare with A-D statistics which facilitates an extra weight towards the tail of probability distributions and quite be more consistent in case of large extreme modelling. K-S & A-D test statistics of Table 9(a) revealing that the performance of Log-Pearson-3P and Lognormal-2P distribution are much satisfactory for flood peak flow samples in comparison with other candidate functions. Such that the K-S value (${}^{KS}{}_n$(d- max) = 0.05178 with p-value 0.99833) for Log-Pearson-3P distribution and (${}^{KS}{}_n$ (d- max) = 0.05293 with p-value 0.9977) for Lognormal-2P distribution and where, the D-critical value for K-S test for sample size 50 is 0.1884 at 5% significance level. Similarly, the A-D value (${}^{AD}{}_n$(d-max) = 0.18403) for Log-Pearson-3P and (0.19412) for Lognormal-2P distribution where, the D-critical for A-D test for sample size 50 is 2.5018 at 5% significance level. By comparing the LP-3P and LN-2P distribution it must be revealed that LP-3P distribution poses much consistent and quite better performance than LN-2P for capturing the peak flow series.

In the second stage, information criteria-based statistics are incorporated to find out the acceptability of the distribution functions pointed on the basis of K-S and A-D test as presented in Table 9(a). The relative performance measuring between the LP-3P and LN-2P distribution is further analyzed based on AIC, BIC & HQC, which usually highlights the trade-off relationship between model bias or uncertainty with the number of fitted parameters and estimated based on Eqs. (6)-(9). Minimum the value of AIC, BIC & HQC could indicate for the best-fitted model. From Table 9(b) it is indicating that the AIC, BIC and HQC values are at minimum for

Table 9 Comparison of flood variables for different statistical distributions based on (a) K-S & A-D distance (or d_{max} criteria) (b) Information criteria statistics (i.e., AIC, BIC & HQC)

(a)

	peak			volume			durations		
Parametric candidate's functions	P-value	KSn (d- max)	and (d-max)	P-value	KSn (d- max)	ADn (d-max)	P-value	KSn (d- max)	ADn (d-max)
FRETCH(2P)	0.32428	0.13147	1.0751	0.28744	0.1359	1.1173	0.36268	0.1272	0.58456
FRETCH(3P)	0.99732	0.05351	0.21153	0.96141	0.06828	0.3033	0.68038	0.09849	0.36878
GEV(3P)	0.99655	0.05451	0.21667	**0.99931**	**0.04897**	0.24945	0.82259	0.086	0.35244
LOG-GAMMA(2P)	0.97557	0.06486	0.22646	0.95247	0.07004	0.26683	0.85726	0.08255	0.3451
LOG-LOGISTIC(2P)	0.96909	0.06655	0.24216	0.88242	0.07982	0.32827	0.73162	0.09416	0.49615
LOG-LOGISTIC(3P)	0.9968	0.05421	0.23129	0.9471	0.07101	0.36216	0.6921	0.09751	0.38531
LOG-PEARSON(3P)	**0.99833**	**0.05178**	**0.18403**	0.9919	0.05836	0.21229	0.80879	0.08731	0.33025
GUMBEL(2P)	0.4966	0.11417	0.90135	0.62555	0.10307	0.74771	0.51472	0.11255	1.0798
GAMMA(2P)	0.81376	0.08684	0.44712	0.94562	0.07126	0.34627	0.54764	0.10968	1.1617
GAMMA(3P)	0.8802	0.08007	0.26953	0.98701	0.06089	0.21109	0.89254	0.07865	0.37708
EXPONENTIAL(1P)	0.03558	0.19698	2.3258	0.0784	0.17643	2.1603	0.0000381	0.32306	6.9597
EXPONENTIAL (2P)	0.45829	0.11768	2.3535	0.9265	0.07425	2.094	0.25721	0.13985	1.661
LOG-NORMAL(2P)	0.9977	0.05293	0.19412	0.98539	0.06157	0.2338	0.60127	0.10511	0.4602
Log-normal (3p)	0.99466	0.05638	0.20029	0.93057	0.07365	0.28195	0.79396	0.08867	0.33032
BURR(3P)	0.99539	0.05573	0.20556	0.99847	0.05147	0.23983	0.564	0.10827	0.44764
WEIBULL(2P)	0.81311	0.0869	0.73212	0.89172	0.07875	0.63575	0.23928	0.14235	1.5472
WEIBULL(3P)	0.86868	0.08134	0.28905	0.99653	0.05454	0.194	0.88156	0.07992	0.45987
INV.GAUSSIAN(2P)	0.98175	0.06293	0.38095	0.81919	0.08633	0.48954	0.87056	0.08114	0.60496
GEN. GAMMA(3P)	0.66896	0.09944	0.45939	0.89941	0.07782	0.36811	0.28097	0.13672	0.91168
GEN. GAMMA(4P)	0.42425	0.12092	2.4528	0.95878	0.06883	2.0178	**0.89623**	**0.07821**	**0.3294**
INV. GAUSSIAN (3P)	0.99323	0.05748	0.2024	0.92454	0.07453	0.27218	0.84885	0.08341	0.33031
PEARSON 5 (3P)	0.99812	0.05219	0.20465	0.96087	0.0684	0.29419	0.7184	0.09529	0.35172
PEARSON 6(3P)	0.84394	0.08391	0.31412	0.61982	0.10355	0.56175	0.82964	0.08532	0.35655
Pearson 6(4p)	0.99812	0.05219	0.20465	7.0585 E-07	0.37656	11.459	0.71263	0.09578	0.35114
JOHNSON SB	0.84788	0.84788	14.822	0.99811	0.05222	**0.17314**	0.56249	0.1084	11.874

[[Note: D-critical for A-D and K-S test statistics @ significance level of 5% or 0.05 for samples size 50 are 2.5018 & 0.1884]

(b)

	Peak		peak	volume			Durations		
Functions	AIC	BIC	HQC	AIC	BIC	HQC	AIC	BIC	HQC
FRETCH(2P)	-284.118	-280.294	-282.66	-274.569	-270.745	-273.11	-307.04	-303.22	-305.588
FRETCH(3P)	-371.057	-365.32	-368.87	-353.796	-348.06	-351.61	-331.1	-325.361	-328.912
GEV(3P)	-374.335	-368.599	-372.15	-268.985	-263.249	-266.8	-336.32	-330.583	-334.135
LOG-GAMMA(2P)	-370.146	-366.322	-368.69	-359.914	-356.09	-358.46	-340.53	-336.709	-339.077
LOG-LOGISTIC(2P)	-360.392	-356.568	-358.94	-294.927	-291.103	-293.47	-321.32	-317.493	-319.861
LOG-LOGISTIC(3P)	-371.549	-365.813	-369.36	-350.302	-344.566	-348.12	-330.41	-324.673	-328.225
LOG-PEARSON(3P)	-379.039	-373.303	-376.85	-375.318	-369.582	-373.13	-339.08	-333.347	-336.899
GUMBEL(2P)	-294.924	-291.1	-291.43	-308.477	-304.653	-307.02	-293.6	-289.775	-292.143
GAMMA(2P)	-335.861	-332.037	-334.4	-360.025	-356.201	-358.57	-260.55	-256.722	-259.089
GAMMA(3P)	-216.301	-210.565	-214.12	-210.107	-204.371	-207.92	-343.62	-337.886	-341.438
EXPONENTIAL(1P)	-242.501	-240.589	-164.07	-248.425	-246.513	-247.7	-179.33	-177.416	-178.6

Continued-

EXPONENTIAL (2P)	-249.916	-246.092	-248.46	-345.901	-342.076	-344.44	-280.58	-276.758	-279.126
LOG-NORMAL(2P)	-379.344	-375.52	-377.89	-371.028	-367.204	-369.57	-327.46	-323.633	-326.001
LOG-NORMAL (3P)	-285.412	-279.676	-283.23	-352.906	-347.17	-350.72	-340.76	-335.026	-338.578
BURR(3P)	-373.626	-367.89	-371.44	-374.982	-369.246	-372.8	-320.24	-314.505	-318.057
WEIBULL(2P)	-329.681	-325.857	-328.23	-342.868	-339.044	-341.41	-292.91	-289.085	-291.453
WEIBULL(3P)	-200.361	-194.625	-200.9	-376.477	-370.741	-374.29	-200.15	-194.418	-197.97
INV.GAUSSIAN(2P)	-362.489	-358.665	-361.03	-344.722	-340.898	-343.27	-325.76	-321.938	-324.306
GEN. GAMMA(3P)	-321.553	-315.817	-319.37	-338.918	-333.182	-336.73	-290.95	-285.21	-291.856
GEN. GAMMA(4P)	-290.892	-283.325	-294.08	-360.62	-352.971	-357.71	-343.42	-335.769	-340.504
INV. GAUSSIAN (3P)	-199.634	-193.898	-197.45	-352.161	-346.425	-349.98	-343.74	-338.007	-341.559
PEARSON 5 (3P)	-371.12	-365.384	-367.9	-354.967	-349.231	-352.78	-334.53	-328.796	-332.348
PEARSON 6(3P)	-337.894	-332.158	-335.71	-237.013	-231.277	-234.83	-333.99	-328.253	-331.805
PEARSON 6(4p)	-371.698	-364.05	-368.79	-143.365	-135.717	-140.45	-332.8	-325.154	-329.89
JOHNSON SB (4P)	-340.899	-333.251	-337.99	-381.821	-374.173	-378.91	-223.65	-216.006	-220.742

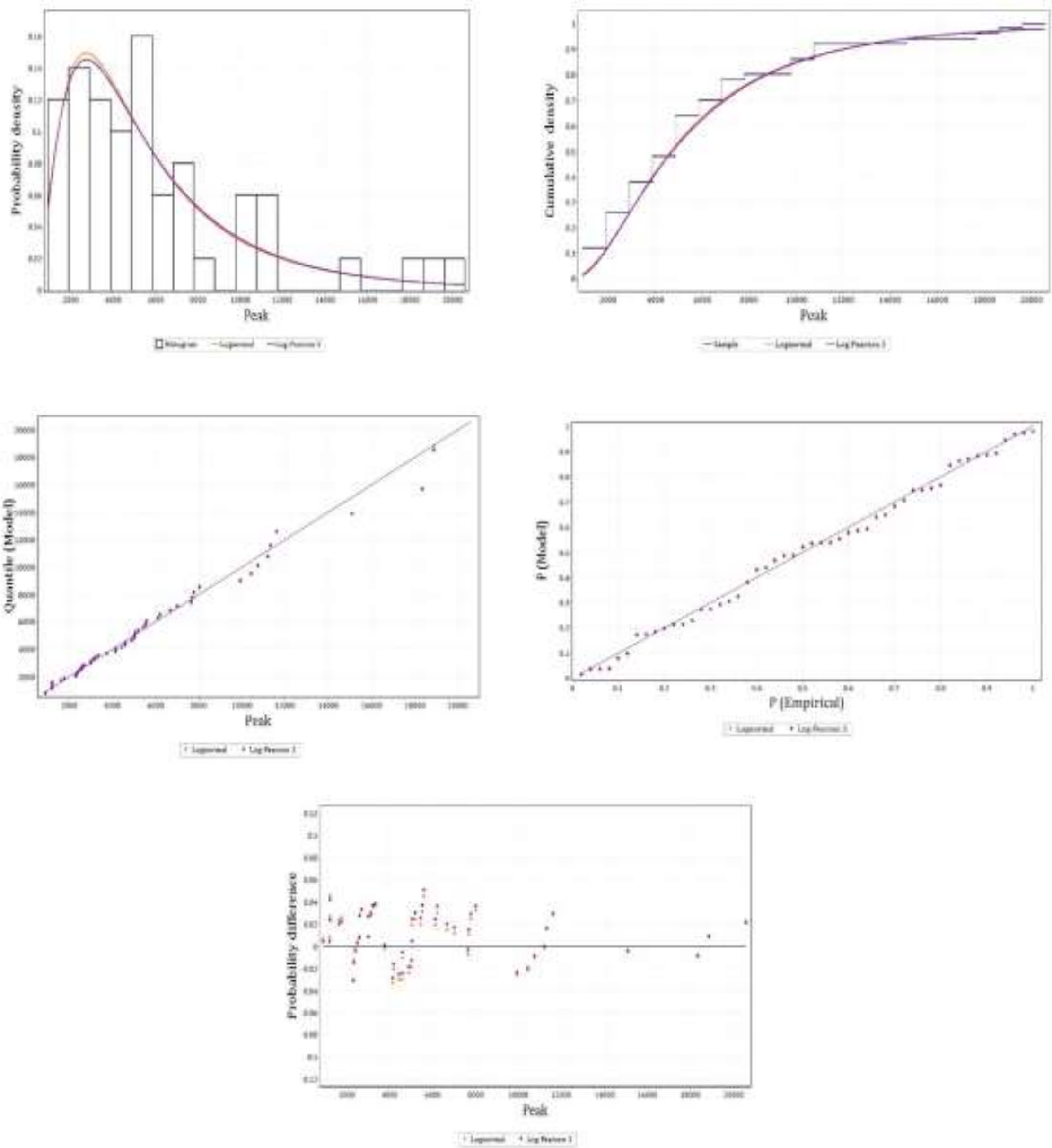
Log-normal-2P (AIC= -379.344, BIC= -375.52, HQC= -377.89) in comparison with Log-Pearson-3P for flood peak series (AIC=-379.039, BIC= -373.303, HQC= -376.85).

In the third stage, the error indices statistics are estimated using the equations which are listed in Table 2 for the distributions selected in the first and second stage. It is seen in the Table 10 that MSE, RMSE, MAE and RSR values are at minimum for LP-3P distribution (MSE=0.0004525, RMSE=0.2127141, MAE=0.01840788, RSR=0.07387798) in comparison with LN-2P distribution (MSE=0.0004681, RMSE=0.02163505, MAE=0.018146725, RSR=0.075140942). Minimum the value of MSE, RMSE, MAE and RSR must be indicated for a better fit. Similarly, the NSE statistics for LP-3P distribution is higher (NSE= 0.994542) than LN-2P distribution (NSE=0.994354) such that value closer to ‘1’ often indicates for better performance.

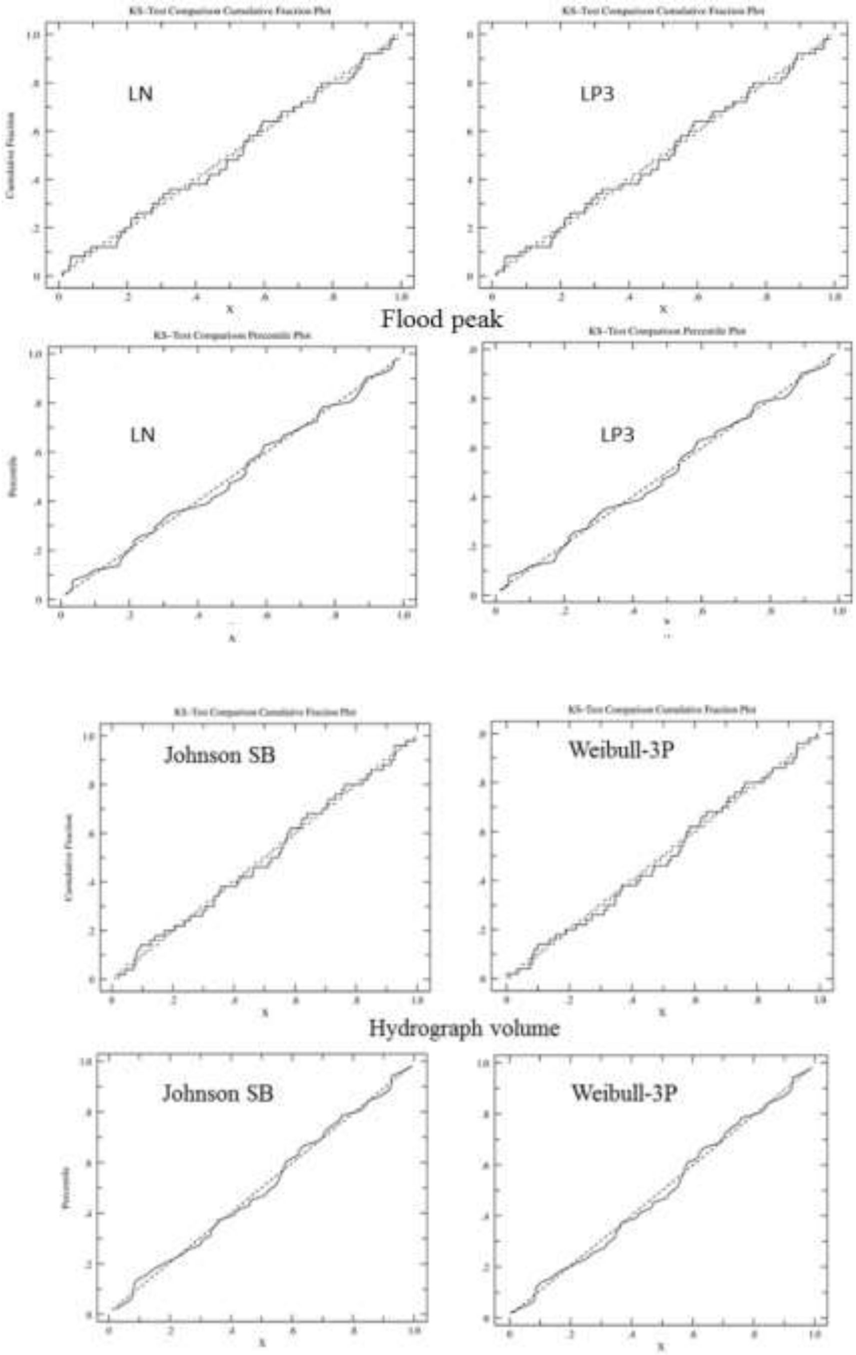
Similarly, based on K-S and A-D test statistics for the hydrograph volume series, Johnson SB-4P and Weibull-3P distribution are selected such that the K-S value ($^{KS}{}_{n}$(d- max) = 0.0522 with p-value 0.99811) for Johnson SB-4P distribution and ($^{KS}{}_{n}$(d- max) = 0.05454 with p-value 0.99653) for Weibull-3P distribution. The A-D value ($^{AD}{}_{n}$(d-max) = 0.17314) for Johnson SB-4P and (0.194) for Weibull- 3P distribution. The performance of Johnson SB-4P is much satisfactory than the Weibull-3P distributions based on both K-S and A-D measures. Similarly, based on the information criteria statistics for volume series, AIC, BIC and HQC values are at a minimum for Johnson

SB-4P distribution (AIC=-381.821, BIC=-374.173, HQC=-378.91) in comparison with Weibull-3P distribution (AIC=-376.477, BIC=-370.741, HQC=-374.29). Also, based on the error indices statistics of Table 10 for hydrograph volume series, it is revealing that the performance of Johnson SB-4P distribution (MSE=0.000412, RMSE=0.02027810, MAE=0.01867, RSR=0.07042) dominating and much consistent over the Weibull-3P distribution (MSE=0.0004762, RMSE=0.02182, MAE=0.01867, RSR=0.07579). Similarly, the NSE statistics of Johnson SB-4P (NSE=0.99450) is much closer to unity than Weibull-3P distribution and thus in support of Johnson SB-4P function.

At first, the Inverse Gaussian-3P, Gamma-3P and Generalized Gamma-4P are selected for the duration samples based on K-S and A-D test statistics such that the K-S value ($^{KS}{}_{n}$(d- max) = 0.08341 with p-value 0.84885) for Inverse Gaussian-3P distribution, ($^{KS}{}_{n}$(d- max) = 0.07865 with p-value 0.89254) and ($^{KS}_{n}$(d- max) = 0.07821 with p-value 0.89623) for Generalized Gamma-4P. Similarly, the A-D value ($^{AD}_{n}$(d-max) = 0.33031) for Inverse Gaussian-3P distribution, ($^{AD}_{n}$(d-max) = 0.37708) for Gamma- 3P distribution and ($^{AD}_{n}$(d-max) = 0.3294) for Generalized Gamma-4P function.

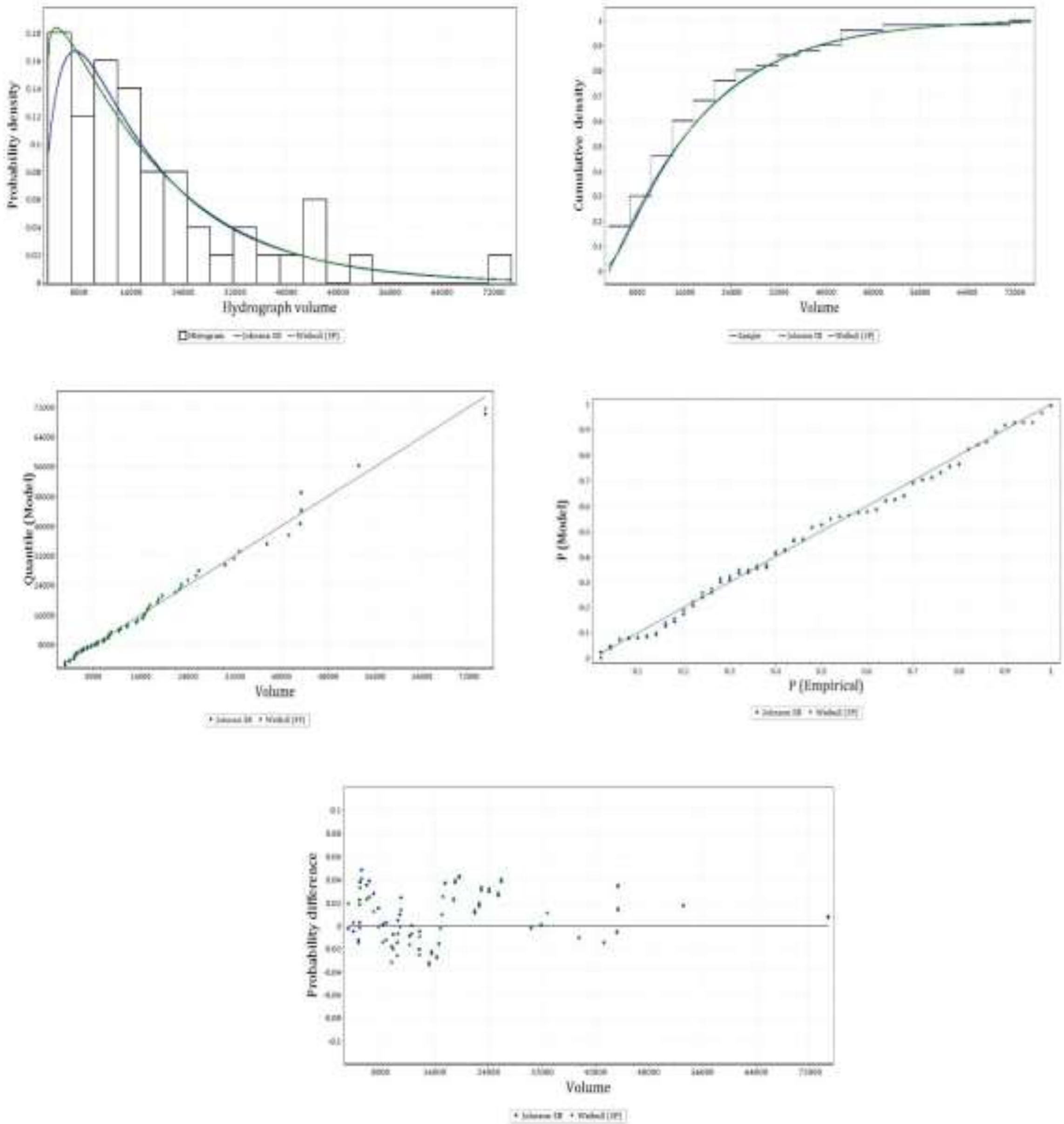


Continued-


LN
LP3
Flood peak
LN
LP3
Johnson SB
Weibull-3P
Hydrograph volume
Johnson SB
Weibull-3P

(a)

Continued-


Probability density
Hydrograph volume
Cumulative density
Volume
Quantile (Model)
Volume
P (Model)
P (Empirical)
Probability difference
Volume

(b)
Continued-

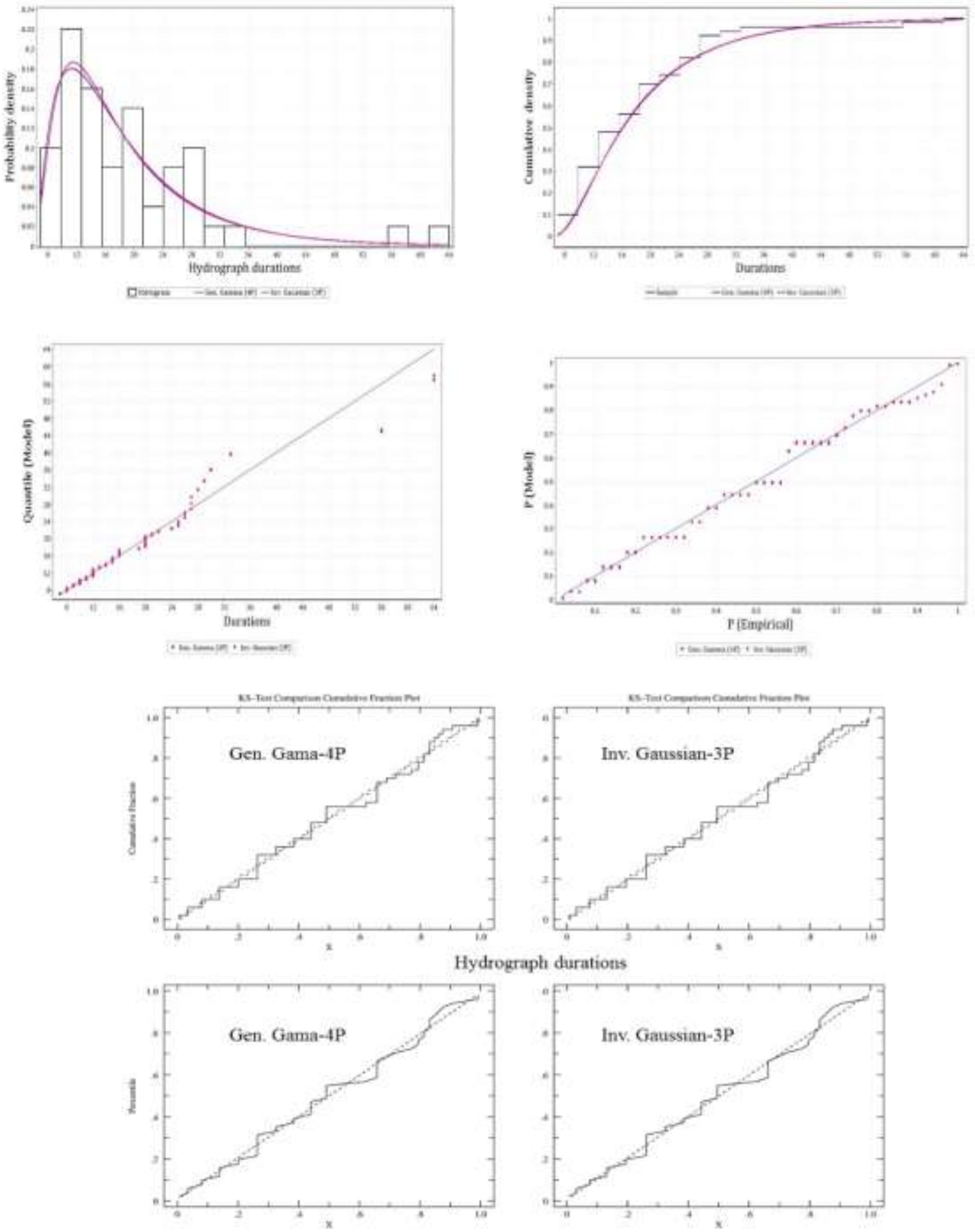


Fig. 5 Model compatibility investigations in the context of pdf, cdf, p-p plot, q-q plot, probability difference plot, K-S test comparison cumulative plot and K-S test comparison percentile plot (a) between Lognormal-2P & Log-Pearsson-3P for flood peak, (b) between Johnson SB & Weibull-3P for volume and (c) and between Generalized Gamma-4P & Inverse Gaussian-3P for duration series

Table 10 Error indices statistics of flood characteristics for different statistical models

Flood vectors	functions	MSE	RMSE	NSE	MAE	RSR
Peak	LP-3P	0.0004525	0.021271419	0.994542	0.01840788	0.07387798
	LN-2P	0.0004681	0.021635059	0.994354	0.018146725	0.075140942
Volume	Weibull-3P	0.0004762	0.02182	0.99364	0.01867	0.07579
	Johnson SB (4P)	0.0004112	0.020278103	0.994507	0.017278038	0.07042
Durations	Gen. Gamma (4P)	0.0008864	0.029772541	0.98877	0.023963239	0.103403315
	Inv. Gaussian (3P)	0.0009166	0.030275057	0.988944	0.024196839	0.105148609
	Gamma (3P)	0.000918804	0.030312	0.988917	0.024593	0.105276

Based on information criteria-based statistics of Table 10, it is pointing the Generalized Gamma-4P (AIC=-343.42, BIC=-335.769, HQC=-340.504) has the minimum values than Inverse Gaussian-3P and Gamma-3P distribution. Similarly, the error indices statistics of Table 10 also indicating and in favour of Generalized Gamma-4P distribution (MSE= 0.0008864, RMSE=0.02977254, MAE=0.023963239, RSR=0.103403315), but NSE statistics for all the three distribution are almost identical.

A qualitative based approach via the graphical investigations are performed for each flood distributions series based on probability density plot, cumulative density plot, p-p plot, q-q plot, probability difference plot, K-S test comparison cumulative fraction plot and K-S test comparison percentile plot as illustrated from Fig. 5(a)-5(c). Based on the probability difference plot and p-p plot for flood peak variables are quite more consistent with the LP-3P distribution in comparison with LN-2P function such that LP-3P exhibited quite better consistencies near the tail of distributions. K-S test comparison cumulative fraction and percentile plot is also in favor of LP-3P, which performing quite better than LN-2P distribution. Similarly, for the hydrograph volume and duration series, the graphical approaches are also in support of the distribution selected based on qualitative approaches such as Johnson SB-4P (for volume series) and Generalized Gamma-4P (for duration series) distribution. Overall, based on analytical and graphical fitness measures, pointing towards flood peak samples seem to follow heavy-tailed Log-Pearson-3P distribution, hydrograph volume seems to follow Johnson SB-4P distribution and the hydrograph duration seems to follow Generalized Gamma-4P distribution to constructing the univariate flood marginals.

5. Conclusions

The higher degree of randomness and complex dependency among the intercorrelated flood characteristics, such as peak, volume & durations, often demanding for multivariate statistical treatment for accounting flow exceedance probabilities or design variable quantiles under the different notations of return periods. Actually, trivariate behaviour of the flood characteristics often limits the reliability of univariate continuous frequency relationship thus could be demanding for joint probability assessments of various possible combinations among flood characteristics such as between peak-volume/or volume-duration/or peak-duration. Flood become one of the most intensive and critical hydro-climatic issues over the Kelantan River basin Malaysia, more likely during the wet monsoon circulations. Multivariate hydrologic risk assessments could be an essential & practical demand for tackling several basin perspective water-related queries in this river basin. From the past decades, such correlation based stochastically hydro-climatic generations become much flexible and robust in the light of copulas multivariate framework in compare with the traditional multivariate functions which already reviewed in section 1. An interactive set of copulas often employed in modelling of extreme samples based on the bivariate or trivariate joint distribution analysis. Actually, copulas-based methodology segregated univariate marginal modelling independently from their joint dependence constructions thus often facilitating to select most justifiable or best-fitted probability density functions.

Multivariate distribution analysis often demands the selection of most justifiable probability distributions for defining the univariate flood marginals before introducing into a joint framework. Therefore, in the present study, an extensive selection of marginal distribution for flood characteristics is performed by parametric estimation procedures. Actually, in hydrologic data modelling, no universally accepted models are assigned from any literature or in favour of any probability distribution functions which solely a trial and error procedure (Adamowaski 1985, 1989, Silverman 1986). Also, several distributions often would fit the data equally well but, each would give different estimates of a given quantile especially in the tails of the distribution, which is solely based on the goodness-of-fit procedure to visualize the compatibility of the fitted distributions. Actually, the model performance evaluations and the selection of best-fitted distributions often demand many precise investigations otherwise inconsistencies might reveal for uncertainty. Also, the strength & weakness of different fitness statistics usually vary and having different extent during demonstrating gaps and dispensary among fitted distributions.

Distinct varieties of 1-parameter, 2-parameters, 3-parameters and 4-parameters parametric family functions are employed and tested for event-based i.e., block (annual) maxima flood characteristics derived from the daily basis streamflow discharge records collected at the Gulliemard Bridge gauge station for Kelantan River basin in Malaysia. Vector of the unknown statistical parameters of fitted distributions for each flood series are estimated based on MLE, MOM, least squares method and L-moment density estimators. Different analytical based goodness-of-fit measures such as based on K-S and A-D distance criteria statistics, information criteria statistics (i.e.., AIC, BIC & HQC) and error indices statistics (i.e., MSE, RMSE, MAE, NSE & RSR) are incorporated for the parametric probability distributions for selecting the possible marginal structure of peak flow, volume and duration series based on the comparative assessments between their empirical cumulative and theoretical probabilities. In the first stage, K-S and A-D test value are used to select the closest distributions for each flood characteristics. In the second & third fitness stage, information criteria statistics based on AIC, BIC and HQC test values as well as error indices measures such as MSE, RMSE, MAE, RSR & NSE are also estimated and compared for each parametric distribution to find out the acceptability of the distribution functions selected on the basis of first stage fitness measures (i.e., based on K-S and A-D test statistics). Overall, after summarizing all the analytical testing measures, it is pointing towards the Log-Pearson (3P) distributions for flood peak discharge flow, Johnson SB (4P) for hydrograph volume and Gen. Gamma (4P) for modelling durations series. Several graphically based visual inspections are also carried out based on PDFs, CDFs, P-P plot, Q-Q plot, probability difference plot, K-S test comparison cumulative plot and K-S test comparison percentile plot, which also in favour of the analytical based judgements. Finally, it is concluded that the flood peak flow series are best fit with the Log-Pearson-3P distribution, hydrograph volume series with the Johnson SB-4P distribution and the durations series are best fit with the Generalized Gamma-4P distribution.

Unless the above univariate constructions are demonstrated in the light of parametric distributions framework, but no one could deny from their limitations such as prior to the distribution information before fitting random observations. As, no universally accepted distributions are assigned from any kinds of literature or either in favour of any probability density functions for tackling any hydro-climatic problems (Adamowaski 1985, Silverman 1986, Yue *et al.* 1999). Also, based on the histogram of the flood vectors from Fig. 5, clearly pointing towards the bimodal distribution behaviour for the peak and volume samples and thus under such distribution environment, nonparametric based smoothing would be much consistency and stable (Sharma 2000, Kim and Yoo 2003). Few attempts adapted the non-parametric i.e., kernel-based data smoothing procedure for solving several extreme consequences and their conclusion often revealing for much practical efforts in defining their marginal behaviour & design estimations in the context of lack of any prior density assumptions over the parametrical density framework (Lall and Moon 1993, Lall 1995, Adamowaski 1996, Bowman and Azzalini 1997, Kim *et al*. 2006). Therefore, future motivations will be extended to model such extreme consequences under the non-parametric concept in order to reproduce the random attributes represented by distributed the flood samples much smoothly.

References

Adamowski, K. (1985), "Nonparametric kernel estimation of flood frequencies", *Water Resour. Res*., **21**(11), 1885-1890.

Abdulkareem, J.H. and Sulaiman, W.N.A. (2015). "*Trend Analysis of Precipitation Data in Flood Source Areas of Kelantan River Basin, Malaysia*", *Proceedings of th*e 3rd International Conference in Water Resources, ICWR-2015.

Alam, A., Bhat, M.S., Hakeem, F., Ahmad, B., Ahmad, S. and Sheikh, A.H. (2018), "Flood risk assessment of Srinagar city in Jammu and Kashmir, India". *Int. J. Disaster Resilience Built Environ*., **2**, 9. https://doi.org/10.1108/IJDRBE-02-2017-0012.

Arshad, M., Rasool, M.T. and Ahmad, M.I. (2003), "Anderson Darling and modified Anderson Darling Tests for generalized Pareto distribution", *Pakistan J. Appl. Sci*., **3**(2), 85-88.

Anderson, T.W. and Darling, D.A. (1954), "A test of goodness of fit", *J. Am. Stat. Assoc*., **49**(268), 765-769.

Adamowski, K. (1989), "A monte Carlo comparison of parametric and nonparametric estimations of flood frequencies", *J. Hydrol*., **108**, 295-308.

Alamgir, M., Ismail, T. and Noor, M. (2018). "Bivariate frequency analysis of flood variables using copula in Kelantan River Basin", *Malaysian J. Civil Eng*., **30**(3), 395-404.

Arnold, J.G. and Allen, P.M. (1999), "Automated methods for estimating baseflow and ground water recharge from streamflow records", *J. Am. Water Resour. Assoc*., **35**, 411-424.

Alexandersson, H. (1986), "A homogeneity test applied to precipitation test", *J. Climatol*., **6**, 661-675.

Adamowski, K. (1996), "Nonparametric estimations of low-flow frequencies", *J Hydraul Eng*., **122**(1), 46-49.

Arora, K. and Singh, V.P. (1988), "On the method of maximum likelihood estimation for the log-pearson type 3 distribution", *Stoch. Hydrol. Hydraul*., **2**(2), 155-160.

Ashkar, F. and Mahdi, S. (2003), "Comparison of two fitting methods for the log-logistic distribution", *Water Resour. Res*., **39**(8), 1217, Doi 0.1029/2002WR001685.

Adnan, N.A. and Atkinson, P.M. (2011), "Exploring the impact of climate and land use changes on streamflow trends in a monsoon catchment", *Int. J. Climatol*., **31**, 815-831.

Akaike, H. (1974). "A new look at the statistical model identification", *IEEE T. Automat. Contr.*, **19**(6), 716-723.

Bobee, B. (1974), "The log Pearson type 3 distribution and its application in hydrology", *Water Resour. Res*., **11**(5), October 1975, 681-689.

Bobee, B. and Rasmussen, P.F. (1994), "Statistical analysis of annual flood series", (Eds., Menon, J.), Trend in Hydrology, 1. Council of Scientific Research Integration, India, 117-135.

Bras, R.L. (1990), *Hydrology: an introduction to hydrologic science*, Addison-Wesley, 0201059223, 9780201059229.

Bennett, N.D., Croke, B.F.W., Guarios, G., Guillaume, J.H.A., Hamilton, S.H., Jakeman, A.J., Marsili-Libeli, S., Newham, L.T.H., Norton, J.P., Perrin, C., Pierce, S.A., Robson, B., Seppelt, R., Voinov, A.A. and Fath, B.D. (2013), "Characterising performance of environmental models", *Environ, Model. Softw*., **40**, 1-20.

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

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

Boughton, W., Srikanthan, S. and Weinmann, E. (2002), "Benchmarking a new design flood estimation system", *Aust. J. Water Resour*., **6**(1), 45-52.

Blazkova, S. and Beven, K. (2004), "Flood frequency estimation by continuous simulation of subcatchmnets rainfalls and discharges with the aim of improving dam safety assessments in a large basin in the Czech Republic", *J. Hydrol*., **292**, 153-172.

Burnham, K.P. and Anderson, D.R. (2004), *Model Selection and Multimodel inference: A Practical Information-Theoretic Approach* (2^{nd} Ed.), Springer-Verlag, ISBN 0-387-9536-7.

Burr, I.W. (1942), "*Cumulative frequency functions*", *Ann. Math. Statist*., **13**, 215-232.

Burnham, K.P. and Anderson, D.R. (2002), *Model Selection and Inference: A Practical Information-Theoretic Approach*, *2nd Ed*., Springer-Verlag, New York. http://dx.doi.org/10.1007/b97636.

Beirlant, J., Teugels. J.L. and Vynckier, P. (1996), *Practical analysis of Extreme values*, Leuven University Press, Leuven, Belgium.

Benth F.E. and Šaltyte-Benth J. (2005), "*Stochastic modelling of temperature variations with a view towards weather derivatives*", *Appl. Math. Finance*, **12**(1), 53-85.

Bowman, A.W. (1984), "An alternative method of cross-validations for the smoothing of kernel density estimates", *Biometrika*, **71**, 353-360.

Bain, L. and Engelhardt, M. (1991), Introduction to Probability and Mathematical Statistics, Duxbury Press.

Buishand, T.A. (1982), "Some methods for testing the homogeneity of rainfall records", *J. Hydrol*., **58**(1-2), 11-12.

Bedford, T. and Cooke, R.M. (2002), "Vines- a new graphical model for dependent random variables", *Ann. Stat*., **30**(4), 1031-1068.

Claeskens, G. and Hjort, N.L. (2008), *Model Selection and Model Averaging*, Cambridge University Press, 2008.

Calver, A. and Lamb, R. (1995), "Flood frequency estimation using continuous rainfall-runoff modelling", *Phys. Chem. Earth*., **20**, 479-483.

Cunnane, C. (1988), "Methods and merits of regional flood frequency analysis", *J. Hydrol*., **100**, 269-290.

Cunnane, C. (1989), "Statistical distributions for flood frequency analysis", Operational Hydrology Report no. 33, WMO no. 718, World Meteorological Organization, Geneva, Switzerland.

Choulakian, V., Jabi, EI. N. and Issa, M. (1990), "On the distribution of flood volume in partial duration series analyses of flood phenomenon", *Stoch. Hydrol. Hydraul*., **4**(3), 217-226.

Chai, T. and Draxler R.R. (2014), "Root mean square error (RMSE) or mean absolute error (MAE)?-

Arguments against avoiding RMSE in the literature", Geoscience Model Development, **7**, 1247-1250.
Cunnane, C. (1978), "Unbiased plotting positions- A review", *J. Hydrol*., **37**(3), 205-222.
Cong, R.G. and Brady, M. (2012), "The interdependence between Rainfall and Temperature: copula Analyses", The Scientific World Journal, Vol 2011, Article ID 405675.
Conover, W.J. (1999), *Practical Nonparametric Statistics*, John Wiley e Sons, New York.
Cugerone, K. and De Michele C. (2005), "Johnson SB as general functional form for raindrop size distribution", *Water Resour. Res*., **51**(8), 6276-6289. http://dx.doi.org/10.1002/2014WR016484.
Chan, N.W. (1995), "Flood disaster management in Malaysia: an evaluation of the effectiveness of government resettlement scheme", *J. Disaster Prevent. Management*, **4**, 22-29.
Chow, V.T., Maidment D.R. and Mays, L.W. (1988), *Applied Hydrology.* McGraw Hill, New York.
Chen, L., Singh, V.P. and Xiong, F. (2017), "An entropy-based generalized gamma distribution for flood frequency analysis", *Entropy*, **19**, 239.
Chambers, J.M, Cleveland, W.S., Kleiner, B. and Tukey, P.A. (1983), *Graphical Methods for Data Analysis*, Wadsworth & Brooks/Cole, Belmont, CA.
Cohn, T.A., Lane, W.L. and Baier W.G. (1997), "An algorithm for computing moments-based flood quantile estimates when historical flood information is available", *Water Resour. Res*., **33**(9), 2089-2096.
Coles, S. (2001), *An introduction statistical modelling of extreme values*, Springer, ISBN 1-85233-459-2.
Choulakian, V., Jabi, EI. N, and Issa, M. (1990), "On the distribution of flood volume in partial duration series analyses of flood phenomenon", *Stoch. Hydrol. Hydraul*., **4**(3), 217-22.
Correia, F.N. (1987), "Multivariate partial duration series in flood risk analysis", (Ed., Singh, V.P.), Hydrologic Frequency Modelling. Reidel, Dordrecht, The Netherlands, 541-554
Durrans, S.R., Eiffe, M.A., Thomas, Jr. W.O. and Goranflo, H.M. (2003), "Joint seasonal/ annual flood frequency analysis", *J. Hydrol. Eng*., **8**, 181-189.
Duins, R.P.W. (1976), "On the choice of smoothing parameters of Parzen estimators pf probability density functions", *IEEE T. Comput*., **25**, 1175-1179.
Duong, T. and Hazelton, M.L. (2003), "Plug-in bandwidth selectors for bivariate kernel density estimations", *J. Nonparametr. Stat*, **15**, 17-30.
Dooge, J.C.E. (1986)," Looking for hydrologic laws", *Water Resour. Res*, **22**(9), 465-485.
De Michele, C. and Salvadori, G. (2003), "A generalized Pareto intensity-duration model of storm rainfall exploiting 2-copulas", *J. Geophys. Res*., **108**(2), 4067. Doi: 10.1029/2002JD002534.
Daneshkhan, A., Remesan R., Omid C. and Holman, I.P. (2016), "Probabilistic modelling of flood characteristics with parametric and minimum information pair-copula model", *J. Hydrol.*, **540**, 469-487.
De Michele, C., Salvadori, G., Canossi, M., Petaccia A. and Rosso, R. (2005), "Bivariate statistical approach to check the adequacy of dam spillway", *J. Hydrol. Eng*., **10**(1), 50-57.
DID (Drainage and Irrigation Department Malaysia). (2003), "Annual flood report of DID for Peninsular Malaysia. Unpublished report", DID, Kuala Lumpur.
DID (Drainage and Irrigation Department Malaysia). (2004), "Annual flood report of DID for Peninsular Malaysia", Unpublished report, DID, Kuala Lumpur.
DID (Drainage and Irrigation Department). (2000), "Annual flood report of DID for Peninsular Malaysia", Unpublished report. DID, Kuala Lumpur.
D'Adderio, L.P., Cugerone, K., Porcu, F., De Michele, C. and Tokay, A. (2016), "Capabilities of the Johnson SB distribution in estimating rain variables", *Adv. Water Resour*., **97**, 241-250, https://doi.org/10.1016/j.advwatres.2016.09.017
Dupuis, D.J. (2007), "Using copulas in hydrology: benefits, cautions, and issues", *J. Hydrol. Eng*., **12**(4), 381-393.
Dufour J.M., Farhat, A., Gardiol, L. and Khalaf, L. (1998), "Simulation-based Finite Sample Normality Tests in Linear Regressions", *Econometrics J*., **1**, 154-173.
Efromovich, S. (1999), *Nonparametric curve estimation: methods, theory and applications*. New York: Springer-Verlag.
Eckhardt, K. (2005), "How to construct recursive digital filters for baseflow separation", *Hydrol. Process*.,

19, 507-515.
Ekanayake, S.T. and Cruise, J.F. (1993). “Comparison of Weibull- and exponential-based partial duration stochastic flood models”, *Stoch. Hydrol. Hydraul*., **7**, 283-297.
Eckhardt, K. (2004), “How to construct recursive digital filters for baseflow separation”, *Hydrol. Process*, **19**(2), https://doi.org/10.1002/hyp.5675
Fan, L. and Zheng, Q. (2016), “Probabilistic modelling of flood events using the entropy copula”, *Adv. Water Resour*., **97**, 233-240.
Favre, A.C., Adlouni, S.E., Perreault, L., Thiemonge, N. and Bobee, B. (2004), “Multivariate hydrological frequency analysis using copulas”, *Water Resour. Res*., **40**. Doi: 10.1029/2003WR002456.
Farrel, P.J. and Stewart, K.R. (2006), “Comprehensive study of tests for normality and symmetry: Extending the Spiegelhalter test”, *J. Stat. Comput. Simul*., **76**, 803-816. https://doi.org/10.1080/10629360500109023
Fan, Y.R., Huang, W.W., Huang, G.H., Huang, K., Li, Y.P. and Kong, X.M. (2015), “Bivariate Hydrological risk analysis based on coupled entropy- copula method for the Xiang xi River in the Three Gorges Reservoir area”, *Theor. Appl. Climatol*., China, Doi: 10.1007/s00704-015-1505-z.
Griffis, V.W. and Stedinger, J.R. (2007), “Log-Pearson type 3 distribution and its application in flood frequency analysis. I: Distribution characteristics”, *J. Hydrol. Eng*., **12**(5), 482-491. doi: .10.1061/(ASCE)1084-0699(2007)12:5(482)
Gupta, H.V., Kling, H., Yilmaz, K.K. and Martinez, G.F. (2009), “Decomposition of the mean squared error and NSE performance criteria: implications for improving hydrological modelling”, *J. Hydrol*., **377**(2), 80e91.
Gupta, H.V., Sorooshian, S. and Yapo, P.O. (1999), “*Status of automatic calibration for hydrologic models: Comparison with multilevel expert caliberation*”, *J. Hydrol.Eng*., **4(**2), 135-143.
Goel, N.K., Seth, S.M. and Chandra, S. (1998), “Multivariate modelling of flood flows”, *J. Hydraul. Eng*., **124**(2), 146-155.
Grimaldi, S., Baets B.D. and Verhost, N.E.C. (2013), “Multivariate return periods in hydrology: a critical and practical review focusing on synthetic design hydrograph estimation”, *Hydrol.Earth Syst. Sci.*, 1.
Graler, B., Berg, M.J.V., Vandenberg, S., Petroselli, A., Grimaldi, S., Baets B.D. and Verhost, N.E.C. (2013), “Multivariate return periods in hydrology: a critical and practical review focusing on synthetic design hydrograph estimation”, *Hydrol.Earth Syst. Sci*., **17**, 1281-1296.
Grimaldi, S. and Serinaldi, F. (2006), “Asymmetric copula in multivariate flood frequency analysis”, *Adv. ater Resour*., **29**, 1155-1167.
Gaál, L., Szolgay, J., Kohnová, S., Hlavčová, K., Parajka, J., Viglione, A. and Blöschl, G. (2015). “Dependence between flood peaks and volumes: a case study on climate and hydrological controls”, *Hydrol. Sci. J*., **60**(6), 968-984
Genest, C., Favre A.C., Beliveau, J. and Jacques C. (2007), “Meta-elliptical copulas and their use in frequency analysis of multivariate ydrological data”, *Water Resour. Res*., **43**, W09401, doi: 10.1029/2006WR005275.
Gupta, H.V., Sorooshian, S. and Yapo, P.O. (1999), “Status of automatic calibration for hydrologic models: Comparison with multilevel expert caliberation”, *J. Hydrol. Eng*., **4**(2), 135-143.
Gringorten, I.I. (1963), “A plotting rule of extreme probability paper”, *J. Geophys. Res.*, **68**(3), 813-814
Guo, S.L. (1990), “Unbiased plotting position formulae for historical floods”, *J. Hydrol.*, **121**(1-4), 45-61.
Genest, C. and Rivest, L.P. (1993), “Statistical inference procedures for bivariate Archimedean copulas”, *J. Am. Stat. Assoc*., **88**(423), 1034-1043.
Gonzales, A.L., Nonner J., Heijkers, J. and Uhlenbrook S. (2009), “Comparison of different base flow separation methods in a lowland catchment”, *Hydrol. Earth Syst. Sci*., **13**, 2055-2068
Haggag, M.M.M. (2014), “New Criteria of Model selection and model averaging in linear regression models”, *Am. J. Theor. Appl. Stat.*, **3**(5), 148-166.
Haddad, K. and Rahman, A. (2008), “Investigation on at-site flood frequency analysis in south-east Australia”, Journal - The Institution of Engineers, Malaysia , 69(3).

Hannan, E.J. and Quinn, B.G. (1979), "The determination of the order of an autoregression", *J. R. Stat. Soc. Series B Stat. Methodol*., **41**, 190-195.

Hosking, J.R.M. and Walis, J.R. (1987), "Parameter and quantile estimations for the generalized Pareto distributions", *Technometrics*, **29**(3), 339-349.

Hosking, J.R.M., Wallis, J.R. and Wood, E.F. (1985), "*Estimation of the general extreme value distribution be the method of probability weighted moments*", *Technometrics*, **27**(3), 251-261.

Hameed, K.H. (2008), "Trend detection in hydrologic data: The Mann–Kendall trend test under the scaling hypothesis", *J. Hydrol*., **349**(3-4), 350-363.

Hosking, J.M.R. and Wallis, J.R. (1997), *Regional Frequency Analysis*, Cambridge University Press. Cambridge, UK.

Hamid, A.T, Sharif, M. and Archer, D. (2014), "Analysis of Temperature Trends in Satluj River Basin, India", *J. Earth Sci. Clim. Change*, **5**, 222. Doi: 10.4172/2157-7617.1000222

Hall, M.J. (1984), *Urban Hydrology*. Barking, UK: Elsevier, 299.

Haktanır, T. (1992), "Comparison of various flood frequency distributions using annual flood peaks data of rivers in Anatolia", *J. Hydrol*., **136**, 1-31.

Heo, J., Salas J.D. and Boes D.C. (2001), "Regional Food frequency analysis based on a Weibull model: Part 2. Simulations and applications", *J. Hydrol*., **242**, 171-182.

Haktanir, T. and Horlacher, H.B. (1993), "Evaluation of various distributions for flood frequency analysis", *Hydrol. Sci*., **38**,1-2, 15-32.

Jain, D. and Singh, V. P. (1987), "Comparison of some flood frequency distributions using empirical data", *Proceedings of the International Symp. on Flood Frequency and Risk Analyses, Hydrologie Frequency Modelling*, D. Reidel Publ. Co., Dordrecht, The Netherlands.

Jamaliah, J. (2007), "Emerging Trends of Urbanization in Malaysia [online]", Acessed from: http://www.statistics.gov.my/eng/images/ stories/files/journalDOSM/V104 Article Jamaliah.pdf. [Acessed 20 January 2009].

Jones, M.C., Marron, J.S and Sheather, S.J. (1996), "A brief survey of bandwidth selection for density estimation", *J. Am. Stat. Assoc*., **91**, 401-407.

Jaiswal, R.K., Lohani, A.K. and Tiwari, H.L. (2015), "Statistical analysis for change detection and trend assessment in climatological parameters", *Environ. Process*, **2**, 729-749. DOI 10.1007/s40710-015-0105-3

Johnson, N.L. (1994), *Continuous univariate distribution*, Wiley New York, Vol 1.

Kullback, S. and Leibler, R.A. (1951), "On information and sufficiency", *Anna. Math. Stat*., **22**, 79-86.

Kendall, M. G. (1975), *Rank Correlation Methods*, 4th ed., Charles Griffin: London, 1975.

Kite, G.W. and Stuart, A. (1977), *Frequency and risk analysis in hydrology*, Water Resources pulic. Fort Collins, Co.

Katz, R.W., Parlang, M.B. and Naveau, P. (2002), "Statistics of extremes in hydrology", *Adv. Water Resour*., **25**, 1287-1304.

Karim, M.A. and Chowdhury, J.U. (1995), "A comparison of four distributions used in flood frequency analysis in Bangladesh", *Hydrol. Sci. J*., **40**(1), 55-66.

Kuchment, L.S. and Gelfan, A.N. (2011), "Assessment of extreme flood characteristics based on a dynamic-stochastic model of runoff generation and the probable maximum discharge", *Risk in Water Resources Management (Proceedings of Symposium H03 held during IUGG2011 in Melbourne*, Australia, July 2011) (IAHS Publ. 347, 2011).

Kang, H.O. and Yusof, F. (2012), "Homogeneity tests on daily rainfall series", *Int. J. Contemp. Math. Sci*., **7**, (1), 9-22

Khaliq, M., Ouarda, T., Ondo, J.C., Gachon, P. and 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

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

Krstanovic, P.F. and Singh, V.P. (1987), "A multivariate stochastic flood analysis using entropy", (Ed.,

Singh, V.P.). Hydrologic Frequency Modelling, Reidel, Dordrecht, 515-539.
Kim, T.W., Valdes J.B. and Yoo C. (2003), "Nonparametric approach for estimating return periods of droughts in arid regions", *J. Hydrol. Eng. - ASCE*, **8**(5), 237-246.
Kahya, E. and Kalayci, S. (2004), "Trend analysis of streamflow in Turkey", *J. Hydrol.*, **289**, 128-144, DOI: 10.1016/j.jhydrol.2003.11.006.
Kim, T.W., Valdes, J.B. and Yoo, C. (2006), "Nonparametric approach for bivariate drought characterisation using Palmer drought index", *J. Hydrol. Eng*., **11**(2), 134-143.
Ghosh, S. and Mujumdar, P.P. (2007), "*Nonparametric methods for modeling GCM and scenario uncertainty in drought assessments*", *Water Resour. Res*, **43**, W07405. Doi: 10.1029/2006WR005351.
Kong, X.M., Huang, G.H., Fan, Y.R. and Li, Y.P. (2015), "Maximum entropy-Gumbel–Hougaard copula method for simulation of monthly streamflow in Xiangxi river, China", *Stoch. Environ. Res. Risk A*, **29**, 833-846.
Keshtkaran, P., Sabzevari, T. and Torabihaghighi, A. (2011), "Regional Flood Frequency Analysis of Fars Rivers in Iran Using New Statistical Distributions (Case Study for Ghareaghaj and Kor Rivers)", *Geophys. Res. Abstracts*, **13**,161, 2011.
Karmakar, S. and Simonovic, S.P. (2008), "*Bivariate flood frequency analysis. Part-1: Determination of marginal by parametric and non-parametric techniques*", *J. Flood Risk Manage*., **1**, 190-200.
Karmakar, S. and Simonovic, S.P. (2009), "Bivariate flood frequency analysis. Part-2: A copula-based approach with mixed marginal distributions", *J. Flood Risk Manage*., **2**(1), 1-13.
Lim, Y.H. and Lye, L.M. (2003), "Regional flood estimation for ungauged basins in Sarawak, Malaysia", *Hydrological Sciences–Journal–des Sciences Hydrologiques*, **48**(1).
Ladson, A.R., Brown, R., Neal, B. and Nathan, R. (2013), "A standard approach to baseflow separation using the Lyne and Hollick filter", *Aust. J. Water Resour*., **17**(1), 25-34.
Lim, K.J., Engel, B.A., Tang, Z., Choi, J., Kim, K., Muthukrishnan S. and Tripathy, D. (2005), "Automated web GIS based hydrograph analysis tool, WHAT", *J. Am. Water Resour. Assoc*., 1407-1416.
Lawrence, D., Paquet, E., Gailhard, J. and Fleig, A.K. (2014), "Stochastic semi-continuous simulations for extreme flood estimations in catchments with combined rainfall-snowmelt flood regimes", *Nat. Hazard Earth Sys*., **14**, 1283-1298.
Liu, Q. and Cui, B. (2008), "Spatial and temporal variability of annual precipitation during 1961–2006 in Yellow River Basin, China", *J. Hydrol*, **361**(3-4), 330-338.
Lall, U. (1995), "Recent advances in nonparametric function estimation: Hydrological applications", *Rev. Geophys*., **33**(1), 1093-1102.
Lall, U., Moon, Y.I. and Khalil. A.F. (1993), "Kernel flood frequency estimators: Bandwidth selection and kernel choice", *Water Resour. Res*., **29**(4), 1003-1015.
Ljung, G.M. and Box, G.E.P. (1978), "On a measure of lack of fit in time series models", *Biometrika*, **65**, 297-303.
Lyne, V. and Hollick, M. (1979), "Stochastic time variable rainfall-runoff modelling", *Proceedings of the Hydrology and Water Resources Symposium*, Perth, 10-12 September, Institution of Engineers National Conference Publication, No. 79/10, 89-92
Lall, U., Rajagopalan, B and Tarboton, D.G. (1996), "A nonparametric wet/dry spell model for resampling daily precipitation", *Water Resour. Res*., **32**(9), 2803-2823.
Legates, D.R. and McCabe, G.J. (1999), "Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation", *Water Resour. Res*., **35**(1), 233-241.
Madsen, H., Rasmussen, P.F. and Rosbjerg, D. (1997), "Comparison of annual maximum series and partial duration series methods for modelling extreme hydrologic events. 1. At-site modelling", *Water Resour. Res*., **33**(4), 747-757.
Modarres, R. and Silva, V.P.R. (2007), "Rainfall trends in arid and semi-arid regions of Iran", *J. Arid Environ*., **70**, 344-355.
Mathwave Technologies: http://www.mathwave.com/help/easyfit/html/analyses/graphs/difference.html.
Martins, E.S. and Stedinger, J.R. (2000), "Generalized maximum likehood GEV quantiles estimators for

hydrologic data”, *Water Resour. Res*., **36**, 747-744.

Moriasi, D.N., Arnold, J.G., Van Liew, M.W., Bingner, R.L., Harmel, R.D. and Veith, T.L. (2007), “Model evaluation guidelines for systematic quantification of accuracy in watershed simulations”, *Transactions of the ASABE*, **50**(3), 885- 900.

MMD. (2007), “Malaysian Meteorological Department (MMD). Report on Heavy Rainfall that Caused Floods in Kelantan and Terengganu”, Unpublished report. MMD: Kuala Lumpur.

Madadgar, S. and Moradkhani, H. (2013), “Drought analysis under climate change using copula”, *J. Hydrol. Eng.- ASCE*, **18**, 746-759.

Markiewicz, I., Strupczewski, W.G., Bogdanowicz, E. and Kochanek, K. (2015), “Generalized exponential distribution in flood frequency analysis for Polish rivers”, *PLOS ONE*, DOI:10.1371/journal.pone.0143965

Morrison, J.E. and Smith, J.A. (2002), “Stochastic modeling of flood peaks using the generalized extreme value (GEV) distributions”, *Water Resour. Res*., **38** (12), 1302, DOI: 10.1029/2001WR000502.

Mann, H.B. (1945), “Nonparametric test against trend”, *Econometrics*, **13**, 245-259.

McMahon, T.A. and Srikanthan, R. (1981), “Log-Pearson type 3 distribution - is it applicable to flood frequency analysis of Australian streams?”, *J. Hydrol*., **52**, 139-147.

Mirabbasi, R., Kakheri-Fard, A. and Dinpashoh, Y. (2012), “Bivariate drought frequency analysis using the copula method”, *Theor . Appl. Climatol*., **108**, 191-206.

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

Nashwan, M.S., Ismail, T. and Ahmed, K. (2018). “Flood susceptibility assessment in Kelantan river basin using copula”, *Int. J. Eng. Technol*., **7**(2), 584-590.

Nathan, R.J. and McMahon, T.A. (1990), “Evaluation of automated techniques for base flow and recession analysis”, *Water Resour. Res*., **26**, 1465-1473.

Nadarajah, S. and Shiau, J. (2005), “Analysis of extreme flood events for the Pachang River, Taiwan”, *Water Resour. Manag*., **19**, 363-375.

Nash, J. and Sutcliffe, J. (1970), “River flow forecasting through conceptual models part i e a discussion of principles”, *J. Hydrol*., **10**(3), 282e290.

Owen, C.E.B. (2008), “Parameter Estimation for the Beta Distribution”, All Thesis and Disertation. 1614. https://scholarsarchive.byu.edu/etd/1614.

O’Connor, P.D.T. and Kleyner, A. (2012), *Practical Reliability Engineering*, Fifth Edition, John Wiley & Sons, Ltd. Published 2012 by John Wiley & Sons, Ltd.

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

Poulin, A., Huard, D., Favre, A.C. and Pugin, S. (2007), “Importance of tail dependence in bivariate frequency analysis”, *J. Hydrol. Eng*., **12**(4), 394-403.

Pettitt ,A.N. (1979), “A non-parametric approach to the change-point problem”, *Appl. Statist*., 2**8**, 126-135

Rao, D.V. (1980), “Log Pearson Type 3 Distribution: A Generalized Evaluation”, *J. Hydraulic Div. - ASCE*, **106**(5), 853-872.

Razawi, S. and Vogel, R. (2018), “Pre-whitening of hydroclimatic time series? Implications for inferred change and variability across time scales”, *J. Hydrol*., **557**(2018), 109-115.

Reddy, M.J. and Ganguli, P. (2012b), “Probabilistic assessments of flood risks using trivariate copulas”, *Theor. Appl. Climatol*., **111**, 341-360.

Requena, A., Flores, I., Mediero, L. and 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**, 1363-1378. doi: https://doi.org/10.1007/s00477-015-1138-x.

Rao, A.R. and Hameed, K.H. (2000), *Flood frequency analysis*, CRC Press, Boca Raton, Fla.

Rauf, U.F.A. and Zeephongsekul, P. (2014), “Copula based analysis of rainfall severity and duration: a case study”, *Theor. Appl. Climatol*., **115**(1-2), 153-166.

Rossi, F., Fiorentino, M. and Versaece, P. (1984), “Two component extreme value distribution for flood

frequency analysis", *Water Resour. Res*., **20**(7), 847-856.
Reddy, M.J. and Ganguli, P. (2012a), "*Bivariate Flood Frequency Analysis of Upper Godavari River Flows Using Archimedean Copulas*", *Water Resour. Manage*., DOI. 10.1007/s11269-012- 0124-z.
Scholz, F.W. and Stephens, M.A. (1987), "K-sample Anderson-Darling tests", *J. Am Stat. Assoc*., **82**(399): 918-924.
Singh, V.P. and Singh, K. (1988), "Parameter Estimation for Log-Pearson Type III Distribution by POME", *J. Hydraul. Eng.- ASCE*, **114** (1), 112-122.
Santhosh, D. and Srinivas, V.V. (2013), "Bivariate frequency analysis of flood using a diffusion kernel density estimators", *Water Resour. Res*., **49**, 8328-8343. doi: 10.1002/2011WR0100777.
Sharma, A. (2000), "Seasonal to interseasonal rainfall probabilistic forcasts for improved water supply management", *J. Hydrol*., **239**, 249-258.
Saklar, A. (1959), "Functions de repartition n dimensions et leurs marges", *Publ. Inst. Stat. Univ. Paris*, **8**, 229-231.
Sevat, E., and Dezetter, A. (1991), "Selection of calibration objective functions in the context of rainfall-runoff modeling in a sudanese savannah area", *Hydrol. Sci. J*., **36**(4), 307-330.
Singh, V.P. (1998), "Log-Pearson Type III Distribution. In: Entropy-Based Parameter Estimation in Hydrology", *Water Sci. Technol*., **30**. Springer, Dordrecht
Salvadori, G., De Michele, C. and Durante, F. (2011), "Multivariate design via copulas", *Hydrol. Earth. Syst. Sci. Discuss*, **8**(3), 5523-5558
Serinaldi, F. and Grimaldi, S. (2007), "Fully nested 3-copula procedure and application on hydrological data", *J. Hydrol. Eng*., **12**(4), 420-430.
Salinas, J.L., Castellarin, A., Viglione, A., Kohnova, S. and Kjeldsen, T. (2014), "Regional parent flood frequency distributions in Europe – Part 1: Is the GEV model suitable as a pan-European parent?", *Hydrol. Eart. Syst. Sci*., **18**, 4381-4389. https://doi.org/10.5194/hess-18-4381-2014
Singh, K. and Singh, V.P. (1991), "Derivation of bivariate probability density functions with exponential marginals", *Stochastic Hydrol. Hydraul*., **5**, 55-68.
Stedinger, J.R., Vogel, R.M. and Georgiou, E.F. (1993), *Frequency analysis of extreme events*, Chapter 18 In: Handbook of Hydrology, ed. D. R. Maidment. McGraw-Hill, New York, USA.
Stedinger, J.R., Vogel, R.M. and Foufoula-Georgiou, E. (1992), *Frequency analysis of extreme events*, (Ed., Maidment, D.R.), Handbook of Hydrology, chap. 18: New York, McGraw-Hill.
Silverman, B.W. (1986), *Density Estimation for Statistics and Data Analysis*, 1st edition, Chapman and Hall, London.
Salvadori, G. (2004), "Bivariate return periods via-2 copulas", *J. Roy. Stat. Soc. Ser. B*, **1**, 129-144.
Shiau J.T. (2003), "*Return period of bivariate distributed extreme hydrological events*", Stoch. Environ. Res. Risk Assess. **17**, 42-57.
Shao. Q., Chen. Y.D. and Zhang, L. (2008), "An extension of three-parameter Burr III distribution for low-flow frequency analysis", *Comput. Stat. Data Anal*., **52**, 1304-1314.
Seier, E. (2002), "Comparison of tests for univariate normality", *lnter. Stat. Statistical J*., **1**, 1-17.
Singh, R.S. (1977), "Applications of estimators of a density and its derivatives", *J. R. Stat. Soc. Series B Stat. Methodol*., **39**(3), 357-363.
Sraj, M., Bezak, N. and Brilly, M. (2014), "Bivariate flood frequency analysis using the copula function: a case study of the Litija station on the Sava River", *Hydrol. Process*., Doi:10.1002/hyp.10145.
Singh, J., Knapp, H.V. and Demissie, M. (2004), "Hydrologic modeling of the Iroquois River watershed using HSPF and SWAT. ISWS CR 2004-08. Champaign, Ill.: Illinois State Water Survey. Available at: www.sws.uiuc.edu/pubdoc/CR/ ISWSCR2004-08.pdf. Accessed 8 September 2005.
Singh, V.P., Guo, H. and Yu, F.X. (1993), "Parameter estimation for 3-parameter log-logistic distribution (LLD3) by Pome", *Stoch. Hydrol. Hydraul*., **7**(3), 163-177.
Scott, D.W. (1992), *Multivariate Density Estimations*, Theory, Practice and Visualization, New York: Wiley.
Salvadori, G. and De Michele, C. (2004), "Frequency analysis via copulas: theoretical aspects and

applications to hydrological events", *Water Resour. Res.*, **40**, W12511, doi: 10.1029/2004WR003133.a
Shoukri, M.M., Mian I. and Tracy, D.S. (1988), "Sampling properties of estimators of the log-logistic distribution with application to Canadian precipitation data", *Can. J. Stat.*, **16**, 223-236.
Schwarz, G.E. (1978), "Estimating the dimension of a model", *Ann. Stat.*, **6**(2), 461e464.
Shao, Q. (2004), "Notes on maximum likehood estimations for the three parameter Burr III distribution", *Comput. Stat. Data Anal.*, **45**, 675-687.
Sen, Z. (1999), "Simple risk calculations in dependent hydrological series", *Hydrol. Sci. J.*, **44**(6), 871- 878.
Schwartz, S.C. (1967), "Estimations of probability density by an orthogonal series", *Ann. Math. Stat.*, **38**, 1261-1265.
Serinaldi, F. (2015), "Dismissing return periods!", *Stoch. Environ. Res. Risk Assess*, **29**(4), 1179-1189.
Serinaldi, F. and Grimaldi, S. (2007), "Fully nested 3-copula procedure and application on hydrological data", *J. Hydrol. Eng.*, **12**(4), 420-430.
Selaman, O.S., Said, S. and Putuhena, F.J. (2007), "Flood frequency analysis for Sarawak using Weibull, Gringorten and L-Moments formula", *J. Institution of Engineers*, **68**, 1, 43-52.
Shao, Q.X., Wong, H., Xia, J. and Ip, W.C. (2004), "Models for extreme using the extended three-parameter Burr XII system with application to flood frequency analysis", *Hydrol. Sci. J.*, **49**, 685-702.
Tosunoglu, F. and Kisi, O. (2016), "Joint modelling of annual maximum drought severity and corresponding duration", *J. Hydrol.*, (In Press). http://dx.doi.org/10.1016/j.jhydrol.2016.10.018.
Toreti, A., Kuglitsch, F.G., Xoplaki, E., Della-Marta, P.M., Aguilar, E., Prohom, M. and Luterbacher, J. (2011), "A note on the use of the standard normal homogeneity test to detect inhomogeneities in climatic time series", *Int. J. Climatol.*, **31**, 630-632, DOI: 10.1002/joc.2088.
Vogel, R.M., Thomas, W.O. and McMahon, T.A. (1993), "Flood-flow frequency model selection in the southwestern United States", *Water Resour. Plan. Manage.- ASCE*, **119**(3), 353-366.
Veronika, B.M. and 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.
Wang, F.K., Keats, J.B. and Zimmir, W.J. (1996), "Maximum likelihood estimation of the burr XII parameters with censored and uncensored data", *Microelectron. Reliab.*, **36**, 359-362.
Wilems, P. (2005), "Bias analysis on the tail properties of flood frequency distributions", *Proceedings of the EGU05 Conference (General Assembly of the European Geoscience Union)*, Vienna, 24-29 April 2005; Geophysical Research Abstract, vol 7, 10299.
Willems, P. (1998), Hydrological applications of extreme value analysis, in hydrology in in a changing environment, (Eds., H, Wheater and C. Kirby), John Wiley & Sons, Chichester, vol.III, 15-25.
Wallis, J.R. (1988), "Catastrophes, computing and containment: living in our restless habitat", *Speculations in Science and Technology*, **11**(4), 295-315.
Wooldridge, S., Kalma, J. and Kuczera, G. (2001), "Parameterisation of a simple semi-distributed model for assessing the impact of land-use on hydrologic response", *J. Hydrol.*, **254**, 16-32.
Wan, I. (1996), "Urban growth determinants for the state of Kelantano of the state's policy makers", Penerbitan Akademik Fakulti Kejuruteraan dan Sains Geoinformasi. Buletin Ukur, **7**, 176-189.
Willmott, C. and Matsuura, K. (2005), "Advantage of the Mean Absolute Error (MAE) OVER THE Root Mean Square Error (RMSE) in assessing average model performance", *Clim. Res.*, **30**, 79-82.
Xu, Y., Huang, G. and Fan, Y. (2015), "Multivariate flood risk analysis for Wei River", *Stoch. Environ. Res. Risk Assess.*, DOI 10.1007/s00477-015-1196-0.
Xu, G.Y., Yan, G.X.Q. and Sun, X.G. (2005), "Interdecadal and interannual variation characteristics of rainfall in north china and its relation with the northern hemisphere atmospheric circulations", *Chinese J. Geophys.*, (in Chinese), **48** (2), 511-518.
Xu, C., Yin, J., Guo, S. and Hong, X. (2016), "Deriving design flood hydrograph based on conditional distribution: A case study of danjiangkou reservoir in Hanjiang Basin", *Math. Probl. Eng.*, **11**, 1-16.
Yue, S., Pilon, P. and Cavadias, G. (2002), "Power of the Mann-Kendall and Spearman's rho test for detecting monotonic trends in hydrological series", *J. Hydrol.*, **259**, 254-271.
Yue, S. (1999), "Applying the bivariate normal distribution to flood frequency analysis", *Water Int.*, **24**(3),

248-252.
Yue, S. (2001), "A bivariate gamma distribution for use in multivariate flood frequency analysis", *Hydrol. Process*., **15**, 1033-1045.
Yue, S. (2000), "The bivariate lognormal distribution to model a multivariate flood episode", *Hydrol. Process*., **14**, 2575-2588.
Yue, S., Ouarda, T.M.B.J., Bobee, B., Legendre, P and Bruneau, P. (1999), "The Gumbel mixed model for flood frequency analysis", *J. Hydrol*., **226**(1-2), 88-100.
Yue, S. and Rasmussen, P. (2002), "Bivariate frequency analysis: discussion of some useful concepts in hydrological application", *Hydrol. Processes,* **16**, 2881-2898.
Zhang, L (2005), "Multivariate hydrological frequency analysis and risk mapping", Doctoral dissertation, Beijing Normal University.
Zhang, L. and Singh, V.P. (2006), "Bivariate flood frequency analysis using copula method", *J. Hydrol. Eng*., **11**(2), 150.
Zhang, L. and Singh, V.P. (2007), "Trivariate flood frequency analysis using the Gumbel-Hougaard copula", *J. Hydrol. Eng*., **12**(4), 431- 439.
Zhang, R., Chen, Xi., Cheng, Q., Zhang, Z. and Shi, P. (2016), "Joint probability of precipitation and reservoir storage for drought estimation in the headwater basin of the Huaihe River, China", *Stoch. Environ. Res. Risk Assess*., **30**, 1641-165
Zhang, R., Li, Q., Chow, T.T., Li, S. and Danielescu, S. (2013), "Baseflow separation in a small watershed in New Brunswick, Canada, using a recursive digital filter calibrated with the conductivity mass balance method", *Hydrol. Process.*, **27**, 2659-2665, DOI:10.1001/hyp.9417.

MK