Nonparametric Approach to Copula Estimation in Compounding The Joint Impact of Storm Surge and Rainfall Events in Coastal Flood Analysis

Shahid Latif · Slobodan P. Simonovic

Received: 8 August 2022 / Accepted: 2 September 2022 / Published online: 10 October 2022
© The Author(s), under exclusive license to Springer Nature B.V. 2022

Abstract

The joint probability modelling of storm surges and rainfall events is the main task in assessing compound flood risk in low-lying coastal areas. These extreme or non-extreme events may not be dangerous if considered individually but can intensify flooding impact if they occur simultaneously or successively. Recently, the copula approach has been widely accepted in compound flooding but is often limited to parametric or semiparametric distribution settings in a limited number of cases. However, both parametric and semiparametric approaches assume the prior distribution type for univariate marginals and copula joint density. In that case, there is a high risk of misspecification if the underlying assumption is violated. In addition, both approaches suffer from a lack of flexibility. This study uses bivariate copula density in the nonparametric distribution setting. The joint copula structure is approximated nonparametrically by employing the Bernstein copula estimator and Beta kernel copula density, and their performances are also compared. The proposed model is tested with 46 years of rainfall and storm surge observations collected on Canada’s west coast. The marginal distribution of the selected flood variables is modelled using nonparametric kernel density estimation (KDE). Based on the different model compatibility tests, the Bernstein copula with normal KDE margins defined the joint dependence structure well. The selected nonparametric copula model is further employed to estimate joint and conditional return periods. It is found that flood hazard characteristics occurrence simultaneously is less frequent in AND-joint cases than in OR-joint cases. Also, the derived model is further used to estimate failure probability (FP) statistics to assess the variation of bivariate hydrologic risk during the project lifetime. It is found that FP statistics could be underestimated when neglecting the compound effect of storm surge and rainfall in the coastal flood risk.

Keywords Flood · Nonparametric copula · Beta kernel copula estimator · Bernstein copula estimator · Bivariate joint analysis · Return periods
1 Introduction

Flooding in low-lying coastal zones is becoming one of the most devastating hazards worldwide. Compound flooding (CF) events can be defined by a simultaneous or successive combination of multiple extreme or non-extreme factors such as oceanographic, hydrological, or meteorological (Zscheischler et al. 2018; Hendry et al. 2019; Jane et al. 2020; Naseri and Hummel 2022). These events may not be disastrous if they occur individually but have an intensive and devastating coastal impact if they coincide (Archetti et al. 2011; Zheng et al. 2013; Wahl et al. 2015; Santos et al. 2021). In reality, storm surge can be considered the primary driver of coastal flooding (Resio and Westerink 2008). It can become an extreme flooding event when combined with other meteorological factors, such as rainfall, even though mutual concurrency might not be much more substantial. The storm surge and rainfall can be naturally correlated in the coastal regions because of a common forcing mechanism that drives both events, such as low atmospheric pressure and tropical or extra-tropical cyclone (Zheng et al. 2014). In such circumstances, the probability of their joint occurrence will be higher and more appropriate for assessing hydrologic risk than separately considering each flood’s extreme contributing variable. In compounding the joint impact of rainfall and storm surge on the flood risk estimate, earlier studies proposed different statistical models, which usually account for the number of joint extremes or the existence of a joint dependence structure between the variables of interest that define flood hazard types (Coles et al. 1999; Coles 2001).

The multivariate analysis can model the multiple uncertain characteristics using a justifiable multivariate probability framework. The copula approach has been recognized as a highly flexible multivariate statistical tool in joint probability analysis that can eliminate several statistical restrictions of traditional multivariate distributions (Nelsen 2006; Zhang 2005; Shahid and Firuza 2021 and references therein). One of the most significant flexibilities is that the choice of best-fitted univariate marginals and their joint dependence structure do not necessarily need to be from the same distribution family. Recent studies frequently incorporated a copula-based multivariate framework in the joint analysis of compound events by targeting different flood-driving mechanisms (Xu et al. 2014, 2019; Wahl et al. 2015; Masina et al. 2015; Paprotny et al. 2018; Ghanbari et al. 2021; Naseri and Hummel 2022; and references therein). All these studies often introduced parametric class copulas (such as Archimedean or Extreme value (EV)) together with parametric marginals probability density functions or PDFs (such as GEV, Lognormal or Logistic) in the parametric distribution settings.

No evidence from the previous literature can be found that modelling the univariate marginals of any hydrologic extreme should be done either using the same distribution family or any fixed predefined distribution (Silverman 1986; Adamowski 1989; Sharma et al. 1998). For instance, the applicability of parametric distributions in modelling univariate marginals could be challenging and questionable if the hydrologic characteristics exhibit multimodal (or skewed) behaviour. Previous studies, such as Moon and Lall (1993, 1994), Wand and Jones (1995); Tarboton et al. (1998), Kim et al. (2006) and references therein, inferred that the implementation of a nonparametric approach via kernel density estimation (KDE) results in more appropriate density function for skewed data and does not require a prior assumption of the distribution type. Dooge (1986) already stated that no amount of statistical refinement could overcome the consequences due to the lack of prior probability distribution information of the observed random samples.
Similarly, the same statistical facts can apply to the copula function; for instance, it does not require any predefined joint PDF assumption in describing mutual dependency (Joe 1997; Nelsen 2006). The parametric copula approach frequently appears in the literature, as cited above, because of its simplicity. However, it is questionable how accurately selected parametric functions approximate joint density structure and/or univariate marginals in such an approach. To overcome this, Genest and Rivest (1995), Shih and Louis (1995), Liebscher (2005), and Bouezmarni and Rombouts (2008) pointed out the possibility of using a parametric model for the density copula and a nonparametric model for univariate marginals, called the semiparametric copula distribution settings. Applications of the semiparametric copula model are primarily in economics or financial studies and infrequent in hydro-meteorological studies. Only a few studies, such as those by Karmakar and Simonovic (2009), Rauf and Zeephongsekul (2014), Reddy and Ganguli (2012), and Shahid and Firuza (2021), are examples of modelling flood peak discharge, volume, and duration series.

However, the semiparametric framework still uses the parametric class copulas, which might contribute to underestimating the actual joint probability density (Charpentier et al. 2006). The parametric model’s parameter estimations are time-consuming using standard statistical procedures. Also, if the underlying assumption of the fitted parametric model is violated, it could be challenging or have spurious inferences (Rauf and Zeephongsekul 2014). Unless a wide variety of parametric copula classes are available, this approach is not flexible and might bear the risk of uncertainty in their estimated joint exceedance due to a lack of appropriate dependence structure. Different parametric classes, such as Archimedean copulas, have a different extent of dependency measuring capability, thus demanding extra attention and caution when applying copula to given historical observations. The nonparametric copula density can be a much more appropriate approach to adapt to any kind of dependence structure without considering specific PDF form for copula functions or their univariate marginals.

Deheuvels and Hominal (1979) significantly contributed by incorporating a nonparametric approach via empirical copula with univariate marginals. Later on, Gijbels and Mielniczuk (1990) study incorporated a 2-D copula using the smoothing kernel methods. This approach suggested a reflection method to alleviate the boundary bias problem of the kernel methods. Chen and Huang (2007) studies highlighted the importance of bivariate kernel copula joint density based on the local linear estimator. In hydro-meteorological studies, Rauf and Zeephongsekul (2014) study incorporated a nonparametric copula framework in the joint modelling of rainfall severity and duration characteristics in Victoria, Australia. Later, Latif and Mustafa (2020) introduced nonparametric copula joint density in the bivariate modelling of flood peak flow, volume and duration series for the Kelantan River in Malaysia. In modelling of multivariate copula density nonparametrically, Beta kernel estimators can be a good choice (Harrell and Davis 1982; Chen 1999). One of the most significant statistical advantages is that they are naturally free of boundary bias and can generate estimates with a minimum variance (Charpentier et al. 2006). The Bernstein estimator is another way to approximate copula density nonparametrically (Pfeifer et al. 2009). Sancetta and Satchell (2004) found an upper bound of its asymptotic bias and variance in developing a Bernstein polynomial estimator of the copula function. However, this study also establishes the asymptotic normality of the Bernstein density copula estimator without an explicit expression of the variance. The Bernstein copula can provide a higher rate of consistency (Kulpa 1999; Weiss and Scheffer 2012) and do not suffer from boundary bias problems; in other words, it can improve the estimation of the underlying dependence structure compared to nonparametric empirical copulas. The Bernstein copula density
approach gained attraction in economics (or insurance modelling or financial risk management) (e.g., Dieres et al. 2012; Pfeifer et al. 2009 and references therein) but has never been used in modelling extreme hydrologic events, to the best of our knowledge.

The present study’s novelty is introducing a nonparametric copula joint density via the Bernstein copula estimator in modelling compounding impacts of rainfall and storm surge events on flood risk in the coastal zones. The performance of the Beta kernel copula estimator is also investigated and compared with the Bernstein copula in the joint dependency simulation. Establishing flood dependence via Bernstein and Beta kernel copula estimator can facilitate a much more realistic modelling approach without considering any prior assumption about their density structure. The Bernstein estimator has never been used to model hydrologic or meteorological observations. The proposed nonparametric joint framework is applied to 46 years of historical observations collected on the west coast of Canada. The derived best-fitted copula joint density is employed further to estimate bivariate joint and conditional return periods. The best-fitted model is used to estimate failure probability (FP) to analyze the bivariate hydrologic risk of rainfall and storm surge. The FP statistics can provide practical input in compound flood risk assessments. The motivation of the present study is the lack of an efficient probabilistic framework for modelling coastal flooding on the western coast of Canada.

The low-lying regions of Canada’s east and west coasts are experiencing a significant threat from coastal flooding, especially under sea level rise (SLR) (Atkinson et al. 2016). The west coast of British Columbia (B.C.), is expected to see an increase in SLR between ½ and 1 m by the end of 2050 and 2100 (British Columbia Ministry of Environment 2013). Pirani and Najafi (2020) study already claimed a higher risk of compound flooding on the Pacific (west) Canada’s coast due to the combination of precipitation and extreme water level. The high risk of coastal water levels can further increase the risk of storm surge events. The complex interplay between storm surges and rainfall can significantly impact flood risk. This paper’s proposed nonparametric extreme modelling can be applied to any coastal region worldwide.

The remainder of this manuscript is organized as follows. The brief mathematical approach to the nonparametric copula density using the Beta kernel and Bernstein copula estimator is discussed in Sect. 2. This section also discusses the nonparametric distribution approach via 1-D Kernel density estimation (KDE) in modelling univariate flood margins. The third section presents a case study application of nonparametric copula density with KDE marginals to the joint analysis of rainfall and storm surge events. The model adequacy of the Bernstein copula is compared with the Beta kernel copula in the nonparametric joint density of flood structure. Subsequently, the bivariate return periods and failure probability statistics are estimated to explore the bivariate compound flood risk caused by rainfall and storm surge events. The last section of the paper provides the research summary and conclusions that can be used to reference coastal flood control and mitigation practices.

2 Methods

2.1 Nonparametric Estimation of 2-D Copula Joint Density

The copula function begins from Saklar’s theorem (Saklar 1959), which states that any multivariate joint probability distribution can be separated into its univariate
distributions and copula functions, which capture the mutual concurrency between them independently from marginal approximation (Nelsen 2006). The copula function eliminates the restriction in selecting the best-fitted univariate marginals and their joint structure from the same family distribution functions. Incorporating the multivariate copula in the parametric and semiparametric settings requires distributional assumptions for their univariate marginals PDFs and/or joint density (refer to Sect. 1). It could be a risk of misspecification if the underlying statistical assumptions of the selected predefined marginals PDF and/or copula density are violated and can lack flexibility. A nonparametric copula density framework is a flexible alternative to previously used parametric and semi-parametric approaches. This study introduces a nonparametric approach via kernel density estimation (KDE) in modelling univariate flood marginal distributions. The best-fitted marginals are then employed with two different nonparametric estimates of the joint copula density, the Bernstein copula estimator and the Beta kernel copula estimator (refer to Fig. 1).

2.1.1 Beta Kernel Copula Estimator

In this study, a nonparametric via the Beta kernel estimator is introduced and tested in the joint probability analysis of storm surge and rainfall characteristics. The earlier study, such as Behnen et al. (1985) and Gijbels and Mielenzuk (1990), pointed out that the nonparametric approach procedures for the density of a copula function often rely on symmetric kernels. Unfortunately, such procedures often suffer from boundary bias problems, depending on bandwidth. Different statistical approaches are highlighted and discussed in the estimation of joint copula density via nonparametric settings, such as Mirror image modification (Schuster 1985), transformed kernels (Wand et al. 1991), and boundary kernels (Müller 1991). The development of the nonparametric copula framework used in this study is based on the Beta kernel function, which was discussed in literature by Brown and Chen (1999), Harrell and Davis (1982), and Chen (1999, 2000). So-called boundary kernels achieve the Beta kernel. This nonparametric density approach can alleviate severe boundary bias problems (or is naturally free of boundary bias) often encountered in the standard kernel estimators and can perform better than other estimators (Jones 1993; Müller 1991; Renault and Scaillet 2004). The Beta kernel estimators are consistent if the actual density is unbounded at the boundary (Bouezmarni and Rolin 2003). In other words, it can be selected as a candidate to simulate a justifiable nonparametric estimator of the copula density.

Mathematically, the univariate Beta-kernel density function based on the sample of univariate variable \( M_1, M_2, \ldots, M_p \) with known compact support \([0,1]\) is given as:

\[
    s_h(m) = \frac{1}{p} \sum_{i=1}^{p} K(M_i, \frac{m}{h} + 1, \frac{1-m}{h} + 1)
\]

In Eq. (1), \( K(., a \text{ and } b) \) represent the density of the Beta distribution function with associated parameters \( a \) and \( b \), which can be estimated by

\[
    K(m, a, b) = \frac{m^a(1-m)^b}{B(a, b)}, \; m \in [0, 1]
\]
Fig. 1 Methodological flowchart for the bivariate copula joint analysis under nonparametric distribution approach in compound flood risk assessments.
where,

\[ B(a, b) = \frac{\Gamma(a + b)}{\Gamma(a)\Gamma(b)} \quad (3) \]

By putting the value of \( B(a, b) \) of Eq. (3) in Eq. (2), get

\[ K(m, a, b) = \frac{m^a(1 - m)^b}{B(a, b)} = \frac{\Gamma(a)\Gamma(b)m^a(1 - m)^b}{\Gamma(a + b)} \quad (4) \]

In Eq. (1), \( h \) is the kernel smoothing parameter called kernel bandwidth.

Mathematically, the bivariate copula joint density structure can be derived from the product of Beta kernel densities called Beta kernel copula at a point \((m, n)\) (Charpentier et al. 2006) given below:

\[ c_h(m, n) = \frac{1}{ph^2} \sum_{i=1}^{p} K(M_i, \frac{m}{h} + 1, \frac{1 - m}{h} + 1) \times K(N_i, \frac{n}{h} + 1, \frac{1 - n}{h} + 1) \quad (5) \]

In Eq. (5), the Beta kernel copula density’s smoothing parameter or bandwidth \( h \) is estimated by the rule of thumb (ROT) method by minimizing the asymptotic mean integrated squared error or AMISE using the Frank copula as the reference copula (Nagler 2014). Mathematically, ROT for the Beta kernel estimator \( c_h(m, n) \) is

\[ h = \left( \frac{1}{8\pi} \int \frac{c(c)^{1/3}}{\xi(c)^{1/3}} n^{-1/3} \right) \quad (6) \]

In Eq. (6), \( c \) can be taken as Frank copula.

### 2.1.2 Bernstein Copula Estimator

Besides the applicability of the above-discussed nonparametric framework, this study firstly introduced and tested the efficacy of the Bernstein estimator in the compound flood probability analysis. In the earlier study, Lorentz (1953) demonstrated that any continuous function within a range of \([0, 1]\) could be approximated using the Bernstein polynomials. A later study by Vitale (1975) used this polynomial for density function, and Tenbusch (1994) examined the Bernstein estimator for bivariate density function. According to Sancetta and Satchell (2004) and Diers et al. (2012), using Bernstein copulas provides a higher consistency and does not suffer from boundary bias. Compared to empirical copula, this nonparametric copula density can better estimate underlying mutual dependence and have a good estimator at an extreme and asymmetric dependence structure (Bouezmarni et al. 2013).

The \( d \) degree Bernstein polynomials can be mathematically estimated as

\[ B(d, a, b) = \binom{d}{a} b^a(1 - b)^{d-a} \quad (7) \]

where \( a = 0, 1, \ldots, d \in \mathbb{N} \) and \( 0 \leq b \leq 1 \).

Again, if \( U = (U_1, U_2) \) represents bivariate random observation with uniform marginals over \( P_i = \{0, 1, 2, \ldots, \ldots, d_i\} \) and having grid size \( d_i \in \mathbb{N} \) and \( i = 1, 2 \). Then,

\[ p(a_1, a_2) = P(\bigcap_{i=1}^{2} \{U_i = a_i\}), (a_1, a_2) \in [0, 1]^2 \quad (8) \]
Hence, the Bernstein copula density for a bivariate joint case is estimated by (Sancetta and Satchell 2004; Bouezmarni et al. 2013)

\[ c(u_1, u_2) = \sum_{d_1=0}^{d_1-1} \sum_{d_2=0}^{d_2-1} p(a_1, a_2) \prod_{i=1}^{2} d_i B(d_i - 1, a_i, u_i), (u_1, u_2) \in [0, 1]^2 \]  \hspace{1cm} (9)

Also, the 2-D Bernstein copula of Eq. (9) can be used in the estimation of the empirical copula (Deheuvels and Hominal 1979) as given below

\[ C_n(x, y) = \begin{cases} 
C_n\left( \frac{i-1}{n}, \frac{j-1}{n} \right), & \frac{i-1}{n} \leq x < \frac{i}{n}, \frac{j-1}{n} \leq y < \frac{j}{n} \\
1, & x = y = 1 
\end{cases} \]  \hspace{1cm} (10)

where

\[ C_n(x, y) = \frac{1}{n} \sum_{k=1}^{n} 1 \{ X_j \leq X(i), Y_j \leq Y(i) \} \]  \hspace{1cm} (11)

### 2.2 Nonparametric Kernel Density in Flood Marginals

Parametric families-based approximations of univariate flood marginals are based on the prior distribution assumption of their marginal pdf. In reality, each flood characteristic can exhibit different distribution behaviour and needs to be modelled separately without imposing any fixed or predefined density functions. The kernel density estimation (KDE) is a data-driven model having no prior assumption of the probability density type, especially suited for skewed distributions. The kernel estimator concept was introduced by Rosenblatt (1956). It can be employed to approximate the probability density function or PDF of the hydrologic variables. The univariate kernel density function can be employed in the estimation of the probability density of targeted univariate random observation having the following statistical property

\[ \int_{-\infty}^{+\infty} K(x)dx = 1 \]  \hspace{1cm} (12)

In Eq. (12), \( K(x) \) is the univariate kernel function, and the PDF represented here can be used as a kernel function. According to Härdle (1991), the kernel function can be described using a general equation

\[ K_t(x) = \frac{1}{t} K\left( \frac{x}{t} \right) \]  \hspace{1cm} (13)

In Eq. (13), \( 't' \) is the kernel smoothing parameter called kernel bandwidth which can help to control the degree of roughness and smoothness of the kernel density function’s shape (Moon and Lall 1993, 1994). Table 1 illustrates some standard univariate kernel functions (Silverman 1986; Karmakar and Simonovic 2008) introduced and tested in this study for the modelling of flood margins. The univariate kernel density estimator \( \hat{f}_t(x) \) of a pdf can be estimated by taking the mean of Eq. (13) of the given random observations as given below:

\[ \hat{f}_t(x) = \frac{1}{nt} \sum_{i=1}^{n} K_t\left( \frac{x - X_i}{t} \right) \]  \hspace{1cm} (14)
In Eq. (14), n is the observation count and $X_i$ is the $i^{th}$ observation. The value of 't' (kernel bandwidth) is a critical concern. Over-smoothing can bypass essential contents, and under-smoothing (or insufficient smoothing) can lead to rough density. Readers are advised to follow Jones et al. (1996) and Sharma et al. (1998) for a comprehensive overview of the kernel bandwidth selection procedure. In the present study, the direct plug-in (DPI) method is used to select the bandwidth of a kernel density estimate (Sheather and Jones 1991). The Plug-in kernel bandwidth approach is simple and performs well (Wand and Jones 1995; Chen 2015). For extended mathematical details about the bandwidth estimation approach, readers are advised to follow Sheather and Jones (1991) and Chen (2015). Our recent study (Shahid and Simonovic 2022) selected parametric GEV and Normal distributions as best-fitted in defining the marginal distributions of annual maximum 24-h rainfall and maximum storm surge (time interval $= \pm 4$ days) characteristics. These estimated results are compared with this study’s most parsimonious nonparametric KDE.

### 3 An Application

#### 3.1 Data Preparation

This study examined the joint dependency analysis of storm surge and rainfall events in the coastal flood risk using the proposed nonparametric copula distribution. The complex interplay between storm surge and rainfall can be devastating and exacerbate the impact of flooding on the coastal region. Therefore, to investigate the dependency, at first, we search for the 24-h rainfall events and their associated highest (or maximum) storm surge within a time lag of $\pm 0$ (simultaneous events), $\pm 1$, $\pm 2$, $\pm 3$, $\pm 4$ and $\pm 5$ days during annual maximum 24-h rainfall events. In the previous attempts, such as Xu et al. (2014) and Xu et al. (2019), the highest value of the second variable was selected (e.g., storm tide or storm surge) within a time lag of $\pm 0$ day (simultaneous) from the date of the highest value of the first variable (rainfall). A few studies, for instance, Wahl et al. (2015), searched for the dependency of precipitation within a time lag of $\pm 1$ day from the highest storm surge events date. It may be possible that the dependency between rainfall and storm surge is highest or more significant at $\pm 2$ or $\pm 3$ or at $\pm 5$ days. In other words, flooding may be devastating if the second event (e.g., rainfall) occurs in close succession or a few days apart before or after the first event (e.g., storm surge).

The tidal gauge collects the observed coastal water level (CWL) data at the New Westminster station (British Columbia, Canada) with the geographical location of 49.2° N Lat and 122.91° W Lon, from 1970 to 2018. The data was provided by the Department of

| SI no. | Kernel functions | $K(x)$ |
|--------|------------------|--------|
| 1      | Normal           | $(2\pi)^{-1/2}e^{-x^2/2}$ |
| 2      | Epanechnikov (or parabolic) | $0.75(1-x^2)$, $|x| \leq 1$
|        |                  | $=0$, otherwise |
| 3      | Biweight (or Quartic) | $0.9375(1-x^2)^2$, $|x| \leq 1$
|        |                  | $=0$, otherwise |
| 4      | Triweight        | $1.09375(1-x^2)^3$, $|x| \leq 1$
|        |                  | $=0$, otherwise |
Fisheries and Oceans, Canada. The predicted astronomical high tide observations are collected by and obtained from the Canadian Hydrographic Service (CHS). It should be noted that west Canada’s coast exhibited two high peaks each day. The storm surge data are estimated by the difference with proper time synchronization or matching between high tide and CWL data, which can be either positive or negative.

Similarly, the rainfall gauge station, Haney UBC RF, with a geographical location 49° 15′ 52.100″ N Lat and 122° 34′ 400″ W Lon is selected within 50 km distance from the selected tidal gauge station. The daily rainfall data is converted into the block (annual) maximum form, called the annual maximum 24-h rainfall observations. The storm surge values are selected within the above-specified time lags of the annual maximum rainfall. Their strength of dependency is measured using the parametric Pearson ($\tau$) and nonparametric Kendall tau ($\tau$) and Spearman rho ($\rho$) correlation coefficients. The investigation confirms that the maximum dependency is exhibited when the maximum value of storm surge is observed within a time lag of ±4 days (refer to Supplementary Fig. SF1). Finally, our study selected annual maximum 24-h rainfall and maximum storm surge (time interval = ±4 days) to define flood-paired observations.

Our previous study (Shahid and Simonovic 2022) confirms that the storm surge observations (time interval = ±4 days) exhibited monotonic time trend behaviour (via Mann–Kendall or MK test) but no serial correlation (via Ljung-Box test using Q-statistics). Therefore, the pre-whitening procedure is adopted to detrend storm surge time series before introducing them into the univariate and multivariate probability framework. Conversely, the annual maximum 24-h rainfall series did not have any time-trend and autocorrelation. Supplementary Table ST1 lists essential descriptive statistics of the selected compound flood driving factors.

### 3.2 Nonparametric Marginal Distributions for Compound Flood Characteristics

Selecting the best-fitted univariate probability function in describing their marginal behaviour is a mandatory pre-requisites before introducing them into the multivariate distribution framework. In this regard, parametric family distributions’ applicability is frequently tested in previous studies but can be problematic due to prior distributional assumptions about their marginal pdf (refer to Sect. 1). The nonparametric kernel density estimation, KDE, does not require any distributional type assumption and can be directly derived from the given observations. Therefore, it is especially suited for unsymmetrical data.

In this study, KDE is employed in estimating univariate flood marginals and whose performance is compared using the best-fitted parametric distributions, such as GEV and Normal distributions which are selected from our previous study (Shahid and Simonovic 2022). The bandwidth of the candidate kernel functions fitted to both flood characteristics are estimated using a direct plug-in (DPI) estimation (refer to Sect. 2.2). Tables 2 and 3 list the fitted kernel functions and their estimated bandwidth. The cumulative distribution functions (CDFs) of the KDE are estimated using an empirical approach that is based on numerical integration due to the lack of a closed-form of probability density function (PDF) and cumulative distribution function (CDF) (Kim and Heo 2002; Kim et al. 2006).

Selecting the most parsimonious univariate distribution can demand a better selection procedure by adopting a series of analytical and graphical visual inspections. The performance measures of the candidate function fitted to observed flood characteristics are undertaken, such as error indices statistics (Mean square error (MSE) and Root mean square error (RMSE)) (Moriasi et al. 2007), and information criterion statistics (such as Akaike information criterion (AIC), Bayesian information criterion (BIC) and Hannan-Quinn information...
Table 2  Summary table of fitted 1-D kernel density estimation (KDE) and its goodness-of-fit (GOF) test statistics and its comparison with parametric functions in defining the univariate marginal distribution of the annual maximum 24-h rainfall series

| Nonparametric kernel functions | Estimated bandwidth (via direct-plug-in method) | MSE (Mean square error) | RMSE (Root mean square error) | AIC (Akaike information criterion) | BIC (Bayesian information criterion) | HQC (Hannan-Quinn information criterion) | MAE (Mean absolute error) | NSE (Nash Sutcliffe model efficiency) |
|--------------------------------|-----------------------------------------------|-------------------------|-------------------------------|---------------------------------|------------------------------------|----------------------------------------|--------------------------|-----------------------------------|
| Normal*                        | 11.251                                        | 0.0003                  | 0.019                         | -358.22                         | -356.39                            | -357.53                                | 0.015                    | 0.998                             |
| Epanechnikov (or parabolic)    | 24.908                                        | 0.0007                  | 0.028                         | -326.36                         | -324.53                            | -325.67                                | 0.023                    | 0.997                             |
| Biweight (or Quartic)          | 29.507                                        | 0.0010                  | 0.031                         | -315.56                         | -313.73                            | -314.88                                | 0.026                    | 0.996                             |
| Triweight                      | 33.507                                        | 0.0011                  | 0.034                         | -308.42                         | -306.59                            | -307.73                                | 0.027                    | 0.995                             |
| Parametric GEV (Generalized extreme value distribution) (Shahid and Simonovic 2022) | Estimated parameters via Maximum likelihood estimation (MLE) location (mu = mu) = 70.38683185; scale (sigma = sigma) = 20.42282895; shape (xi = xi) = -0.07904404 |
| Nonparametric Approach to Copula Estimation in Compounding... |

Normal kernel (highlighted in bold letters with an asterisk) is the most parsimonious function in approximating the univariate marginal distribution of annual maximum 24-h rainfall observations. It exhibited a minimum value of MSE, RMSE, AIC, BIC and HQC, and also a high value of NSE fitness test statistics. The performance of parametric GEV distribution is inferior to nonparametric normal KDE in modelling rainfall marginals.
Table 3 Summary table of fitted 1-D kernel density estimation (KDE) and its goodness-of-fit (GOF) test statistics and its comparison with parametric functions in defining univariate marginal distribution of the maximum storm surge (time interval = ±4) series

| Nonparametric Kernel Functions | Estimated bandwidth (via direct-plug-in method) | MSE (Mean square error) | RMSE (Root mean square error) | AIC (Akaike information criterion) | BIC (Bayesian information criterion) | HQC (Hannan-Quinn information criterion) | MAE (Mean absolute error) | NSE (Nash Sutcliffe model efficiency) |
|--------------------------------|-----------------------------------------------|-------------------------|-------------------------------|----------------------------------|--------------------------------------|----------------------------------------|------------------------|-------------------------------|
| Normal*                        | 0.076                                         | 0.0003                  | 0.017                         | -369.83                         | -368.00                              | -369.15                                | 0.014                  | 0.999                         |
| Epanechnikov (or parabolic)    | 0.169                                         | 0.0007                  | 0.026                         | -331.95                         | -330.12                              | -331.26                                | 0.020                  | 0.997                         |
| Biweight (or Quartic)         | 0.201                                         | 0.0008                  | 0.028                         | -324.17                         | -322.34                              | -323.48                                | 0.022                  | 0.996                         |
| Triweight                      | 0.228                                         | 0.0009                  | 0.030                         | -319.30                         | -317.47                              | -318.61                                | 0.023                  | 0.997                         |
| Parametric Normal distribution (Shahid and Simonovic 2022) | Estimated parameters via Maximum likelihood estimation (MLE) | 0.0011                  | 0.034                         | -306.56                         | -302.90                              | -305.19                                | 0.026                  | 0.985                         |

Normal kernel (highlighted by in bold letters with an asterisk) is recognized as the most parsimonious function in the approximation of the univariate marginal distribution of annual maximum 24-h rainfall observations. It exhibited minimum value of MSE, RMSE, AIC, BIC and HQC, and also a high value of NSE fitness test statistics. Also, the performance of parametric GEV distribution is inferior than to nonparametric Normal KDE in modelling rainfall marginals.
criterion (HQC) (Akaike 1974; Schwarz 1978; Hannan and Quinn 1979). The AIC statistics included the lack of the model’s fit on the one hand and the model’s unreliability due to the number of model parameters on the other hand (Zhang and Singh 2007). On the other side, HQC statistics is another alternative to AIC and BIC statistics and can exhibit a higher level of consistency in determining the model’s performance (Claeskens and Hjort 2008, Haggag 2014). Besides this, Mean absolute error (MAE) and Nash–Sutcliffe model efficiency coefficient (NSE) test statistics are also taken to measure model compatibility levels to observed flood characteristics (Willmott and Kenji 2005; Nash and Sutcliffe 1970). The MSE, RMSE and MAE usually define error statistics in the units of interest such that a value closer or zero often indicates consistency or optimum model performance (Singh et al. 2004; Moriasi et al. 2007). Also, when the value of NSE statistics is nearer to 1, indicating higher predictive skill. Similarly, the NSE statistics compared the model performance by comparing the data and residual variance structure. In hydro-climate studies, the NSE test can significantly evaluate model performance (Sevat and Dezetter 1991; Legates and McCabe 1999).

Similarly, the vector of unknown statistical parameters of the fitted parametric distributions in our previous study is estimated via the maximum likelihood estimation (MLE) (refer to Tables 2 and 3). The empirical non-exceedance probabilities are estimated using the Gringorten plotting positioning approach (Gringorten 1963). For a better model performance, all the estimated statistics must be lower, except the value of NSE statistics must be nearer to 1. Based on the above investigation, it is found that the Normal KDE is identified as best fitted to define the univariate marginal distribution of both flood variables (refer to Tables 2 and 3). It exhibits the minimum value of MSE, RMSE, MAE, AIC, BIC and HQC statistics and a higher value of NSE statistics closer to 1. Also, the performance of selected Normal KDE dominates and performs much better than parametric GEV and Normal distribution selected in our previous study.

The graphical visual inspections are carried out using the cumulative distribution plots and probability-probability (P-P) plots to cross-check the fitted distributions’ data reproducing capabilities (refer to Supplementary Figs. SF2 and SF3). This investigation also supports the results obtained via analytical fitness measures. In conclusion, the best-fitted Normal KDE is selected for both flood-contributing variables and used with a nonparametric copula density estimator in flood dependency analysis.

3.3 Joint Dependency Modelling Using Nonparametric Copula Estimators

Section 3.1 and Supplementary Fig. SF1 already confirm a positive and significant (at a 5% significance level and 95% confidence interval) correlation between the flood contributing variables, annual maximum 24-h and the highest storm surge value at the time lag of ±4 days. The correlation statistics are measured using both parametric via the Pearson correlation (r) coefficient and nonparametric, via Kendall tau (τ) and Spearman coefficient (ρ) (Favre et al. 2004; Klein et al. 2011). This investigation confirms that the copula function can be employed to model joint dependence between them. In other words, compounding the combined impact of storm surge and rainfall can significantly impact the flooding effect in our study area due to the observed significant positive correlation. The parametric copula's efficacy could be attributed to underestimating actual joint density and lacking flexibility (refer to the introduction in Sect. 1). It can be demanding for the prior subjective assumption about the copula joint pdf or having a parametric class function. The nonparametric copula joint density can adapt to any type of dependence.
structure to the given observed flood characteristics without having a distributional assumption.

As discussed in Sect. 2 of the paper, the Beta kernel copula can alleviate the problem of boundary bias. On the other side, Bernstein copula-based joint approximation can provide higher consistency, lack boundary bias problems, and provides a better estimate of joint dependence structure than empirical copula. The best-fitted KDE flood marginals in the previous section are now introduced in the bivariate joint distribution analysis using the Bernstein and Beta kernel copula estimators (referring to Fig. 1). The bandwidth of the fitted Beta kernel estimator is estimated using the rule of thumb (ROT), followed by Eq. (6) (refer to Sect. 2.1.2). In fitting the Bernstein copula estimator with Normal KDE marginals (refer to Sect. 3.1 and Tables 3 and 4), the coefficients are adjusted using the procedure of Weiss and Scheffer (2012).

The model adequacy of both fitted nonparametric joint densities is assessed using the good-of-fit or GOF test statistics, such as MSE, RMSE, MAE and NSE, as shown in Table 4. The Mean absolute error or MAE between theoretical observation (obtained from the fitted 2-D copula estimators) and empirical observations (obtained from the corresponding empirical copula (Deheuvels and Hominal 1979) is also estimated. It is found that the Bernstein copula estimator with Normal KDE flood marginals outperformed the Beta kernel copula with the same marginal structure. The performance of both nonparametric copula densities is also evaluated graphically using the overlapped scatterplots between a set of (sample size N = 1000) generated random flood observations from the fitted model and observed samples—refer to Fig. 2. It is found that both nonparametric copulas adequately capture the observed mutual concurrency of storm surge and rainfall characteristics. The performance of the Bernstein copula estimator is slightly better than the Beta kernel copula in compounding the joint impact of both flood variables. The generated random observations (shown by light blue circles in Fig. 2) coincide with the natural dependence of historical flood characteristics (shown in red) much better. The reliability of the developed Bernstein copula is compared with Beta copula density using the theoretical Kendall’s τ correlation estimated from the generated sample (sample size, N = 1000) of selected models with the empirical Kendall’s value obtained from observed samples. From Table 4, it is found that, for the Bernstein copula, the minimum gap is observed between the empirical and theoretical Kendall’s τ. This result confirms that the selected model comprehensively regenerates the mutual dependence of historical flood characteristics. Supplementary Fig. SF4a, b also confirm that Bernstein copula’s performance is better in the bivariate flood dependency modelling unless both nonparametric joint density performs well. Figure 3 illustrates the derived bivariate Bernstein copula density of annual maximum 24-h rainfall and maximum storm surge (time interval = ± 4 days). Supplementary Fig. SF5 illustrates the contour plot derived from 2-D Bernstein copula density fitted to bivariate flood dependence of rainfall and storm surge events.

The performed investigation shows how effectively the nonparametric copula framework captures flood dependence structure between storm surge and rainfall without having in advance PDF in their univariate marginals and copula density. In conclusion, the selected copula is employed further in estimating exceedance probabilities and associated joint and conditional return periods. The case study results clearly show that there is merit in using a nonparametric copula joint distribution framework for west coast compound flooding analyses.
Table 4 Nonparametric copula density estimation in the bivariate flood dependence modelling and goodness-of-fit test statistics

| Nonparametric copula density estimators | MAE (Mean absolute error) | MSE (Mean square error) | RMSE (Root mean square error) | NSE (Nash Sutcliffe model efficiency) | Theoretical Kendall’s tau (\(\tau\)) | Observed (or empirical) Kendall’s tau (\(\tau\)) |
|----------------------------------------|----------------------------|-------------------------|-------------------------------|---------------------------------------|--------------------------------------|-----------------------------------------------|
| Beta Kernel                            | 0.0295                     | 0.0014                  | 0.037                         | 0.979                                 | 0.17                                 | 0.283                                         |
| Bernstein*                             | 0.0290                     | 0.0013                  | 0.036                         | 0.980                                 | 0.26                                 |                                               |

Bernstein kernel copula density (highlighted in bold with an asterisk) exhibited a minimum value of MSE, RMSE, MAE and high-value of NSE statistics. Bernstein copula also exhibits a minimum gap between theoretical and empirical Kendall’s tau, which indicates that the selected copula captures or regenerates the historical flood dependence structure much better than Beta kernel copula density.
Coastal flooding is an extreme multidimensional phenomenon characterized by multiple flood drivers. This multidimensional behaviour limits the applicability of single variable frequency analysis or return periods. The return period defines an average inter-arrival time between two successive occurrences of hydrologic events (Salvadori 2004).

The low-lying coastal regions can experience a devastating situation when either of the targeted compound flooding variables, such as storm surge or rainfall, exceeds a specific threshold value. This hydrologic situation can be analyzed using the OR-joint probability distribution relationship and their associated return period (Shiau 2003; Brunner et al. 2016).

\[ T_{\text{Rainfall, Storm surge}}^{\text{OR}} = \frac{1}{P(\text{Rainfall} \geq x \text{ OR Storm surge} \geq y)} = \frac{1}{1 - H(\text{Rainfall, Storm surge})} = \frac{1}{1 - C(F(\text{Rainfall}), F(\text{Storm surge}))} \]  

(15)

where \( C(F(\text{Rainfall}), F(\text{Storm surge})) \) is the copula-based bivariate joint cumulative distribution functions (CDFs) derived from the best-fitted nonparametric 2-D copula joint density. \( F(\text{Rain}) \) and \( F(\text{Storm surge}) \) are the univariate marginal CDFs derived from the best-fitted Normal KDE.

The situation may be worst when both flood characteristics simultaneously exceed a specific threshold value. In order to handle this hydrologic situation, the analysis of the
AND-joint probability and its associated return periods could be practical and can be mathematically estimated by

\[
T^{\text{AND}}_{\text{Rainfall, Storm surge}} = \frac{1}{P(\text{Rainfall} \geq x \text{ AND Storm surge} \geq y)} = \frac{1}{(1 - F(\text{Rainfall}) - F(\text{Storm surge}) + H(\text{Rainfall, Storm surge}))} = \frac{1}{(1 - F(\text{Rainfall}) - F(\text{Storm surge}) + C(F(\text{Rainfall}), F(\text{Storm surge}))}
\]

(16)

The OR-joint and AND-joint return periods are estimated using Eqs. (15) and (16); refer to Table 5. In this table, different flood characteristics are first estimated for the return periods of 5, 10, 20, 50, 100, 200, 500, and 1000 years. It is found that the AND-joint return
period for any bivariate flood event is higher than the OR-joint return period. This result further confirms that the occurrence of bivariate flood variables simultaneously is more frequent in the case of OR-joint and less frequent in the case of AND-joint. For instance, refer to Table 5, 100-yr flood events having the following characteristics, rainfall = 147.540 mm and storm surge = 0.337 m, the bivariate OR- and AND-joint return periods is 50.92 years and 2744.23 years. Similarly, for a 200-year flood event with the following properties, rainfall = 149.799 mm, and storm surge = 0.353 m, the OR- and AND-joint return periods are 101.06 years and 9514.74, respectively. The same table shows that the OR return period is lower than the univariate return period.

From the case study outputs, it is clear that the use of single variable probability distributions in analyzing the coastal flooding phenomena is not justified. In other words, flooding events in coastal regions may not be severe if they occur alone but can be devastating if combined. Table 5 shows that a better understanding of coastal compound flooding can be achieved by combining storm surge and rainfall instead of independently visualizing the probabilistic behaviour of these variables. Both OR and AND-joint distribution relationships and their return periods are essential. In practice, the appropriate notation of return will be decided based on the nature of the water resources problem undertaken.

For water infrastructure design, the conditional joint probability distribution relationship is necessary (Zhang and Singh 2007; Sraj et al. 2014; Brunner et al. 2016 and references therein). The nonparametric Bernstein copula joint density is used further to estimate the conditional joint return periods. Two different conditional joint distribution cases are considered for the bivariate compound events.

**Case 1:** The conditional return periods of rainfall given different percentile values of storm surge, i.e., Storm surge > y(threshold) is given by,

\[ T_{\text{Rainfall}\mid \text{Storm surge}>y(\text{threshold})} = \frac{1}{(1-F(\text{Storm surge}))(1-F(\text{Rainfall})-F(\text{Storm surge})+C(F(\text{Rainfall}), F(\text{Storm surge})))} \]  \hspace{1cm} (17)

where ‘y’ is the threshold value when considering storm surge as a conditioning variable, which is defined for different percentile values such as 25th, 50th, 75th, 90th, 95th, and 99th.

**Case 2:** The conditional return periods of rainfall given different percentile values of storm surge events, i.e., Storm surge ≤ y(threshold) is estimated by

| Return period(years) | Annual maximum 24-h rainfall (mm) | Storm surge (m) (time interval = ± 4 days) | OR-joint return period, \( T_{\text{OR Rainfall,Storm surge}} \) | AND-joint return period, \( T_{\text{AND Rainfall,Storm surge}} \) |
|----------------------|-----------------------------------|--------------------------------------------|---------------------------------|---------------------------------|
| 5                    | 97.043                            | 0.151                                      | 3.08                            | 13.17                           |
| 10                   | 110.734                           | 0.228                                      | 5.69                            | 40.97                           |
| 20                   | 128.046                           | 0.277                                      | 10.77                           | 139.73                          |
| 50                   | 143.423                           | 0.316                                      | 25.85                           | 760.16                          |
| 100                  | 147.540                           | 0.337                                      | 50.92                           | 2744.23                         |
| 200                  | 149.799                           | 0.353                                      | 101.06                          | 9514.74                         |
| 500                  | 151.689                           | 0.371                                      | 251.44                          | 43478.26                        |
| 1000                 | 152.710                           | 0.381                                      | 502.08                          | 120481.90                       |

Table 5 Comparison of univariate and bivariate return periods for different combinations of compound flooding characteristics.
The conditional return periods for both cases (Case 1 & Case 2) are estimated using Eqs. (17) and (18), where we considered different percentile values, for instance, 25th, 50th, 75th, 90th, 95th and 99th, of storm surge events as a conditioning variable, refer to Figs. 4 and 5. It is found that the conditional return period of rainfall events decreases with an increase in the percentile value of storm surge, refer to Fig. 4. For instance, in

$$T_{\text{Rainfall} \mid \text{Storm surge} \leq y(\text{threshold})} = \frac{1}{1 - \left( C(F(\text{Rainfall}), F(\text{Storm surge})) / F(\text{Storm surge}) \right)}$$

(18)
a flood event with rainfall = 146 mm, the return period is 77.29 years (when considering the 25th percentile value of storm surge ≤ 0.187 m), 73.56 years (at the 50th percentile value of Storm surge ≤ 0.301 m), and 72.58 years (when considering the 90th percentile of Storm surge ≤ 0.556 m). In other words, higher return periods are obtained at a low percentile value of storm surge than higher percentile values for the same rainfall events.

Figure 5 (conditional joint Case 1) shows that higher return periods are obtained at higher percentile values of storm surge events than lower return periods for the same rainfall event. For instance, in a flood event with rainfall = 100.8 mm, the return period is 180.39 years (when considering the 25th percentile of Storm surge > 0.187 m) and
4139.09 years (when considering the 50th percentile of Storm surge > 0.301m). After considering both joint and conditional return periods, it is concluded that the knowledge of both return periods is essential for water resources management decision-making. For example, in most hydraulic or engineering-based flood defence infrastructure designs, it could be required to consider an extreme hydrologic situation by giving importance to one design variable over another. This requirement can be easily satisfied using a conditional joint probability relationship.

3.5 Failure Probability in the Evaluation of Compound Flood Risk

The traditional definition of return periods, which considers the average inter-arrival time between two successive occurrences of extreme events, would be ineffective in characterizing the risk of potential flooding hazards during the entire project design lifetime of the infrastructure (Salvadori et al. 2011; Huang and Chen 2015). In this study, the developed nonparametric joint density is applied further in the estimation of failure probability (FP) statistics, which is considered a practical approach in hydrologic risk assessments by Serinaldi (2015) and Salvadori et al. (2016).

For the bivariate flood hazard scenarios, if Rain$_1$, Rain$_2$ … … Rain$_T$ and Storm surge$_1$, Storm surge$_2$ … … , Storm surge$_T$ are the paired flood observations, and T is the arbitrary project design lifetime; the FP statistics are estimated for different joint cases:

Case 1: In the OR-joint distribution case is given by

$$FP^{OR}_T = 1 - (1 - P(Rain \geq x \ OR \ Storm \ surge \geq y))^T$$

And

Case 2: In the AND-joint distribution is given by

$$FP^{AND}_T = 1 - (1 - P(Rain \geq x \ AND \ Storm \ surge \geq y))^T$$

The failure probability (FP) statistics are estimated for bivariate OR flood hazard scenarios (indicated by red colour in Fig. 6a–g) using Eq. (19). The same Figure also shows the comparison with the value of univariate hazard scenarios. These figures illustrate the variation in the bivariate (and univariate) hydrologic risk or hazard according to the variation in the service design lifetime (different return periods, for instance, 500 years, 200 years, 100 years, 50 years, 20 years, 10 years and 5 years). It is found that bivariate events produce higher FP statistics when compounding the collective impact of rainfall and storm surge than when considering these flood drivers individually. For example, at the return period (RP = 200 years), the value of FP statistics for the bivariate and univariate hazard scenarios is 0.328 and 0.181 (when the design lifetime is 100 years). Nevertheless, when the return periods are reduced to (RP = 100 years), the value of FP statistics for bivariate and univariate scenarios are 0.630 and 0.394. These results show that with the increase in the value of return periods, the bivariate (and univariate) flood risk decreases, at the same time, bivariate (and univariate) hydrologic risk increases with an increase in the value of the service design lifetime of the infrastructure. Thus, ignoring the compound effect of rainfall and storm surge can result in an underestimation of FP in the coastal flood risk.

This study also examined the variation of bivariate compound flood risk with changes in the annual maximum 24-h rainfall events for different synthetic maximum storm surge
Bivariate flood hazard scenario for OR-joint case (annual maximum 24-hr rainfall (mm)-maximum storm surge (m) (Time interval = ±4 days))

Univariate flood hazard scenario (annual maximum 24-hr rainfall (mm))

Univariate flood hazard scenario (maximum storm surge (m) (Time interval = ±4 days))

Fig. 6 Bivariate flood risk through compounding the joint impact of storm surge and rainfall for different return periods a 500 years, b 200 years, c 100 years, d 50 years, e 20 years, f 10 years, and g 5 years
(c) Return period = 100 years

- Bivariate flood hazard scenario for OR-joint case (annual maximum 24-hr rainfall (mm)-maximum storm surge (m) (Time interval = ±4 days))
- Univariate flood hazard scenario (annual maximum 24-hr rainfall (mm))
- Univariate flood hazard scenario (maximum storm surge (m) (Time interval = ±4 days))

(d) Return period = 50 years

- Bivariate flood hazard scenario for OR-joint case (annual maximum 24-hr rainfall (mm)-maximum storm surge (m) (Time interval = ±4 days))
- Univariate flood hazard scenario (annual maximum 24-hr rainfall (mm))
- Univariate flood hazard scenario (maximum storm surge (m) (Time interval = ±4 days))

**Fig. 6** (continued)
Fig. 6 (continued)
The storm surge is estimated at 200 years, 100 years, 50 years, 20 years, and 10 years return periods. Also, three different service times are considered, 100 years, 50 years and 30 years, refer to Fig. 7a–c. It is found that for the higher return periods of the storm surge value, the bivariate hydrologic risk (FP statistics) decreases. At the same time, the FP statistics increase with an increase in the service design lifetime of the infrastructure under consideration. It is also found that when the rainfall is above 80 mm, the bivariate flood risk decreases rapidly. When the rainfall exceeds 140 mm, the flood risk value would attain a minimum value and decreases slowly.

In conclusion, the simultaneous accounting of the joint impact of rainfall and storm surge enables a better understanding of the extreme compound phenomena and provides needed information for coastal flood risk assessments. All the above analytical and graphical investigations are vital for effective and sustainable flood management strategies in coastal regions.

4 Results and Discussions

Compounding the joint probability distribution of storm surge and rainfall events is essential in assessing the hydrologic risk of coastal flooding. The complex interplay between them can exacerbate flooding events, as both drivers are usually interconnected through a common forcing mechanism (tropical or extra-tropical cyclone events). The coastal inundation can be devastating if the storm surge and rainfall occur in close succession or
Fig. 7 Variation of bivariate compound flood risk with change of annual maximum 24-h rainfall for different designed maximum storm surge (time interval = ±4 days) when a Service time = 100 years b Service time = 50 years c Service time = 30 years

(a) Service time = 100 years

(b) Service time = 50 years

- Maximum storm surge (Time interval = ±4 days) (200-year return period)
- Maximum storm surge (Time interval = ±4 days) (100-year return period)
- Maximum storm surge (Time interval = ±4 days) (50-year return period)
- Maximum storm surge (Time interval = ±4 days) (20-year return period)
- Maximum storm surge (Time interval = ±4 days) (10-year return period)
simultaneously. Therefore, the potential for flooding in the case of their joint occurrence is higher than considering each driver individually.

The copula-based methodology is frequently used under the parametric distribution settings in the current joint distribution modelling of extreme hydro-climatological events. The parametric approach relies on prior knowledge of 'marginals' distribution assumptions (PDFs) and joint copula density in the parametric fitting procedure. Limited literature highlighted the significance of semiparametric-based joint simulation, introducing nonparametric marginal PDFs co-joined using a parametric class copula density in the parametric fitting process. A parametric fitting approach may contribute to underestimating the actual joint density. Also, it may be misleading or result in spurious inference if the underlying distribution assumption is wrong. No evidence from the existing literature points to modelling either univariate marginal distributions or copula dependence structure by employing any fixed or predefined structure.

In conclusion, parametric and semiparametric copula distribution approaches can have statistical constraints that can not be neglected. To alleviate those limitations, the nonparametric density framework is proposed that can adjust any dependence structure to the given observed flood characteristics without distributional assumption. This paper presents a nonparametric approach to the copula joint estimation by incorporating and comparing the performance of the Bernstein copula and Beta kernel copula estimator in compounding the joint impact of 46 years of storm surge and rainfall data for the west coast of Canada. The main findings of the present study are given below:
1. The paired compound flood observations are obtained by observing the annual maximum 24-h rainfall events and the highest value of storm surge events within a time lag of ± 0 (simultaneous occurrence of both events), ± 1 days, ± 2 days, ± 3 days, ± 4 days and ± 5 days. At first, the dependence between rainfall and the maximum value of storm surge extracted under different time lags is examined. It is found that the strongest correlation is observed at ± 4 days from the date of annual maximum 24-h rainfall events. Finally, this study investigates the dependency between annual maximum 24-h rainfall and maximum storm surge (Time interval = ± 4 days) in developing a joint probability distribution framework nonparametrically.

2. The maximum storm surge (Time interval = ± 4 days) observations exhibited monotonic time trend behaviour with zero serial correlation (or autocorrelation) observed in our previous study. Therefore, a pre-whitening procedure is adopted to detrend given the time series before introducing them into a univariate or multivariate distribution framework. The annual maximum 24-h rainfall series exhibited zero time trend behaviour with no autocorrelation.

3. The nonparametric 1-D kernel density estimations (KDE) are introduced in the modelling of the univariate marginal distribution. Their performance is tested and compared using different analytical and graphical goodness-of-fit (GOF) test statistics. Results confirm that the Normal KDE best describes univariate marginals of both targeted flood variables, annual maximum 24-h rainfall and maximum storm surge (Time interval = ± 4 days). The performance of the best-fitted parametric probability distribution obtained from our recent study (Shahid and Simonovic 2022) is also included in the comparison. It is observed that Normal KDE captures the marginal behaviour of both flood driving variables much better than GEV and Normal parametric function. The obtained results further reveal: how effectively one can approximate marginal behaviour without knowing the prior distribution assumption of marginal PDFs. In other words, the adequacy of the nonparametric approach is much more practical, especially when hydrologic characteristics exhibit unsymmetrical (or multimodal) distribution behaviour.

4. The nonparametric Bernstein copula and Beta kernel copula density are employed separately in the joint density approximation with the Normal KDE, which defines marginal behaviour of annual maximum 24-h rainfall and maximum storm surge (Time interval = ± 4 days) events For instance, a series of analytical model compatibility tests are employed via MSE, RMSE, MAE and NSE statistics. It is found that the Bernstein copula estimator’s performance with Normal KDE margins describes flood dependence structure better than the Beta kernel copula joint density with the same marginal structure. The reliability and model adequacy of the developed Bernstein copula density is further examined and compared analytically with Beta copula density, comparing the theoretical Kendall’s tau ($\tau$) correlation measures from the simulated random sample (sample size, N = 1000) with the empirical Kendall’s tau value obtained from the observed samples. It is found that the Bernstein copula framework regenerates the joint dependence structure of historical flood characteristics much better than the Beta kernel copula joint density. Besides this, performance between both developed nonparametric copula densities is investigated through visual inspection, using the scatterplots between a set of (sample size N = 1000) simulated flood samples from the fitted copula density and observed data. It is confirmed that the simulated flood characteristics using the Bernstein copula density better match the historical flood characteristics’ natural dependence.

5. In conclusion, both nonparametric joint density frameworks adequately capture the observed mutual dependence of flood events, with the performance of Bernstein density being slightly better. This investigation reveals that the nonparametric approach can...
provide better outcomes without subjective distributional assumptions about joint copula density and their univariate marginal PDFs. There is much scope in incorporating this nonparametric framework in modelling compound extremes such as floods, droughts and rainfall characteristics worldwide.

6. Finally, the selected Bernstein copula with a normal KDE margin is employed to estimate joint and conditional joint return periods. First, the joint return periods for OR- and AND-joint cases are considered. It is found that the OR-joint return period for any bivariate flood event is lower than the AND-joint return period. In other words, the occurrence of bivariate flood hazard characteristics simultaneously is less frequent in the case of AND-joint and more frequent in the case of OR-joint.

7. The conditional joint distribution relationships for bivariate events are also examined. It is found that the conditional return period of rainfall events decreases with an increase in the percentile value of storm surge. Higher return periods are obtained at a low percentile value of storm surge than higher percentile values for the same rainfall event. Both joint and conditional return periods are essential for water management practice. In the design of hydraulic infrastructure, conditional return periods often play a crucial role in assessing hydrologic risk.

8. The selected bivariate joint density function is employed further in the estimation of failure probability or FP for the bivariate OR hazard scenario. The FP statistics are used to assess the risk of failure associated with the bivariate return periods due to the joint occurrence or combined effect of storm surge and rainfall events. The traditional definition of return periods, for instance, OR- or AND-joint cases, would be ineffective in demonstrating the risk of potential food events occurring during the entire project design lifetime of the infrastructure. The variation in the FP statistics is observed by changes in the service design lifetime under different return periods. It is observed that bivariate flood hazard scenarios produce a higher FP than considering each flood variable independently. Also, the bivariate (and univariate) hydrologic risk increases with an increase in the design lifetime of the infrastructure and decreases with an increase in return periods. In conclusion, it could be an underestimation of failure probability in assessing coastal flood risk if the compound effect of storm surge and rainfall events is ignored.

9. Similarly, the variation of bivariate flood risk with changes in rainfall value in differently designed storm surges (i.e., 200, 100, 50, 20 and 10 years return periods) is also examined. It is found that the failure probability (or bivariate hydrologic risk) decreases when considering a designed storm surge at a higher return period. At the same instant, FP statistics increase with an increase in the service design lifetime of the infrastructure under consideration. Overall, it must be revealed that the joint simulation of rainfall and storm surge facilitates a better understanding of the coastal region’s critical hydrological behaviour during the flooding.

Using the vine copula methodology, the nonparametric bivariate copula density estimator incorporated in the present study can be projected into a higher dimensional or trivariate distribution. Shahid and Simonovic (2022) developed trivariate distribution under parametric settings in the recent study. It is confirmed that compared to bivariate analysis, trivariate flood events explain the risk of coastal flooding much more comprehensively by integrating river discharge observations’ impact in conjunction with rainfall and storm surge events. This study also confirms that neglecting the trivariate joint probability analysis can significantly underestimate the failure probability or hydrologic
risk. Our present study demonstrated that a nonparametric framework could better approximate joint density without having a particular form of univariate marginals and copula joint density. Future work will consider introducing nonparametric copula methodology in the trivariate compound flood distribution analysis using the concept of pair-copula constructions.

**Abbreviations**

KDE: Kernel Density Estimation; FP: Failure Probability; CF: Compound Flooding; JRP(s): Joint Return Period(s); PDF(s): Probability Density Function(s); CDF(s): Cumulative Distribution Function(s); EV: Extreme Value; GEV: Generalized Extreme Value; SLR: Sea Level Rise; MSE: Mean Square Error; RMSE: Root Mean Square Error; MAE: Maximum Absolute Error; AIC: Akaike Information Criterion; BIC: Bayesian Information Criterion; HQC: Hannan-Quinn Information Criterion; NSE: Nash-Sutcliffe Model Efficiency; CWL: Coastal Water Level; MLE: Maximum Likelihood Estimation; DPI: Direct Plug-in; AMISE: Asymptotic Mean Integrated Squared Error; MK: Mann-Kendall; CHS: Canadian Hydrographic Service

**Supplementary Information**
The online version contains supplementary material available at https://doi.org/10.1007/s11269-022-03321-y.

**Acknowledgements**
The work presented in the paper has been funded by the Natural Sciences and Engineering Council of Canada’s discovery grant to the second author. We thank Fisheries and Ocean Canada for providing the coastal water level observations and Environment and Climate Change Canada for daily streamflow discharge records. Special thanks to Canadian Hydrographic Service for providing the tidal data.

**Author Contributions**

Shahid Latif (SL): Conceptualization, Methodology, Software, Visualization, Investigation, Validation, Writing-Original draft preparation. Slobodan P. Simonovic (SPS): Supervision. Conceptualization, Writing- Reviewing and Editing.

**Funding**
This research was funded by the Natural Sciences and Engineering Research Council of Canada (NSERC) collaborative grant with the Institute for Catastrophic Loss Reduction (ICLR) to the second author (Slobodan P. Simonovic).

**Data Availability Statement**
Data used in the presented research are available at https://tides.gc.ca/eng/data (CWL data); https://wateroffice.ec.gc.ca/search/historical_e.html (Daily streamflow discharge records); https://climate.weather.gc.ca/ (Daily rainfall data).

**Declarations**

**Ethical Approval**
Not Applicable.

**Consent to Participate**
Not Applicable.

**Consent for Publication**
All authors have duly consented to publish this manuscript.

**Conflict of Interest**
The authors declare that they have no conflict of interest.

**References**

Adamowski K (1989) A Monte Carlo comparison of parametric and nonparametric estimations of flood frequencies. J Hydrol 108:295–308

Akaike H (1974) A new look at the statistical model identification. IEEE Trans Autom Control 19(6):716–723

Archetti R, Bolognesi A, Casadio A, Maglionioco M (2011) Development of flood probability charts for urban drainage network in coastal areas through a simplified joint assessment approach. Hydrol Earth Syst Sci 15(10):3115–3122
Atkinson DE, Forbes DL, James TS (2016) Dynamic coasts in a changing climate. In Canada's marine coasts in a changing climate, (ed.) D.S. Lemmen, F.J. Warren, T.S. James and C.S.L. Mercer Clarke; Government of Canada, Ottawa, Ontario, p. 27–68

Behnen K, Huskova M, Neuhaus G (1985) Rank estimators of scores for testing independence. Stat Decis 3:239–262

Bouezmarni T, Ghouch E, Taamouti A (2013) Bernstein estimator for unbounded copula densities. Stat Risk Model 30(4):343–360. https://doi.org/10.1524/strm.2013.2003

Bouezmarni T, Rolin JM (2003) Bernstein estimator for unbounded density function. Université Catholique de Louvain-la-Neuve

Bouezmarni T, Rombouts J (2008) Semiparametric density estimation using copulas for multivariate positive data. Comput Stat Data Anal 53:2040–2054

British Columbia Ministry of Environment (2013) Sea level rise adaptation primer, a tool kit to build adaptive capacity on Canada’s South coasts. Available online: https://www2.gov.bc.ca/assets/gov/environment/climate-change/adaptation/resources/slr-primer.pdf (Accessed on 14 Apr 2021)

Brown BM, Chen SX (1999) Beta-bernstein smoothing for regression curves with compact support. Scand J Stat 26(1):47–59

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

Chen S (2015) Optimal Bandwidth Selection for Kernel Density Functionals Estimation. J Probab Stat 2015:1–21. https://doi.org/10.1155/2015/242683

Chen SX (1999) Beta kernel estimators for density functions. Comput Stat Data Anal 31(2):131–145

Chen SX (2000) Beta kernel for regression curve. Stat Sin 10:73–92

Chen SX, Huang T (2007) Nonparametric estimation of copula functions for dependent modeling. Can J Stat 35:265–282

Claeskens G, Hjort NL (2008) Model selection and model averaging. Cambridge University Press

Coles S, Heffernan J, Tawn J (1999) Dependence measures for extreme value analyses. Extremes 2(4):339–365

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

Charpentier A, Fermanian J, Scaillet O (2006) Copulas: from theory to application in finance, 1st edn, Risk Books, Torquay, UK, chap The Estimation of Copulas: Theory and Practice

Deheuvels P, Hominal P (1979) Estimation non paramétrique de la densité compte tenu d’informations sur le support. Rev Stat Appl 27:47–68

Diers D, Eling M, Marek S (2012) Dependence modeling in non-life insurance using the Bernstein copula. Insur Math Econ 50:430–436

Dooge JCE (1986) Looking for hydrologic laws. Water Resour Res 22(9):465–485

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

Genest CKG, Rivest L (1995) Semiparametric estimation procedure of dependence parameters in multivariate families of distributions. Biomетrika 82:543–552

Ghanbari M, Arabi M, Kao S, Obeysekera J, Sweet W (2021) Climate change and changes in compound coastal-riverine flooding hazard along the U.S. coasts. Earth’s Future 9(5):1. https://doi.org/10.1029/2021e002055

Gijbels I, Mielniczuk J (1990) Estimating the density of a copula function. Commun Stat Theory Methods 19:445–464

Gringorten II (1963) A plotting rule of extreme probability paper. J Geophys Res 68(3):813–814

Hannan EJ, Quinn BG (1979) The Determination of the order of an autoregression. J Roy Stat Soc B 41:190–195

Haggag MMM (2014) New Criteria of Model selection and model averaging in linear regression models. Am J Theor Appl Stat 3(5):148–166

Härdle W (1991) Kernel density estimation. In Smoothing Techniques. Springer Series in Statistics. Springer, New York. https://doi.org/10.1007/978-1-4612-4432-5_2

Harrell FE, Davis CE (1982) A new distribution-free quantile estimator. Biomетrika 69(3):635–640

Hendry A, Haigh ID, Nicholls RJ, Winter H, Neal R, Wahl T, Joly-Laugel A, Darby SE (2019) Assessing the characteristics and drivers of compound flooding events around the UK coast. Hydrol Earth Syst Sci 23:3117–3139. https://doi.org/10.5194/hess-23-3117-2019

Huang Q, Chen Z (2015) Multivariate flood risk assessment based on the secondary return period. J Lake Sci 27(2):352–360. https://doi.org/10.18307/2015.0221

Jane R, Cadavid L, Obeysekera J, Wahl T (2020) Multivariate statistical modelling of the drivers of compound flood events in south Florida. Nat Hazard 20(10):2681–2699. https://doi.org/10.5194/nhess-20-2681-2020
Joe H (1997) Multivariate models and dependence concept. CRC Press, Boca Raton, Fla
Jones MC (1993) Simple boundary correction for kernel density estimation. Stat Comput 3:135–146
Jones MC, Marron JS, Sheather SJ (1996) A brief survey of bandwidth selection for density estimation. J Am Stat Assoc 91:401–407
Karmakar S, Simonovic SP (2008) Bivariate flood frequency analysis. Part-1: determination of marginal by parametric and non-parametric techniques. J Flood Risk Manag 1:190–200. https://doi.org/10.1111/j.1753-318X.2008.00022.x
Karmakar S, Simonovic SP (2009) Bivariate flood frequency analysis. Part-2: a copula-based approach with mixed marginal distributions. J Flood Risk Manag 2(1):1–13. https://doi.org/10.1111/j.1753-318X.2009.01020.x
Kim KD, Heo JH (2002) Comparative study of flood quantiles estimation by nonparametric models. J Hydrol 260:176–193
Kim TW, Valdes JB, Yoo C (2006) Nonparametric approach for bivariate drought characterization using Palmer drought index. J Hydrol Eng 11(2):134–143
Klein B, Schumann AH, Pahlow M (2011) Copulas-New risk assessment methodology for dam safety, food risk assessment and management. Springer, pp. 149–185
Kulpa T (1999) On approximation of copulas. Int J Math Math Sci 22:259–269
Latif S, Mustafa F (2020) A nonparametric copula distribution framework for bivariate joint distribution analysis of flood characteristics for the Kelantan River basin in Malaysia. AIMS Geosci 6(2):171–198. https://doi.org/10.3934/geosci.2020012
Legates DR, McCabe GJ (1999) Evaluating the use of “goodness-of-fit” measures in hydrologic and hydroclimatic model validation. Water Resour Res 35(1):233–241
Liebscher E (2005) Semiparametric density estimators using copulas. Commun Stat Theory Methods 34(6):318–348
Lorentz G (1953) Bernstein polynomials. University of Toronto Press
Masina M, Lamberti A, Archetti R (2015) Coastal flooding: A copula based approach for estimating the joint probability of water levels and waves. Coast Eng 97:37–52. https://doi.org/10.1016/j.coastaleng.2014.12.010
Moon YI, Lall U (1993) A kernel quantile function estimator for flood frequency analysis. Rep Pap 194. https://digitalcommons.usu.edu/water_rep/194
Moon Y-I, Lall U (1994) Kernel function estimator for flood frequency analysis. Water Resour Res 30(11):3095–3103
Moriasi DN, Arnold JG, Van Liew MW, Bingner RL, Harmel RD, Veith TL (2007) Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. Trans ASABE 50(3):885–900
Müller HG (1991) Smooth optimum kernel estimators near endpoints. Biometrika 78:521–530
Nagler T (2014) Kernel methods for vine copula estimation. Master’s thesis, Technische Universität München. https://mediatum.ub.tum.de/doc/1231221/1231221.pdf
Naseri K, Hummel MA (2022) A Bayesian copula-based nonstationary framework for compound flood risk assessment along US coastlines. J Hydrol 610:128005. https://doi.org/10.1016/j.jhydrol.2022.128005
Nash J, Sutcliffe J (1970) River flow forecasting through conceptual models part i e a discussion of principles. J Hydrol 10(3):282e290
Nelsen RB (2006) An introduction to copulas. Springer, New York
Paprotny D, Voudoukas MI, Morales-Nápoles O, Jonkman SN, Feyen L (2018) Compound flood potential in Europe. Hydrol Earth Syst Sci Discuss. https://doi.org/10.5194/hess-2018-132
Pfeiffer D, Strasserger D, Philipp JS (2009) Modelling and simulation of dependence structures in non-life insurance with Bernstein copulas. Working Paper, Carl von Ossietzky University, Oldenburg
Pirani FJ, Najafi MR (2020) Recent trends in individual and multivariate flood drivers in Canada’s coasts. Water Resour Res 56(8). https://doi.org/10.1029/2020WR027785
Rauf UFA, Zeephongsekul P (2014) Analysis of rainfall severity and duration in Victoria, Australia using nonparametric copulas and marginal distributions. Water Resour Manag 28:4835–4856. https://doi.org/10.1007/s11269-014-0779-8
Reddy MJ, Ganguli P (2012) Bivariate flood frequency analysis of Upper Godavari River flows using Archimedean copulas. Water Resour Manag. https://doi.org/10.1007/s11269-012-0124-z
Renault O, Scaillet O (2004) On the way to recovery: a nonparametric bias free estimation of recovery rate densities. J Bank Finance 28:2915–2931
Resio DT, Westerinck JJ (2008) Modeling the physics of storm surges. Phys Today 61(9). https://doi.org/10.1063/1.2982120
Rosenblatt M (1956) Remarks on some nonparametric estimates of a density function. Ann Math Stat 27(3):832–837. https://doi.org/10.1214/aoms/1177728190

Saklar A (1959) Fonctions de repartition n dimensions et leurs marges. Publications De L’institut De Statistique De L’université De Paris 8:229–231

Salvadori G (2004) Bivariate return periods via-2 copulas. J Royal Stat Soc Series B 1:129–144. https://doi.org/10.1045/j.stamet.2004.07.002

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

Salvadori G, Durante F, De Michele C, Bernardi M, Petrella L (2016) A multivariate copula-based framework for dealing with hazard scenarios and failure probabilities. Water Resour Res 52(5):3701–3721. https://doi.org/10.1002/2015WR017225

Sancetta A, Satchell S (2004) The Bernstein copula and its applications to modeling and approximations of multivariate distributions. Economet Theor 20:1–38

Santos VM, Casas-Prat M, Poschlod B, Ragno E, van den Hurk B, Hao Z, Najafi H (2021) Statistical modeling and climate variability of compound surge and precipitation events in a managed water system: A case study in the Netherlands. Hydrol Earth Syst Sci 25(6):3595–3615. https://doi.org/10.5194/hess-25-3595-2021

Schuster E (1985) Incorporating support constraints into nonparametric estimators of densities. Comm Stat Theory Methods 14:1123–1136

Schwarz GE (1978) Estimating the dimension of a model. Ann Stat 6(2):461e464

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

Sevat E, 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

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

Shahid L, Firuza M (2021) Bivariate joint distribution analysis of the flood characteristics under semiparametric copula distribution framework for the Kelantan River basin in Malaysia. J Ocean Eng Sci 6(2):128–145. ISSN 2468–0133. https://doi.org/10.1007/s00477-003-0125-9

Shahid L, Simonovic SP (2022) Compounding joint impact of rainfall, storm surge and river discharge on coastal flood risk: An approach based on 3D fully nested archimedean copulas. Environ. Earth Sci; preprint

Sharma A, Lall U, Tarboton DG (1998) Kernel bandwidth selection for a first order nonparametric streamflow simulation model. Stoch Hydrol Hydraul 12:33–52

Sheather SJ, Jones MC (1991) A reliable data-based bandwidth selection method for kernel density estimation. J Roy Stat Soc B 53:683–690

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

Shih J, Louis T (1995) Inferences on the association parameter in copula models for bivariate survival data. Biometrics 51:1384–1399

Singh J, Knapp HV, 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

Sraj M, Beznik N, Brilly M (2014) Bivariate flood frequency analysis using the copula function: A case study of the Litija station on the Sava River. Hydrol Process 29:225–238

Tarboton DG, Sharma A, Lall U (1998) Disaggregation procedures for stochastic hydrology based on nonparametric density estimation. Water Resour Res 34(1):107–119

Tenbusch A (1994) Two-dimensional Bernstein polynomial density estimation. Metrika 41:233–253

Vitale R (1975) A Bernstein polynomial approach to density estimation. In: Puri M (ed) Statistical Inference and Related Topics, vol 2. Academic Press, New York, pp 87–99

Wahl T, Jain S, Bender J, Meyers SD, Luther ME (2015) Increasing risk of compound flooding from storm surge and rainfall for major US cities. Nat Clim Chang 5:1093–1097. https://doi.org/10.1038/nclimate2736

Wand MP, Jones MC (1995) Kernel smoothing. Chapman and Hall, London, UK

Wand MP, Marron JS, Ruppert D (1991) Transformations in density estimation: rejoinder (in theory and methods). J Am Stat Assoc 86:360–361

Weiss GNF, Scheffer M (2012) Smooth nonparametric bernstein vine copulas. SSRN Electron J. https://doi.org/10.2139/ssrn.2154458

Willmott C, Kenji M (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 K, Ma C, Lian J, Bin L (2014) Joint probability analysis of extreme precipitation and storm tide in a coastal city under changing environment. PLoS ONE 9(10):e109341–e109341
Xu H, Xu K, Lian J, Ma C (2019) Compound effects of rainfall and storm tides on coastal flooding risk. Stoch Env Res Risk Assess 33:1249–1261
Zhang L (2005) Multivariate hydrological frequency analysis and risk mapping. Doctoral dissertation, Beijing Normal University.
Zhang L, Singh VP (2007) Trivariate flood frequency analysis using the Gumbel-Hougaard copula. J Hydrol Eng 12(4):431–439. https://doi.org/10.1061/(ASCE)1084-0699(2007)12:4(431)
Zheng F, Seth W, Michael L, Sisson SA (2014) Modeling dependence between extreme rainfall and storm surge to estimate coastal flooding risk. Water Resour Res 50(3):2050–2071
Zheng F, Westra S, Sisson SA (2013) Quantifying the between extreme rainfall and storm surge in the coastal zone. J Hydrol 505:172–187
Zscheischler J, Westra S, van den Hurk BJJM, Seneviratne SI, Ward PJ, Pitman A, AghaKouchak A, Bresch DN, Leonard M, Wahl T, Zhang X (2018) Future climate risk from compound events. Nat Clim Chang 8:469–477. https://doi.org/10.1038/s41558-018-0156-3

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.