A Nonstationary Multivariate Framework for Modelling Compound Flooding

Han Wang[1], Yunqing Xuan[2]

[1] China Institute of Water Resources and Hydropower Research, Beijing, China (wanghan@iwhr.com)

[2] Department of Civil Engineering, Swansea University, Swansea, United Kingdom (y.xuan@swansea.ac.uk)

Keywords: Compound floods; Multivariate analysis; Non-stationarity; Climate change

Abstract: Flooding is widely regarded as one of the most dangerous natural hazards worldwide. It often arises from various sources either individually or combined such as extreme rainfall, storm surge, high sea level, large river discharge or the combination of them. However, the concurrence or close succession of these different source mechanisms can lead to compound flooding, resulting in larger damages and even catastrophic consequences than those from the events caused by the individual mechanism. Here, we present a modelling framework aimed at supporting risk analysis of compound flooding in the context of climate change, where nonstationary joint probability of multiple variables and their interactions need to be quantified. The framework uses the Block Bootstrapping Mann-Kendall test to detect the temporal changes of marginals, and the correlation test associated with the Rolling Window method to estimate whether the correlation structure varies with time; it then evaluates various combinations of marginals and copulas under stationary and nonstationary assumptions. Meanwhile, a Bayesian Markov Chain Monte Carlo method is employed to estimate the time-varying parameters of copulas.

1. Introduction

Flooding is widely regarded as one of the most dangerous natural hazards, often resulting from various sources such as intense rainfall, storm surge, high sea level, and large river discharge either individually or in combination (Bevacqua et al., 2020; Hendry et al., 2019). Particularly, the simultaneous occurrence or close succession of these diverse source mechanisms can lead to compound flooding events, resulting in greater damage than from separate events caused by the individual mechanism (AghaKouchak et al., 2020; Hendry et al., 2019). This is further manifested by the occurrence of several recent events where inland floods were associated with hydrologic drivers (e.g., rainfall, river discharge) combined with oceanographic drivers (e.g., tides, storm surges, waves). Examples include compound floods on the North Carolina Coast of the USA (Gori et al., 2020); in the Shoalhaven estuary of Australia in June 2016 (Kumbier et al., 2018); at Noorderzijlvest of the Netherlands in 2015 (van den Hurk et al., 2015); and in Ravenna of Italy in 2015 (Bevacqua et al., 2017). Traditional approaches for estimating characteristics of such high-impact compound events typically rely on multivariate probability analysis that can calculate desired joint probability, and copula (Sklar, 1959) when considering the interdependence, interaction and associations among different drivers.

In the field of hydroclimatic sciences, numerous studies have already utilised various types of copulas to model the dependence structure among hydrological variables in order to evaluate the risks associated with compound events, e.g., Zhu et al. (2019); Renard and Lang (2007); Zhang and Singh (2006); Favre et al. (2004); Salvadori and De Michele (2004), to name just a few. However, global warming has resulted in significant changes in regional climate (Ricke et al., 2010) which not least can impact the individual hydroclimatic variables such as temperature, precipitation, sea level, snowpack, drought, and heatwave, but can also influence their interactions (Villalobos-Herrera et al., 2021; Zscheischler et al., 2019). Consequently, traditional multivariate probability analyses that assume constant variables and stationary interactions may become less accurate in capturing the characteristics of compound events under climate change.

To account for the non-stationarity in probability analysis, significant efforts have been made in recent decades to develop univariate nonstationary models for hydrological risk assessment, e.g., Cancelliere (2017); Tramblay et al.

(2013). However, progress in multivariate nonstationary studies remains very limited and requires further exploration. Chebana et al. (2013) initially considered a time-varying dependence structure between multivariate hydrological variables to estimate their joint probability; Kwon and Lall (2016) quantified the time-varying joint probabilities of the severity and duration of drought in California by modelling the nonstationary marginal distributions of these two variables using a stationary Gumbel copula; Sarhadi et al. (2018) assessed temporal changes in the joint probability of warm and dry conditions occurring at individual locations as well as multiple locations by assuming time-varying parameters of copulas; and Feng et al. (2020) investigated flood risk under nonstationary conditions resulting from climate change when floods occur simultaneously in the Huai River and Hong River of China, assuming nonstationary marginal distributions of flood magnitudes for both rivers and employing dynamic copulas to calculate the joint probability. It is important to note that these recent studies primarily focused on compound events driven only by the same or similar types of variables such as flood volume and peaks, which are inherently intercorrelated. In consequence, their contribution may be rendered less meaningful when it comes to compound events driven by significantly different variables. Furthermore, there has been limited research on quantification of the temporal changes in joint probability of different variables leading to compound floods within a framework of non-stationarity, particularly at the level of extreme.

Therefore, we developed a feasible nonstationary framework to estimate the joint probability of different types of variables and involves different types of copulas in the context of climate change impact and with various combinations of situations. This framework is able to account for both the marginal probability distributions of single variables and their correlation structure, while recognizing the representing the associated stationary or nonstationary nature. It uses the Block Bootstrapping Mann-Kendall test to detect the temporal changes of marginals, and the correlation test associated with the Rolling Window method to estimate whether the correlation structure varies with time; it then evaluates various combinations of marginals and copulas under stationary and nonstationary assumptions. Meanwhile, a Bayesian Markov Chain Monte Carlo method is employed to estimate the time-varying parameters of copulas. To test the framework, an illustrative case is demonstrated here by applying to estimate the compound flood in Ho Chi Minh City (HCMC) of Vietnam driven by a hydrometeorological factor i.e., high rainfall and an oceanographic factor, i.e., high skew surge. The example given in this chapter can be readily expanded to cover many other similar cases.

The remainder is organized as follows: section 2 describes the proposed nonstationary framework for estimating compound floods; section 3 demonstrates the example case of the application of the framework. Finally, the conclusions and recommendations for further study are given in section 4.

2. Nonstationary framework for compound flood estimation

To demonstrate the strcutre of the framework, we use an example of compound flooding resulted from two drivers, i.e., the hydrometerological driver, e.g., rainfall, and the oceanographic drivcer, e.g., waves and surges. It should be noted that other drivers can be readily included. Figure 1 shows the steps adopted by the framework to analyze floods resulting from both hydrometeorological drivers (e.g., monthly maximum rainfall, MMR of the example case) and oceanographic drivers (e.g., monthly maximum skew surge, MMS of the example case), with a focus on non-stationarity associated with climate change. This framework can be described as consisting of four main steps, which are further elaborated upon in the subsequent subsections.

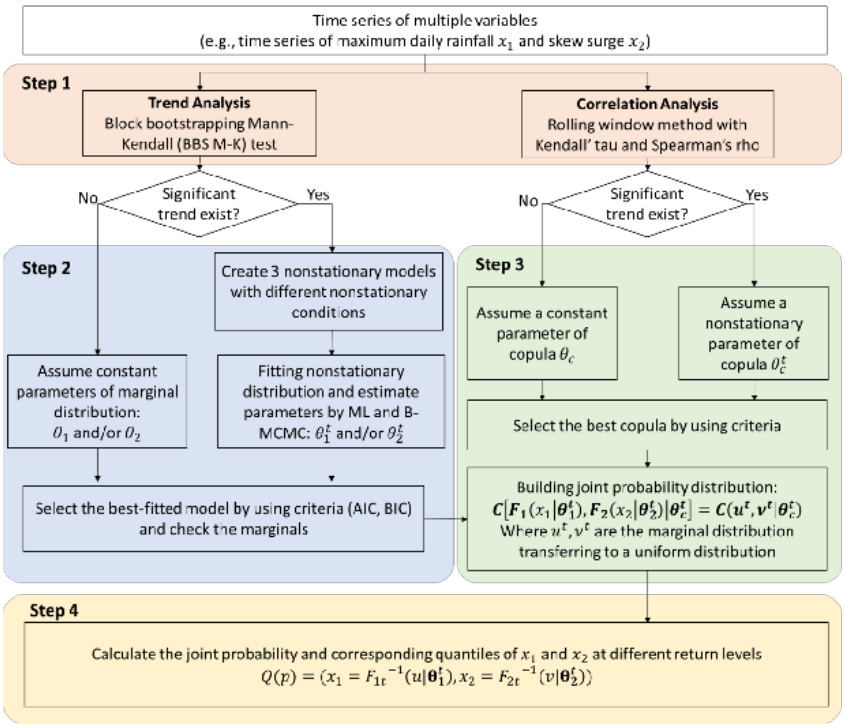

Figure 1 The nonstationary framework of multivariate probability distribution analysis. MMR is the monthly maximum rainfall; MMS is the monthly maximum skew surge.

Step 1 Trend and correlation analysis

The objective of this step is to identify changes in the values of each flood driver and the correlation structure between these drivers when other covariates are considered. Since the focus of this study is to investigate potential temporal nonstationarity in the drivers, the time variable has been selected as a covariate. This selection will serve as the foundation for parameterizing the proposed copula model. To detect monotonic trends in a series of data at a significance level of 0.05 (using R Package 'modifiedmk' Version 1.6), we employ the Block Bootstrapping Mann-Kendall (BBS-MK) test (Kundzewicz and Robson, 2004; Önöz and Bayazit, 2012).

In parallel, to examine the correlation between the series of data and its temporal variations, we employ Kendall's tau and Spearman correlation analyses in conjunction with a Rolling window approach, both of which have been widely utilized (De Winter et al., 2016; Zar, 2005). The correlation coefficients (τ and ρ) indicate respectively potential positive or negative correlations among the time series, while their corresponding p-values are compared against the critical values at a significance level of 0.05 to determine whether to reject the null hypothesis that no significant correlation exists. The Rolling window method is employed to incorporate various predefined widths of the window and progressively move towards the end of the data (Inoue et al., 2017), following the sequential steps:

i. Select the width of the rolling window m in terms of the number of consecutive data points per window. The selection is usually based on the total number S and periodicity of the data ($m < S$). As the result can be sensitive to the block length, several different lengths are tested (i.e., 10, 20 and 30 years) so that the consistency of the results is ensured.

ii. Set the number of increments between the successive rolling window, i.e., the moving step, as 1 year. In other words, I partition the data into $N = S - m + 1$ subsamples and the first rolling window covers the data from 1 to m and the second covers 2 to $m + 1$ and so on.

iii. For each rolling window, the correlation is tested by two tests and calculate the correlation coefficients and the p-value at the significance level of 0.05.

Step 2 Modelling marginal distributions of the series of data

In this step, several widely used types of probability distributions are selected as the candidates of marginal distribution to fit each series, inlcuding the Generalised Extreme Value (GEV) distribution, Generalised Pareto distribution, Gamma distribution, Lognormal distribution and Exponential distribution. Following the outcomes from Step 1, as shown in Figure 1, two assumptions can be made for the marginal distribution, i.e., it is of stationarity, which means that the parameters of the distribution remain constant and independent and can be estimated by the Maximum Likelihood (ML) method; and it is of non-stationarity, which means that multiple parameters of the distribution are assumed to be changing over time and parameters can be estimated by both the ML and the Bayesian Markov-Chain Monte-Carlo (B-MCMC) methods. A stationary and three non-stationary assumptions of these distributions are presented in Table 2. For the time series whose trend is determined to be insignificant at the significance level of 0.05, only stationary distributions are applied, otherwise, both stationary and non-stationary distributions are applied. The best-fitted distribution is finally selected by evaluating the two criteria: the Akaike's information criterion (AIC) and Bayesian information criterion (BIC); and the minimum values indicate the better performance of the model.

$$AIC = -\frac{2}{N} \times LL + 2 \times \frac{k}{N} \tag{1}$$

$$BIC = -2 \times LL + \log(N) \times k, \tag{2}$$

where N is the number of data in the series, LL is the log-likelihood of the model on these data and k is the number of parameters shown in Table 1.

Table 1 Stationary (S) and nonstationary (NS) candidate distributions for time series.

Distribution	Model	Description	Parameters ($\boldsymbol{\theta}$)
Generalized extreme value distribution	S	$F(x;\sigma,\mu,\xi) = \begin{cases} \exp\left[-(1+\xi(\frac{x-\mu}{\sigma}))^{-1/\xi}\right], \xi \neq 0 \\ \exp[-\exp(\frac{x-\mu}{\sigma})], \xi = 0 \end{cases}$ where $1+\xi(x-\mu)/\sigma > 0, -\infty < \mu < \infty, \sigma > 0$ and $-\infty < \xi < \infty$.	$\boldsymbol{\theta} = \{\sigma, \mu, \xi\}$ where, σ, μ, ξ are all constant.
	NS1	$F_t(x;\sigma,\mu_t,\xi)$	$\boldsymbol{\theta_t} = \{\mu_0, \mu_1, \sigma, \xi\}$ where, $\mu_t = \mu_0 + \mu_1 \times t$ σ, ξ are constant
	NS2	$F_t(x;\sigma_t,\mu_t,\xi)$	$\boldsymbol{\theta_t} = \{\mu_0, \mu_1, \sigma_0, \sigma_1, \xi\}$ where, $\mu_t = \mu_0 + \mu_1 \times t$ $\sigma_t = \sigma_0 + \sigma_1 \times t$ ξ is constant
	NS3	$F_t(x;\sigma_t,\mu_t,\xi)$	$\boldsymbol{\theta_t} = \{\mu_0, \mu_1, \sigma_0, \sigma_1, \xi\}$ where, $\mu_t = \mu_0 + \mu_1 \times t$ $\sigma_t = \exp(\sigma_0 + \sigma_1 \times t)$ ξ is constant

Generalized Pareto distribution	S	$F(x;\sigma,\mu,\xi)=\begin{cases}1-(1+\xi(\frac{x-\mu}{\sigma}))^{-1/\xi},\xi\neq 0\\ 1-\exp\left(-\frac{x-\mu}{\sigma}\right),\xi=0\end{cases}$ where $x\geq\mu$ when $\xi\geq 0$ and $\mu\leq x\leq\mu-\sigma/\xi$ when $\xi<0$.	$\boldsymbol{\theta}=\{\sigma,\mu,\xi\}$ where, σ,μ,ξ are constants
	NS1	$F_t(x;\sigma,\mu_t,\xi)$	$\boldsymbol{\theta_t}=\{\mu_0,\mu_1,\sigma,\xi\}$ where, $\mu_t=\mu_0+\mu_1\times t$ σ,ξ are constants
	NS2	$F_t(x;\sigma_t,\mu_t,\xi)$	$\boldsymbol{\theta_t}=\{\mu_0,\mu_1,\sigma_0,\sigma_1,\xi\}$ where, $\mu_t=\mu_0+\mu_1\times t$ $\sigma_t=\sigma_0+\sigma_1\times t$ ξ is constant
	NS3	$F_t(x;\sigma_t,\mu_t,\xi)$	$\boldsymbol{\theta_t}=\{\mu_0,\mu_1,\sigma_0,\sigma_1,\xi\}$ where, $\mu_t=\mu_0+\mu_1\times t$ $\sigma_t=\exp(\sigma_0+\sigma_1\times t)$ ξ is constant
Gamma distribution	S	$F(x;\sigma,\xi)=\frac{1}{\Gamma(\xi)}\gamma(k,\frac{x}{\sigma})$ Where $x>0,\ \sigma>0,\xi>0$.	$\boldsymbol{\theta}=\{\sigma,\xi\}$ where, σ,ξ are all constant
	NS1	$F_t(x;\sigma,\xi_t)$	$\boldsymbol{\theta_t}=\{\xi_0,\xi_1,\sigma\}$ where, $\xi_t=\xi_0+\xi_1\times t$ σ is constant
	NS2	$F_t(x;\sigma_t,\xi_t)$	$\boldsymbol{\theta_t}=\{\xi_0,\xi_1,\sigma_0,\sigma_1\}$ where, $\xi_t=\xi_0+\xi_1\times t$ $\sigma_t=\sigma_0+\sigma_1\times t$
Lognormal distribution	S	$F(x;\sigma,\mu)=\emptyset(\frac{\ln x-\mu}{\sigma})$ where $\emptyset$ is the cumulative distribution function of the standard normal distribution.	$\boldsymbol{\theta}=\{\sigma,\mu\}$ where, σ,μ are all constant
	NS1	$F_t(x;\sigma,\mu_t)$	$\boldsymbol{\theta_t}=\{\mu_0,\mu_1,\sigma\}$ where, $\mu_t=\mu_0+\mu_1\times t$ σ are constant
	NS2	$F_t(x;\sigma_t,\mu_t)$	$\boldsymbol{\theta_t}=\{\mu_0,\mu_1,\sigma_0,\sigma_1\}$ where, $\mu_t=\mu_0+\mu_1\times t$ $\sigma_t=\sigma_0+\sigma_1\times t$
	NS3	$F_t(x;\sigma_t,\mu_t)$	$\boldsymbol{\theta_t}=\{\mu_0,\mu_1,\sigma_0,\sigma_1,\xi\}$ where, $\mu_t=\mu_0+\mu_1\times t$ $\sigma_t=\exp(\sigma_0+\sigma_1\times t)$
Exponential distribution	S	$F(x;\sigma)=\begin{cases}\frac{1}{\sigma}\exp\left(-\frac{x}{\sigma}\right),x>0\\ 0,x<0\end{cases}$ where the scale parameter $\sigma=1/\lambda$ and $\lambda>0$ is the rate parameter of the exponential distribution.	$\boldsymbol{\theta}=\{\sigma\}$ where, σ is constant
	NS2	$F_t(x;\sigma_t)$	$\boldsymbol{\theta_t}=\{\sigma_0,\sigma_1\}$

			Where, $\sigma_t = \sigma_0 + \sigma_1 \times t$
	NS3	$F_t(x;\sigma_t)$	$\boldsymbol{\theta_t} = \{\sigma_0, \sigma_1\}$ Where, $\sigma_t = \exp(\sigma_0 + \sigma_1 \times t)$

Noted that μ, σ and ξ indicate location, scale and shape parameter of distribution respectively; $\boldsymbol{\theta}$ and $\boldsymbol{\theta_t}$ are symbols to indicate the parameters for each model needing to be estimated and the subscript t is used for indicating the nonstationary model; S is short for "stationarity" case and NS1, NS2 and NS3 indicate three "non-stationarity" cases.

As previously mentioned, there are several types of candidate distributions. Without losing generality, the GEV distribution is used in this example to demonstrate the process which is also followed for other types of distribution.

If a GEV distribution is the best fit for the time series in question, its stationary cumulative probability function F is defined as:

$$F(x;\sigma,\mu,\xi) = \Pr(X \le x) = \exp\left[-\left(1+\xi\left(\frac{x-\mu}{\sigma}\right)\right)^{-\frac{1}{\xi}}\right]. \tag{3}$$

F is defined for $1+\xi(x-\mu)/\sigma > 0, -\infty < \mu < \infty, \sigma > 0$ and $-\infty < \xi < \infty$, where μ is the location parameter, σ is the scale parameter, and ξ is the shape parameter which defines the three types of distribution in the GEV family: Type I, also known as the Gumbel distribution, refers to the case where $\xi = 0$ for which the F is shown in Table 1; types II and III are known as the Fréchet and the reversed Weibull distributions, corresponding to the cases where $\xi > 0$ and $\xi < 0$, respectively. The ML method is employed to estimate the three constant parameters by maximizing the likelihood function which is the product of the probability density function:

$$L(x;\boldsymbol{\theta}) = \prod_{t=t_0}^{t_{0+N}} f(x;\boldsymbol{\theta}) = \left(\frac{1}{\sigma}\left(1+\xi\left(\frac{x-\mu}{\sigma}\right)\right)^{-\frac{1}{\xi}-1}\exp\left(-\left(1+\xi\left(\frac{x-\mu}{\sigma}\right)\right)^{-1/\xi}\right)\right)^N \tag{4}$$

where t indicates the time from t_0 to t_{0+N} of the time series, N is the length of the time period and the parameter of the stationary model is denoted as $\boldsymbol{\theta} = (\sigma, \mu, \xi)$.

The nonstationary model is built when the time series shows a significant trend. In this case, the nonstationary cumulative probability function can be written as:

$$F_t(x;\sigma_t,\mu_t,\xi) = \exp\left[-\left(1+\xi\left(\frac{x-\mu_t}{\sigma_t}\right)\right)^{-\frac{1}{\xi}}\right] \tag{5}$$

F_t follows the same form as the stationary one except that an additional subscript t is added to the location and scale parameters, which indicates that both parameters are time-dependent and controlled by several hyper-parameters, i.e., for MMS $\sigma_t = (\sigma_{S0}, \sigma_{S1})$ and $\mu_t = (\mu_{S0}, \mu_{S1})$, and for MMR $\sigma_t = (\sigma_{R0}, \sigma_{R1})$ and $\mu_t = (\mu_{R0}, \mu_{R1})$. The shape parameter, ξ_S and ξ_R, is assumed to be constant. Therefore, the parameters of nonstationary models can be denoted as $\boldsymbol{\theta_S} = (\sigma_{S0}, \sigma_{S1}, \mu_{S0}, \mu_{S1}, \xi_S)$ and $\boldsymbol{\theta_R} = (\sigma_{R0}, \sigma_{R1}, \mu_{R0}, \mu_{R1}, \xi_R)$. Both ML and B-MCMC methods are used to estimate the value of them. The B-MCMC method makes use of Bayesian inference to estimate the posterior distribution of the time-varying parameters $\boldsymbol{\theta}$ of the nonstationary model. In this study, the estimated parameters of the stationary model are used to define the initial prior values of the nonstationary model. The prior distribution of parameters is assumed to be a uniform distribution. The transformation from prior distribution to

posterior distribution is done by multiplying its likelihood, which is given by (Rasmussen and Ghahramani, 2003):

$$p(\boldsymbol{\theta}|x,t) \propto p(\boldsymbol{\theta}|t) \times p(x|\boldsymbol{\theta},t) = p(\boldsymbol{\theta}|t) \times \prod_{t=1}^{N} p(x_t|\boldsymbol{\theta}_t,t) \quad (6)$$

where $p(x|\boldsymbol{\theta},t) \propto L(x;\boldsymbol{\theta},t)$ is the likelihood function and $p(\boldsymbol{\theta}|t)$ is the prior probability distribution of the parameters $\boldsymbol{\theta}$; t indicates the state. Numerical iterations for processing the posterior distribution are carried out by using MCMC simulation (Manly, 2018; Metropolis and Ulam, 1949; Murthy, 2004) The final simulation results are compared with those estimated using the ML method.

The Bayesian MCMC algorithm starts with numerical iterations to explore the posterior distribution. It is carried out using MCMC simulations with Metropolis within Gibbs sampling. The essence of the MCMC algorithm is to generate a trial move from the current state of the Markov Chain with a prior probability of parameters $p(\boldsymbol{\theta}|t)$ to the next proposed state with a prior probability of the proposed parameters $p(\boldsymbol{\theta}'|t)$. In this study, to make full use of the knowledge, the estimated parameters of the stationary model were used to define the initial prior values of the nonstationary parameters which are drawn from uniform distributions using Latin Hypercube Sampling (LHS). The Metropolis ratio is calculated to accept or reject proposal status and the convergence of simulation is monitored by Gelman-Rubin diagnostic.

This algorithm starts as a random search over the entire prior distribution ($p(\boldsymbol{\theta}|t)$ of D parameters using the LHS method then *d* samples are randomly assigned to N Markov chains and the sample with the highest likelihood value will be selected as the starting point for each chain. To diversify the probability of the jumping direction, we broadly followed to use two approaches to updating the chain: some chains (N1) follow the Adaptive Metropolis (AM) approach which is effective for searching direction at the early stage of MCMC and the rest (N-N1) follow the Differential evolution (DE) approach which has a stronger potential in converging to the target distribution. The details are shown below:

1) For each chain, randomly select d samples from D parameter spaces with Gibbs sampling.
2) For the N1 chains, propose a new state S_{t+1} with a proposed set of parameters $\boldsymbol{\theta}'$ by using AM approach, i.e., $S_{t+1}(\boldsymbol{\theta}') = S_t(\boldsymbol{\theta}') + (1-\beta)N(0_{\boldsymbol{\theta}'}, \gamma_1^2 \sum_{\boldsymbol{\theta}'}) + \beta N(0_{\boldsymbol{\theta}'}, \gamma_2^2 I_{\boldsymbol{\theta}'})$ where $\sum_{\boldsymbol{\theta}'}$ is the covariance matrix of $\boldsymbol{\theta}'$.
3) For the rest N-N1 chains, the new state is $S_{t+1}(\boldsymbol{\theta}') = S_t(\boldsymbol{\theta}') + \gamma_3(S_{r_2} - S_{r_1}) + e$.
   where γ indicates the jump factors defined as γ_1 is a number randomly selected from $[1.2,2.2]$; $\gamma_2 = 2.38/\sqrt{d}$, $\gamma_3 = 0.1/\sqrt{d}$ and $\gamma_4 = 2.38/\sqrt{2d}$ and S_{r_1} and S_{r_2} are two samplers drawn from parameter space D just for pre-defining the chain update direction.
4) Compute the Metropolis ratio $\frac{p(x|\boldsymbol{\theta}',t)}{p(x|\boldsymbol{\theta},t)}$; if $\min\left(1, \frac{p(x|\boldsymbol{\theta}',t)}{p(x|\boldsymbol{\theta},t)}\right) \geq p^*$, then accept S_{t+1} and update the current chain where p^* is the random number drawn from $\boldsymbol{N}(0,1)$. If not, reject and go back to the previous step to re-propose the state.
5) Check whether the iteration convergence or not by Gelman-Rubin convergence diagnostic. End the process if it converges.

Step 3 Building copulas and calculating the joint probability

Let J denote the joint cumulative distribution function of the two series of data, and C denote the copula function parameterized by θ_C. Then, the basic joint probability can be calculated by:

$$J(x_S, x_R|\theta_C) = C(F_1(x_1|\boldsymbol{\theta_1}), F_2(x_2|\boldsymbol{\theta_2})|\theta_C) = C(u,v|\theta_C) \quad (7)$$

where F_1 and F_2 indicate the marginal cumulative probability function of the two series of data x_1 (in the case

study, MMS) and x_2 (MMR) with their estimated parameters $\boldsymbol{\theta_1}$ and $\boldsymbol{\theta_2}$ respectively, and θ_C indicates the set of parameters of the copula. u and v are the marginal probabilities of F_1 and F_2 in the unit hypercube with uniform marginal distributions $U(0,1)$. According to the trend analysis of the individual time series and their mutual correlation structure, four contexts are relevant in this framework:

- Both marginal distributions ($\boldsymbol{\theta}$) are stationary, and the correlation structure (θ_C) is stationary.
- Both marginal distributions ($\boldsymbol{\theta}$) are stationary, while the correlation structure (θ_C^t) is nonstationary.
- At least one of the marginal distributions ($\boldsymbol{\theta_t}$) is nonstationary, while the correlation structure (θ_C) is stationary.
- At least one of the marginal distributions ($\boldsymbol{\theta_t}$) is nonstationary, while the correlation structure (θ_C) is nonstationary.

In this framework, several well-known one-parameter copulas are selected as the candidates to characterize the dependence structure between two series of data, namely, Gaussian, Clayton, Frank, Gumbel, Joe, Plackett and Raftery copulas whose parameter θ_C is estimated by using both the local optimization method and MCMC approach and processed by using a modified MvCAT toolbox (Sadegh et al., 2017) that incorporates the nonstationary terms. However, if there is no significant correlation identified, i.e., the two variables are independent, we also involve an independent copula which is simply reduced to the form where the joint probability is calculated by the product of two marginal probabilities of the variables. The algorithm can be demonstrated below:

1) Read Data: Input two samples (e.g., MMR and MMS) and the parameters of their best-fitted stationary cumulative probability function (CDF).
2) Rank each data and transfer MMR and MMS to uniform marginal.
3) Compute the empirical univariate probability distribution of each marginal u and v; and the empirical bivariate probability distribution of (u, v).
4) Conduct a random search over the entire prior distribution $(p(\boldsymbol{\theta_C^t}|t)$ of D parameters.
5) Randomly assign d $(d < D)$ samples to the N Markov chains and the sample with the highest likelihood value will be selected as the starting point for each chain.
6) Adopt the MCMC algorithm presented in 1) to 6). However, the difference exists in calculating the CDF of the copula family during MCMC simulation. Instead of using one constant parameter, the parameter of the copula is changing over time controlled by two constant parameters $(\theta_{C0}, \theta_{C1})$ and the equation is:

$$\theta_C^t = \theta_{C0} + \theta_{C1} \times t \tag{8}$$

Therefore, the joint probability equation becomes:

$$J_t(x_{1t}, x_{2t}|\theta_C^t) = C\left(F_{1t}(x_{1t}|\boldsymbol{\theta_{1t}}), F_{2t}\left(x_{2(t)}|\boldsymbol{\theta_{2t}}\right)\middle|\theta_C^t\right) = C(u_t, v_t|\theta_C^t) \tag{9}$$

where $\boldsymbol{\theta_{1t}}$ and $\boldsymbol{\theta_{2t}}$ indicate the time-varying parameters of the two marginal distributions shown in Table 1 and u_t, v_t are the nonstationary marginal probabilities converting in the uniform U [0,1]. These parameters can also be estimated by the B-MCMC method, and the posterior joint distribution can be calculated as (Ausin et al., 2010):

$$p(\emptyset|x_{1t}, x_{2t}) \propto p(\emptyset|t) \times \prod_{t=1}^{N} p(x_{1t}, x_{2t}|\emptyset, t) \tag{10}$$

where $p(x_{1t}, x_{2t}|\emptyset, t) = c\left(F_{1t}(x_{1t}|\boldsymbol{\theta_{1t}}), F_{2t}\left(x_{2(t)}|\boldsymbol{\theta_{2t}}\right)\middle|\theta_C^t\right) \times f_{1t}(x_{1t}|\boldsymbol{\theta_{1t}}) \times f_{2t}(x_{2t}|\boldsymbol{\theta_{2t}})$ is the copula density function, f_{1t} and f_{2t} are the marginal probability density functions and the parameters of the joint posterior are $\emptyset = (\mu_{S0}, \mu_{S1}, \sigma_{S0}, \sigma_{S1}, \xi_S, \mu_{R0}, \mu_{R1}, \sigma_{R0}, \sigma_{R1}, \xi_2, \theta_{C0}, \theta_{C1})$. $p(\emptyset|t)$ is the prior distribution of the parameters $\emptyset$ and according to the prior knowledge for which we assume a uniform distribution for all parameters, i.e., the parameters of the marginal distributions are assumed to be uniformly distributed around the values estimated by stationary assumption and the copula parameters are assumed to be within the maximum and minimum limits subject to copula

types. To reduce the computing time of the MCMC algorithm, the nonstationary marginal parameters are firstly estimated by applying the approach in step 2 before being transformed into u_t and v_t. The MCMC algorithm in this step is only used to estimate the copula parameter by:

$$p(\theta_C^t|u_t, v_t) \propto p(\theta_C^t|t) \times \prod_{t=1}^{N} c(u_t, v_t|\theta_C^t, t) \tag{11}$$

Finally, the best copula is selected by evaluating the goodness of fit measure AIC.

Step 4 Generating the quantiles.

The final step of the framework is to calculate the quantiles of the joint exceedance probability that has been defined in the last step. Depending on the outcomes from the analysis of the marginal distributions and correlation structure obtained from the previous steps, the stationary context will lead to a single quantile with a given probability of p, while the nonstationary context will obtain a series of quantiles changing over covariate (i.e., time in the study) for the same given p. The quantiles can be expressed as:

$$Q(p) = (x_1 = {F_{1t}}^{-1}(u|\boldsymbol{\theta_1^t}), x_2 = {F_{2t}}^{-1}(v|\boldsymbol{\theta_2^t})) \tag{12}$$

where $p = C_t(u, v|\theta_C^t)$ and t indicates that the parameters or variables are changing over time. If the context is stationary and the two variables are independent, $Q(p)$ can be simply calculated by inverting the marginal distributions, i.e., $x_1 = {F_1}^{-1}(u|\boldsymbol{\theta_1})$ and $x_2 = {F_2}^{-1}(v|\boldsymbol{\theta_2})$.

There are several approaches for selecting the cases of combination of marginals (e.g., the variable of x_1 and x_2) and arguably the mostly-used approach is to get the most likely combined with the highest joint density level (Sadegh et al., 2018; Salvadori et al., 2014). For example, Figure 6 illustrates all quantile curves corresponding to the probability from $p = 0.01$ to $p = 0.99$. The horizontal *x*-axis and vertical *y*-axis are two marginal variables where all the combinations of these two variables along the same curve correspond to the same p. The most likely scenarios method is to peak up the combination where the joint density of this quantile curve is the highest. In this example, the *z*-axis of Figure 2 indicates the joint density level uniformed to the range of (0,1) and the blue circles indicate the location where the density level is 1.0 and the combination of MMS and MMR is regarded as the most likely one. The other commonly used approach is to sample all the combinations on the same quantile curve instead of selecting only one combination, however, this means the scenario selection has stochasticity which requires to be further analysed. In this case study, the most likely scenario approach is applied.

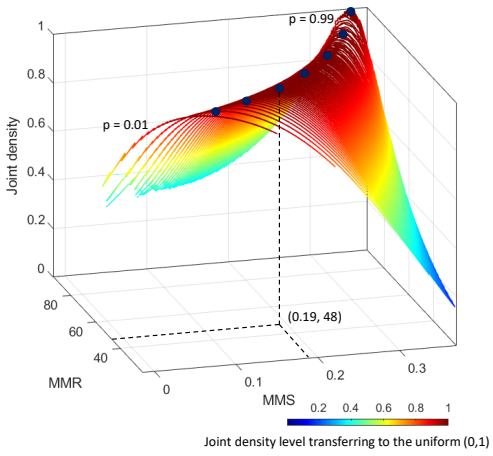

Figure 2 Joint density level of quantile curves corresponding to the probability from 0.01 to 0.99.

3. An illustrative case study

As a major economic centre of Vietnam, Ho Chi Minh City (HCMC) is located in the downstream reach of the Saigon and Dong Nai rivers (Figure 3a) with nearly 10 million inhabitants, contributing more than 20% of the GDP of Vietnam (Hallegatte et al., 2013; Kontgis et al., 2014). Yet, due to its geographical location and the ageing infrastructure, the city is vulnerable to frequent floods resulting from concurrent intense rainfall and strong surges, commonly known as compound flooding (Binh et al., 2019; Nguyen et al., 2019; Vachaud et al., 2019). The rising sea levels have been driving the threats of compound flooding to an even higher level with further complications. Therefore, it is imperative to apply a quantitative nonstationary framework to evaluate compound floods taking climate change into account.

Two observed datasets are applied in this study, i.e., daily rainfall collected from the six rain gauges in the vicinity of HCMC and hourly sea level at the estuary of the Vung Tau water level gauge (see Figure 3a). Both datasets cover a period of 38 years from 01/01/1980 to 31/12/2017 and are provided by the Southern Regional Hydrometeorological Center, Vietnam. The gauged rainfall is firstly converted to areal rainfall by applying the Thiessen polygon method and the surge is calculated as the difference between the highest observed sea level and high tide within a tidal period, where the details of data processing are given in Couasnon et al. (2022). Finally, the monthly maxima from the series of daily areal rainfall (henceforth, MMR) and of daily surge (henceforth, MMS) are extracted before being used to estimate the dependence and risk, again, in a seasonal fashion. Boxplots of MMS and MMR are depicted in Figure3b where MMS shows much fewer variations compared with MMR which has a strong dry-wet season variation. As to the seasonal variation, the extreme cases in MMR (i.e., daily rainfall higher than 90 mm) occur frequently in the wet season (i.e., August, September, October); in comparison, the extreme cases of skew surge appear mainly in the three dry months (February, March, April) and one wet month (July, which has the largest deviation (around 0.4m) from the 5th to 95th percentiles).

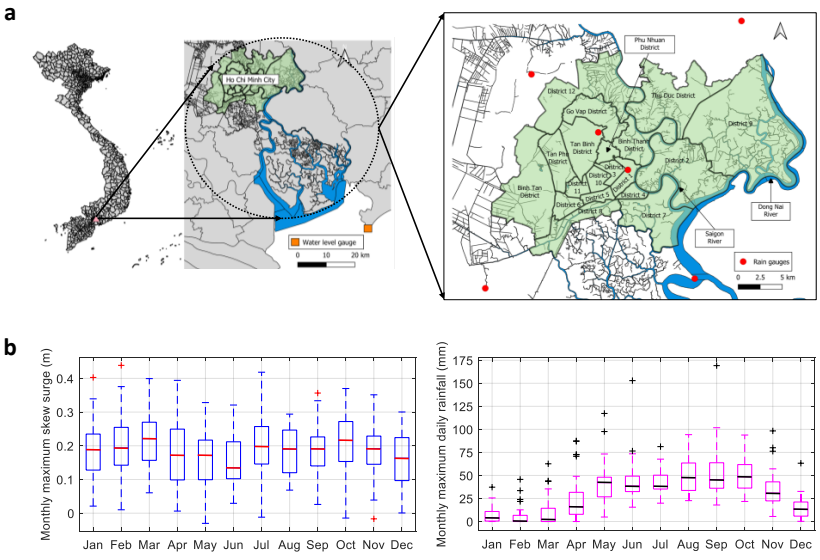

Figure 3 (a) The study area (Ho Chi Minh City) where the six rain gauges for collecting the daily rainfall over the city centre are denoted by red circles and a water level gauge by a yellow rectangle and (b) boxplots of monthly maximum time series of surge and rainfall (1980-2017). In each box, the central mark indicates the median, and the bottom and top edges of the box indicate the 25th and 75th percentiles, respectively, and the outliers are indicated by a '+'.

3.1 Trend and time-varying correlation of MMS and MMR

The first step in the application of the framework (see Figure 1) is to detect whether the series of MMS and MMR and their correlation structure vary with time. This is achieved by using the Block Bootstrapping Mann-Kendall test, and the Rolling Window combined with Kendall and Spearman test respectively. The results (see Table 2) show that MMS of all months but May were detected to be increasing during the 38-year period at a significance level of 0.05 and there is a small difference in the magnitude of these positive trends between the dry and wet seasons of HCMC. For MMR, there are only one month of dry season and two out of the eight months of wet season (May-November), i.e., April, October and November showing a significant trend over time. Regarding the correlation between MMS and MMR, the two dry months March and April witnessed a significant, positive correlation and such correlation varies with time (see Figure 4).

Table 2 Test and estimation results of monthly maximum rainfall (MMR) and skew surge (MMS) in HCMC.

Month	Time series	BBS-MK test		Correlation test (with all datasets)				Best-fitted marginal distribution		Best-fitted copula	
		Kendall's tau	*p*-value	τ	*p*-value	ρ	*p*-value	Type	**θ**	Copula	θ_C
Jan	MMS	0.34	**0.003**	0.10	0.379	0.13	0.427	GEV	S	Independence	S
	MMR	0.11	0.326					Gamma	S		
Feb	MMS	0.45	**0.000**	0.09	0.445	0.14	0.403	GEV	S	Independence	S
	MMR	0.08	0.491					Gamma	S		
Mar	MMS	0.35	**0.002**	0.21	0.062	0.34	**0.038**	GEV	S	Gaussian	NS
	MMR	0.17	0.137					Gamma	S		
Apr	MMS	0.38	**0.001**	0.27	**0.017**	0.36	**0.026**	GEV	NS1	Clayton	NS
	MMR	0.47	**0.000**					GEV	NS1		
May	MMS	0.16	0.167	0.02	0.860	0.04	0.823	GEV	S	Independence	S
	MMR	-0.01	0.940					GEV	S		
Jun	MMS	0.41	**0.001**	-0.08	0.490	-0.10	0.559	LogN	S	Plackett	S
	MMR	-0.08	0.497					GEV	S		
Jul	MMS	0.49	**0.000**	-0.08	0.469	-0.10	0.562	GEV	S	Independence	S
	MMR	-0.06	0.615					LogN	S		
Aug	MMS	0.49	**0.000**	0.09	0.453	0.14	0.392	GEV	S	Raftery	S
	MMR	-0.04	0.763					LogN	S		
Sep	MMS	0.44	**0.000**	0.07	0.532	0.10	0.568	GEV	S	Raftery	S
	MMR	0.20	0.083					GEV	S		
Oct	MMS	0.35	**0.002**	0.16	0.168	0.24	0.140	GEV	S	Joe	S
	MMR	0.23	**0.039**					GEV	NS1		
Nov	MMS	0.25	**0.031**	0.13	0.257	0.18	0.276	GEV	S	Joe	S
	MMR	0.22	**0.050**					GEV	NS3		
Dec	MMS	0.38	**0.001**	0.21	0.059	0.29	0.075	GEV	S	Raftery	S
	MMR	0.22	0.053					GEV	S		

Noted that GEV, Gamma, LogN are short for “Generalized extreme value distribution”, “Gamma distribution” and “Log-normal distribution” respectively while S indicates stationary (constant) assumption and NS indicates the nonstationary (time-varying) assumption.

Figure 4 presents the p-value of both correlation tests using three different widths of the window in March and April, indicated by different colours. It can be observed that there is little difference between the results tested by the two methods (Kendall and Spearman). Although the correlation tested on entire data is significant, it can change over time and the most correlated periods can be viewed. The correlation between MMS and MMR in March weakens continuously initially before getting strengthened during the period around 1983-2012 to 1986-2015 and then decreases again in the final periods of 1987-2016 and 1988-2017. In April, the correlation in the first period (1980-2009) is strong and becomes weak in the periods afterwards.

It is rather intriguing to see such a stronger and time-varying correlation between MMS and MMR in March and April. One of the most plausible reasons is that both MMS and MMR are directly affected by the easterly wind flow. This flow blows perpendicularly toward the coastal area in the South of Vietnam only in these two months such that it strongly stresses the surface layer water into the mainland. In addition, rainfall in these two months comes mainly from the perturbation and moisture from the easterly wind that also facilitates convection. However, in other dry months, the wind gradually moves from easterly to north-easterly and parallel to the coastal area; while in wet months, this easterly component retreats to the middle and the northeast of Vietnam's East Sea when the rainfall events in Vietnam are dominated by the summer monsoon from the Bay of Bengal.

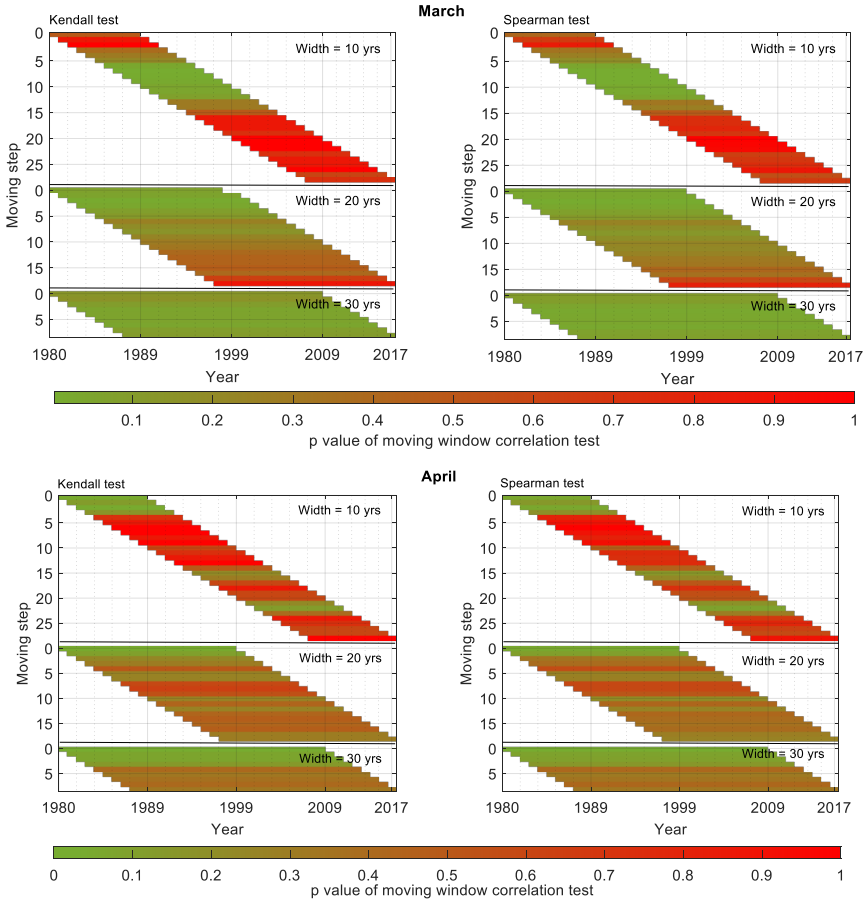

Figure 4 p-value of two rolling window correlation tests (Kendall test and Spearman test) between monthly maxima of daily rainfall and skew surge in March and April with different width of window.

3.2 Stationary and nonstationary joint probability analysis

Following the rest three steps (steps 3-5, Figure 1) of the framework, the joint probability of MMS and MMR is estimated using the most suitable candidates of marginal distributions and copulas. For the case that both the

marginal distribution and copula are best fitted by stationary models, e.g., July, the result is shown to be the same as that from the traditional multivariate probability analysis. However, when either the marginals or copula is best fitted nonstationarily, the features of quantiles corresponding to joint exceedance probabilities are different. For example, October is best estimated by a time-varying distribution of MMR but the correlation structure does change significantly over time. Figure 5 demonstrates how the quantile curves change over time with only a single distribution varying with time where it can be observed that the time-varying marginal distribution can only cause an upward or downward movement of the quantile curves at each exceedance probability, but the shape of these curves remains unchanged. However, the shape of the quantile curves can be affected by a time-varying copula, which is demonstrated by the case of March whose marginal distributions are constant, but the correlation structure varies with time. The results show that the changes in shapes are more significant in the middle than in the tails. Furthermore, April has both time-varying marginals and correlation structure and the results show that the quantile curves twist over time where such change is translational at the lower tail of quantile curves (e.g., the combination of the same skew surge with higher rainfall) while at the middle, the angle of the curves shrinks (the combination of both lower skew surge and rainfall).

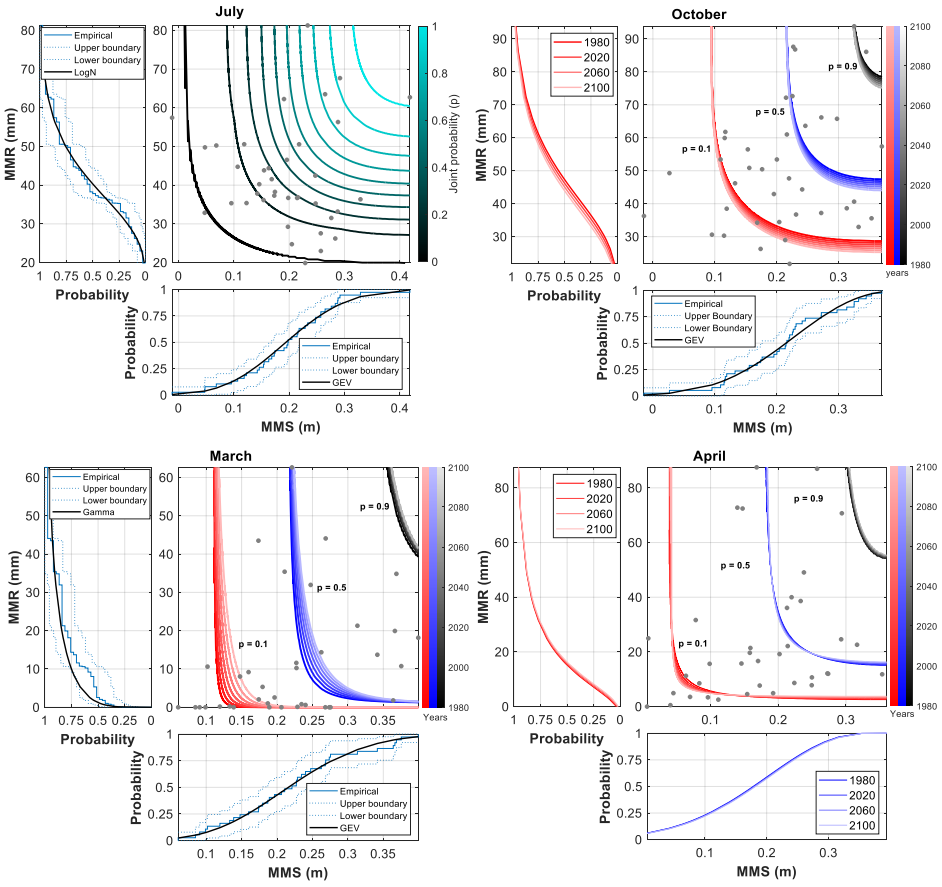

Figure 5 Quantiles corresponding to different joint cumulative probabilities (p = 0.1, 0.5 and 0.9) of the monthly maximum skew surge and rainfall in the selected four months where the best-fitted distributions of the two marginals (MMS and MMR) are shown in the left and lower panels of each sub-figure respectively: when the best-fitted distribution is stationary, the comparison between the empirical and best-fitted distribution is shown while it is nonstationary, only the best-fitted distribution is plotted and colour indicates the changes over time (year).

There are several approaches to selecting the scenarios of the combination of marginals (e.g., rainfall and skew surge in this case) and the most used approach is to get the scenarios under the most-likely compound events. In other

words, the combination with the highest joint density level is the most likely scenario. The other commonly used approach is to sample many scenarios on the same quantile curve instead of selecting only one combination, however, it means the scenario selection has stochasticity which requires to be further analysed. Therefore, in this study, we only used the most likely scenario to be the expected combination and fed it to the hydrodynamic model. We first calculated the joint density of all quantile curves shown in Figure 5 and the designed pair for each case is presented in Table 3.

Table 3 the cases for compound flood simulation.

<table>
<tr><td rowspan="2">Dry Season (Mar)</td><td colspan="6">Return level: 1-in-50 years</td></tr>
<tr><td colspan="2">Time series</td><td>Nonstationary</td><td colspan="2">Time series</td><td>Stationary</td></tr>
<tr><td rowspan="2">2020</td><td rowspan="2">Case 1</td><td>S (m)</td><td>0.49</td><td rowspan="3">Case 3</td><td>S (m)</td><td>0.52</td></tr>
<tr><td>R (mm)</td><td>93.36</td><td rowspan="2">R (mm)</td><td rowspan="2">90.87</td></tr>
<tr><td rowspan="2">2050</td><td rowspan="2">Case 2</td><td>S (m)</td><td>0.43</td></tr>
<tr><td>R (mm)</td><td>117.58</td><td colspan="2"></td></tr>
<tr><td>Wet Season (Oct)</td><td colspan="2">Time series</td><td>Nonstationary</td><td colspan="2">Time series</td><td>Stationary</td></tr>
<tr><td rowspan="2">2020</td><td rowspan="2">Case 4</td><td>S (m)</td><td>0.37</td><td rowspan="3">Case 6</td><td>S (m)</td><td>0.37</td></tr>
<tr><td>R (mm)</td><td>114.51</td><td rowspan="2">R (mm)</td><td rowspan="2">106.81</td></tr>
<tr><td rowspan="2">2050</td><td rowspan="2">Case 5</td><td>S (m)</td><td>0.37</td></tr>
<tr><td>R (mm)</td><td>113.66</td><td colspan="2"></td></tr>
</table>

4. Conclusion

This chapter presents the development of a nonstationary multivariate modelling framework for quantifying the time-varying joint probability of two meteorological and oceanographic drivers of compound flooding in view of their non-stationarity. The framework utilises the Bootstrapping Mann-Kendall trend test and the rolling window correlation tests to determine whether the drivers and/or their dependence structure should be treated in a stationary or non-stationary fashion. The Bayesian Markov-Chain Monte-Carlo (B-MCMC) method is applied to estimate the nonstationary parameters. The best marginal distribution of drivers and copula is selected by evaluating Akaike's information criterion (AIC) and Bayesian information criterion (BIC). The Ho Chi Minh City (HCMC) of Vietnam is used for a case to demonstrate the application of the framework and the result highlights the significance to incorporate non-stationarity into the statistical analysis when estimating the compound floods.

The developed nonstationary framework offers much flexibility for modelling complicated hydroclimatic extreme variables as far as the possible combinations of stationary and nonstationary assumptions are concerned. Compared with other approaches, the B-MCMC method employed in the framework is more effective to process the complex time-varying phenomena under climate change such as in estimating the correlation structure between continuous and discrete variables. Except for HCMC, other low-lying coastal cities and countries may also confront a similar predicament where a comprehensive regional risk assessment of the compound flooding potential is currently missing. This modelling framework, with the flexibility it has, will be of substantial use in this regard.

Further work is recommended to investigate compound flood and the correlation between rainfall and surge with longer-term observations which are likely to make the conclusions more robust. And for the multivariate nonstationary framework, more copulas and types of distribution candidates can be involved alongside uncertainty quantifications. Apparently, linking climate model projections into the framework will be another important and challenging area to explore.

Reference

AghaKouchak, A., Chiang, F., Huning, L. S., Love, C. A., Mallakpour, I., Mazdiyasni, O., Moftakhari, H., Papalexiou, S. M., Ragno, E., Sadegh, M. 2020. Climate extremes and compound hazards in a warming world. Annual Review of Earth and Planetary Sciences, 48, 519-548.

Ausin, M. C., Lopes, H. F., Analysis, D. 2010. Time-varying joint distribution through copulas. Computational Statistics, 54(11), 2383-2399.

Bevacqua, E., Maraun, D., Hobæk Haff, I., Widmann, M., Vrac, M. 2017. Multivariate statistical modelling of compound events via pair-copula constructions: analysis of floods in Ravenna (Italy). Hydrology and Earth System Sciences, 21(6), 2701-2723.

Bevacqua, E., Vousdoukas, M. I., Zappa, G., Hodges, K., Shepherd, T. G., Maraun, D., Mentaschi, L., Feyen, L. 2020. More meteorological events that drive compound coastal flooding are projected under climate change. Communications earth & environment, 1(1), 1-11.

Binh, L. T. H., Umamahesh, N., Rathnam, E. V. 2019. High-resolution flood hazard mapping based on nonstationary frequency analysis: case study of Ho Chi Minh City, Vietnam. Hydrological Sciences Journal, 64(3), 318-335.

Cancelliere, A. 2017. Non stationary analysis of extreme events. Water Resources Management, 31(10), 3097-3110.

Chebana, F., Ouarda, T. B., Duong, T. C. 2013. Testing for multivariate trends in hydrologic frequency analysis. Journal of Hydrology, 486, 519-530.

Couasnon, A., Scussolini, P., Tran, T., Eilander, D., Muis, S., Wang, H., Keesom, J., Dullaart, J., Xuan, Y., Nguyen, H. 2022. A Flood Risk Framework Capturing the Seasonality of and Dependence Between Rainfall and Sea Levels—An Application to Ho Chi Minh City, Vietnam. Water Resources Research, 58(2), e2021WR030002.

De Winter, J. C., Gosling, S. D., Potter, J. 2016. Comparing the Pearson and Spearman correlation coefficients across distributions and sample sizes: A tutorial using simulations and empirical data. Psychological methods, 21(3), 273.

Favre, A. C., El Adlouni, S., Perreault, L., Thiémonge, N., Bobée, B. 2004. Multivariate hydrological frequency analysis using copulas. Water Resources Research, 40(1).

Feng, Y., Shi, P., Qu, S., Mou, S., Chen, C., Dong, F. 2020. Nonstationary flood coincidence risk analysis using time-varying copula functions. Scientific reports, 10(1), 1-12.

Gori, A., Lin, N., Smith, J. 2020. Assessing compound flooding from landfalling tropical cyclones on the North Carolina coast. Water Resources Research, 56(4), e2019WR026788.

Hallegatte, S., Green, C., Nicholls, R. J., Corfee-Morlot, J. 2013. Future flood losses in major coastal cities. Nature Climate Change, 3(9), 802-806.

Hendry, A., Haigh, I. D., Nicholls, R. J., Winter, H., Neal, R., Wahl, T., Joly-Laugel, A., Darby, S. E. 2019. Assessing the characteristics and drivers of compound flooding events around the UK coast. Hydrology and Earth System Sciences, 23(7), 3117-3139.

Inoue, A., Jin, L., Rossi, B. 2017. Rolling window selection for out-of-sample forecasting with time-varying parameters. Journal of Econometrics, 196(1), 55-67.

Kontgis, C., Schneider, A., Fox, J., Saksena, S., Spencer, J. H., Castrence, M. 2014. Monitoring peri-urbanization in the greater Ho Chi Minh City metropolitan area. Applied Geography, 53, 377-388.

Kumbier, K., Carvalho, R. C., Vafeidis, A. T., Woodroffe, C. D. 2018. Investigating compound flooding in an estuary using hydrodynamic modelling: a case study from the Shoalhaven River, Australia. Natural Hazards and Earth System Sciences, 18(2), 463-477.

Kundzewicz, Z. W., Robson, A. 2004. Change detection in hydrological records—a review of the methodology/revue méthodologique de la détection de changements dans les chroniques hydrologiques. Hydrological Sciences Journal, 49(1), 7-19.

Kwon, H. H., Lall, U. 2016. A copula-based nonstationary frequency analysis for the 2012–2015 drought in California. Water Resources Research, 52(7), 5662-5675.

Manly, B. F. 2018. Randomization, Bootstrap and Monte Carlo Methods in Biology: Texts in Statistical Science: chapman and hall/CRC.

Metropolis, N., Ulam, S. 1949. The monte carlo method. Journal of the American statistical association, 44(247), 335-341.

Murthy, K. 2004. Monte Carlo Methods in Statistical Physics: Universities Press.

Nguyen, H. Q., Radhakrishnan, M., Bui, T. K. N., Tran, D. D., Ho, L. P., Tong, V. T., Huynh, L. T. P., Chau, N. X. Q., Ngo, T. T. T., Pathirana, A. 2019. Evaluation of retrofitting responses to urban flood risk in Ho Chi Minh City using the motivation and ability (MOTA) framework. Sustainable cities Society, 47, 101465.

Önöz, B., Bayazit, M. 2012. Block bootstrap for Mann–Kendall trend test of serially dependent data. Hydrological Processes, 26(23), 3552-3560.

Rasmussen, C. E., Ghahramani, Z. 2003. Bayesian monte carlo. Advances in neural information processing systems, 505-512.

Renard, B., Lang, M. 2007. Use of a Gaussian copula for multivariate extreme value analysis: some case studies in hydrology. Advances in Water Resources, 30(4), 897-912.

Ricke, K. L., Morgan, M. G., Allen, M. R. 2010. Regional climate response to solar-radiation management. Nature Geoscience, 3(8), 537-541.

Sadegh, M., Moftakhari, H., Gupta, H. V., Ragno, E., Mazdiyasni, O., Sanders, B., Matthew, R., AghaKouchak, A. 2018. Multihazard scenarios for analysis of compound extreme events. Geophysical Research Letters, 45(11), 5470-5480.

Sadegh, M., Ragno, E., AghaKouchak, A. 2017. Multivariate C opula A nalysis T oolbox (MvCAT): describing dependence and underlying uncertainty using a B ayesian framework. Water Resources Research, 53(6), 5166-5183.

Salvadori, G., De Michele, C. 2004. Frequency analysis via copulas: Theoretical aspects and applications to hydrological events. Water Resources Research, 40(12).

Salvadori, G., Tomasicchio, G., D'Alessandro, F. 2014. Practical guidelines for multivariate analysis and design in coastal and off-shore engineering. Coastal Engineering, 88, 1-14.

Sarhadi, A., Ausín, M. C., Wiper, M. P., Touma, D., Diffenbaugh, N. S. 2018. Multidimensional risk in a nonstationary climate: Joint probability of increasingly severe warm and dry conditions. Science Advances, 4(11), eaau3487.

Sklar, M. 1959. Fonctions de repartition an dimensions et leurs marges. Publ. inst. statist. univ. Paris, 8, 229-231.

Tramblay, Y., Neppel, L., Carreau, J., Najib, K. 2013. Non-stationary frequency analysis of heavy rainfall events in southern France. Hydrological Sciences Journal, 58(2), 280-294.

Vachaud, G., Quertamp, F., Phan, T. S. H., Ngoc, T. D. T., Nguyen, T., Luu, X. L., Nguyen, A. T., Gratiot, N. 2019. Flood-related risks in Ho Chi Minh City and ways of mitigation. Journal of Hydrology, 573, 1021-1027.

van den Hurk, B., van Meijgaard, E., de Valk, P., van Heeringen, K.-J., Gooijer, J. 2015. Analysis of a compounding surge and precipitation event in the Netherlands. Environmental Research Letters, 10(3), 035001.

Villalobos-Herrera, R., Bevacqua, E., Ribeiro, A. F., Auld, G., Crocetti, L., Mircheva, B., Ha, M., Zscheischler, J., De Michele, C. 2021. Towards a compound-event-oriented climate model evaluation: a decomposition of

the underlying biases in multivariate fire and heat stress hazards. Natural Hazards and Earth System Sciences, 21(6), 1867-1885.

Zar, J. H. 2005. Spearman rank correlation. Encyclopedia of biostatistics, 7.

Zhang, L., Singh, V. 2006. Bivariate flood frequency analysis using the copula method. Journal of hydrologic engineering, 11(2), 150-164.

Zhu, Y., Liu, Y., Wang, W., Singh, V. P., Ma, X., Yu, Z. 2019. Three dimensional characterization of meteorological and hydrological droughts and their probabilistic links. Journal of Hydrology, 578, 124016.

Zscheischler, J., Fischer, E. M., Lange, S. 2019. The effect of univariate bias adjustment on multivariate hazard estimates. Earth System Dynamics, 10(1), 31-43.