https://doi.org/10.5194/egusphere-2026-2118
Preprint. Discussion started: 22 April 2026

Estimating extreme sea levels using a copula joint probability method

Zhi Yang Koh[1], Benjamin S. Grandey[1], Justin Dauwels[2], Benjamin P. Horton[3], Lock Yue Chew[1]

[1]School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore
[2]Department of Microelectronics, Faculty of Electrical Engineering, Mathematics, and Computer Science, Delft University of Technology (TU Delft), The Netherlands
[3]School of Energy and Environment, City University of Hong Kong, Hong Kong

Correspondence to: Zhi Yang Koh (kohz0034@e.ntu.edu.sg)

Abstract. Extreme sea levels pose significant risks to coastal communities and infrastructure. Joint probability methods are widely used to estimate return levels of extreme sea levels by combining tidal and non-tidal components, but most implementations assume independence between tide and surge. This assumption is not always valid, and neglecting correlation can affect the estimation of return levels and their associated uncertainties. Here, we introduce a Copula Joint Probability Method (CJPM) that explicitly accounts for non-linear correlation between peak tides and skew surges, generalising the Skew Surge Joint Probability Method (SSJPM). Using long tide gauge records (≥100 years) from 23 locations, we assess how incorporating correlation affects both central estimates and confidence intervals of estimated return levels. We find that accounting for this correlation can shift estimated return levels by up to approximately 10 cm at some locations. Importantly, uncertainty in the peak tide–skew surge correlation can be a statistically significant contributor to the width of confidence intervals, in some cases exceeding the contribution from uncertainty in the extreme skew surge distribution. At other locations, correlation has a negligible effect, and CJPM and SSJPM estimates are indistinguishable. These results demonstrate that explicitly representing correlation and its uncertainty provides a more complete quantification of return levels and their associated confidence intervals, and helps determine whether correlation materially affects return level estimation. The CJPM provides a flexible framework that can be applied across a wide range of settings without requiring assumptions about the strength or cause of any correlation.

Short summary (500-character incl. spaces). Extreme sea levels pose a risk to coastal communities. We present a joint probability method that uses a copula to account for non-linear correlation between peak tides and skew surges. Accounting for this correlation can shift central estimates of return levels by up to about 10 cm. We further quantify how uncertainty in this correlation affects confidence intervals, allowing assessment of whether correlation plays a statistically significant role in return level estimation.

1 Introduction

Extreme sea levels (ESLs) pose an increasing risk to coastal communities and ecosystems (Diaz, 2016; Hinkel et al., 2014). In 2018, about 748 million people (10 % of 2018 global population) lived within 5 km of the coast and 2.86 billion (38 %) within

https://doi.org/10.5194/egusphere-2026-2118
Preprint. Discussion started: 22 April 2026

100 km (Cosby et al., 2024). In 2020, 896 million people (11 % of 2020 global population) lived in areas less than 10 m above sea level and hydrologically connected to the sea (Haasnoot et al., 2021). This number may increase beyond 1 billion by 2050 (Oppenheimer et al., 2019). Furthermore, infrastructure and assets worth 6.5–11 trillion USD are located at sites assessed to flood with a probability of at least 1 % per year (Haasnoot et al., 2021; Kirezci et al., 2020; Kulp et al., 2019). Accurate estimation and prediction of ESLs help ensure that flood defences are neither under- nor over-engineered, both of which have social and financial cost implications (Batstone et al., 2013; Tebaldi et al., 2021).

ESLs refer to the occurrence of exceptionally high or low sea levels relative to mean sea levels (Gregory et al., 2019). These deviations from the mean are caused by tides, storm surges, and waves. For many purposes, including this study, it is useful to regard sea level as the sum of three components: the mean sea level, tides, and non-tidal residuals (Pugh and Woodworth, 2014a). Tide gauges are a common source of sea level data, typically providing observations at a temporal resolution of 15 minutes to 1 hour. This temporal resolution is insufficient to capture the instantaneous effects of waves. Consequently, the term “sea level” commonly excludes these waves in studies using tide gauge data. Tides refer to the sea level component driven by astronomical forces, most notably in relation to our Moon, followed by our Sun. Tides are largely periodic and can be predicted with sufficient precision to be treated as deterministic (Pugh and Woodworth, 2014b). Non-tidal residuals are the differences obtained after subtracting mean sea level and tides from the observed sea level. Non-tidal residuals are mainly driven by meteorological effects such as storms but can also include tide–surge interaction, tide gauge errors, or tidal prediction errors. Hence, non-tidal residuals are also referred to as storm surges, or simply surges, and are treated as a stochastic component.

Various methods have been developed to estimate ESLs, based on extreme value theory (Coles, 2001), joint probability methods (Pugh and Vassie, 1978), combinations of both (Tawn, 1992; Batstone et al., 2013; Baranes et al., 2020), or mixture models (Caruso and Marani, 2022; D’Arcy et al., 2023; Palmer et al., 2024). A commonly used extreme value theory-based approach involves fitting declustered peak-over-thresholds—hourly (or more frequent), mean-detrended sea level observations exceeding a specified height threshold and occurring at least a given time interval apart—to the generalised Pareto distribution (Buchanan et al., 2016; Kopp et al., 2023). However, the peak-over-threshold-generalised Pareto distribution approach does not explicitly account for the deterministic tidal component of sea level. Ignoring this deterministic tidal component has been shown to result in biased ESL estimates (Dixon and Tawn, 1999).

In contrast, joint probability methods (JPMs) separate the deterministic tides and the stochastic surges before computing their combined distribution. In JPMs, tide and surge are treated as random variables. Their sum, the storm tide, is used as a measure of ESLs. The distribution of annual maximum storm tide is obtained after integrating the joint probability density function (PDF) of tide and surge. However, the first proposed JPM (Pugh and Vassie, 1978) has three weaknesses. First, it does not account for the temporal dependence of hourly storm tides, which is generally significantly autocorrelated. Second, it uses the empirical distributions of tide and surge for the integration, limiting the range of estimable storm tides, particularly at the extremes which are of interest. Tawn (1992) introduced the Revised JPM, which addressed the first weakness using the

https://doi.org/10.5194/egusphere-2026-2118
Preprint. Discussion started: 22 April 2026

extremal index (Leadbetter, 1983) and the second by modelling the upper tail of the surge distribution using an extreme value distribution.

The third weakness is the assumption that tide and surge are mutually independent random variables. However, in reality, tide and surge can be correlated at certain coasts due to drivers such as tide–surge interaction (Dixon and Tawn, 1994; Olbert et al., 2013) and confounding variables such as seasonality (Pirazzoli and Tomasin, 2007; Williams et al., 2016). Batstone et al. (2013) partially addressed this with the Skew Surge JPM (SSJPM), which uses the distribution of peak tide and skew surge instead of hourly tide and non-tidal residuals. Peak tide and skew surge are often approximately independent (Batstone et al., 2013; Williams et al., 2016), partially supporting this assumption of independence. However, correlation between peak tide and skew surge can still be present, particularly at sites where correlation is driven by seasonal relationships (Williams et al., 2016). Furthermore, Santamaria-Aguilar and Vafeidis (2018) identified non-seasonal correlation between peak tide and skew surge at locations within shallow continental shelves with mixed semidiurnal tidal regimes. D'Arcy et al. (2023), Palmer et al. (2024), and Baranes et al. (2020) are some recent studies that propose variations of the SSJPM that account for seasonal correlation.

In this paper, we propose an approach that is agnostic to the underlying causes of the correlation and is therefore more general and potentially applicable under a wider range of circumstances. Correlation affects not only the central estimates of predicted ESLs, but also contributes to uncertainty in these estimates. Accounting for uncertainty in correlation, which can lead to wider confidence intervals in predicted ESLs, provides a more complete characterisation of the uncertainty in predicted ESLs. Our results also show that even weak correlation between peak tide and skew surge can have a statistically significant impact on estimates of extreme annual maximum storm tides.

To address the challenge that correlation between tide and surge poses to JPMs, we propose a Copula JPM (CJPM). The CJPM builds on the SSJPM by introducing a copula to account for correlation between peak tide and skew surge. Copulas have been used to estimate joint probabilities of events such as compound flooding (Lee et al., 2024; Li et al., 2021; Yavuzdoğan and Tanır Kayıkçı, 2020; Galiatsatou et al., 2019; Masina et al., 2015) or compound weather extremes (Tavakol et al., 2020; Mesbahzadeh et al., 2019). However, their use for capturing correlation between tide and surge within JPMs has not been extensively explored. A copula is used to combine the marginal distributions of peak tide and skew surge, derived using the SSJPM, to obtain their joint PDF (Sklar, 1959; van Vliet, 2023). Copulas can take many functional forms and provide a flexible framework for representing a wide range of correlation structures, including but not limited to linear correlation. This includes correlation in the peak tide–skew surge joint PDF, regardless of the underlying drivers of said correlation (e.g. seasonal relationships and tide–surge interaction). In other words, the CJPM does not assume that peak tide and skew surge are mutually independent.

To test the robustness of the CJPM, we use long tide gauge records as they provide ample observations for validation. We apply the CJPM to tide gauge records, from the GESLA 3.0 dataset, with at least a 90 % data completion rate over ≥100 calendar years. The CJPM and SSJPM are fitted to the qualifying tide gauge records, and the resulting return level estimates

https://doi.org/10.5194/egusphere-2026-2118
Preprint. Discussion started: 22 April 2026

are compared with empirical return levels. Cross-validation is also used to compare the performance of the CJPM and the SSJPM. In addition, we test the sensitivity of the CJPM and SSJPM to the generalised Pareto distribution threshold μ.

2 Data and Methodology

2.1 Data

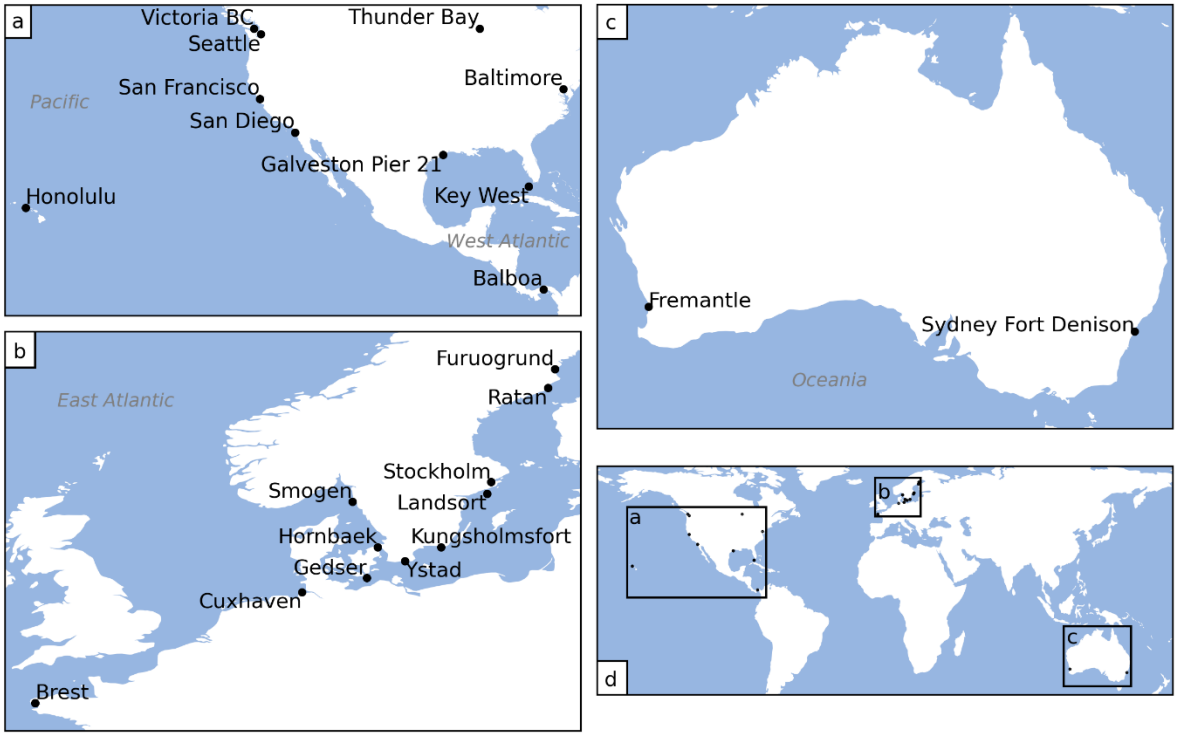

Figure 1. Locations of the tide gauges used in this study. Coordinates and record lengths of each tide gauge are listed in Table 1.

To validate the robustness of the Copula Joint Probability Method (CJPM), numerous long and high-resolution tide gauge records are required. These were obtained from the Global Extreme Sea Level Analysis (GESLA 3.0) dataset (Caldwell et al., 2015; Haigh et al., 2023; Woodworth et al., 2017). We note that the GESLA 4.0 dataset is now available and can be used for future studies. For consistency across tide gauges, records with sampling intervals shorter than hourly were averaged to produce hourly observations. No additional filtering was applied to the downsampled observations to avoid further data loss, which also aligns with the default option in the toolkit of Williams et al. (2019). To calculate unbiased annual means at each tide gauge record, observations in a calendar year were discarded if the calendar year had more than 876 missing hourly observations (i.e. below 90 % completion rate based on 8760 hourly observations per year). The 40 tide gauge records with observations remaining in at least 100 calendar years were shortlisted for this study (Table 1). These records originate from 23 unique locations (Fig. 1). At 15 of the 23 locations, two or three alternative records from different contributors are available.

https://doi.org/10.5194/egusphere-2026-2118
Preprint. Discussion started: 22 April 2026

Table 1. Tide gauge records used in this study, contributed by the Marine Environmental Data Section (MEDS), University of Hawaii Sea Level Center (UHSLC), National Oceanic and Atmospheric Administration (NOAA), Danish Meteorological Institute (DMI), Réseaux de référence des observations marégraphiques (REFMAR), Copernicus Marine Environment Monitoring Service (CMEMS), Swedish Meteorological and Hydrological Institute (SMHI), and the Bureau of Meteorology (BOM). A map of these coordinates is shown in Fig. 1. The length of usable records is given in years. The standard deviations of peak tides σ_X and skew surges σ_Y and the maximum recorded peak tide $X_{\max}$ and skew surge $Y_{\max}$ are provided in metres (m).

Continent	Country	Tide gauge name	Contributor	Lat.	Lon.	Length	σ_X	$X_{\max}$	σ_Y	$Y_{\max}$
Americas	Canada	Thunder Bay Ontario	MEDS	48.4	-89.2	104	0.08	0.33	0.11	0.64
		Victoria BC	UHSLC	48.4	-123.4	107	0.27	1.42	0.13	0.86
	Panama	Balboa	UHSLC	9.0	-79.6	108	0.46	3.26	0.10	0.69
	USA	Baltimore	NOAA	39.3	-76.6	115	0.10	0.50	0.21	2.16
		Galveston Pier 21	UHSLC	29.3	-94.8	105	0.09	0.45	0.14	2.33
		Galveston Pier 21	NOAA	29.3	-94.8	108	0.09	0.45	0.14	2.33
		Honolulu	NOAA	21.3	-157.9	101	0.17	0.64	0.05	0.62
		Honolulu Hawaii	UHSLC	21.3	-157.9	109	0.17	0.64	0.05	0.62
		Key West	NOAA	24.6	-81.8	103	0.12	0.55	0.08	0.92
		Key West Fl	UHSLC	24.6	-81.8	102	0.12	0.55	0.08	0.92
		San Diego	NOAA	32.7	-117.2	110	0.31	1.47	0.07	0.45
		San Diego Ca	UHSLC	32.7	-117.2	108	0.31	1.47	0.07	0.51
		San Francisco	NOAA	37.8	-122.5	120	0.23	1.29	0.08	0.67
		San Francisco Ca	UHSLC	37.8	-122.5	120	0.23	1.31	0.08	0.67
		Seattle	NOAA	47.6	-122.3	122	0.37	1.99	0.12	0.82
Europe	Germany	Cuxhaven	UHSLC	53.9	8.7	101	0.21	2.02	0.34	3.65
	Denmark	Gedser	UHSLC	54.6	11.9	108	0.07	0.32	0.21	1.80
		Gedser	DMI	54.6	11.9	114	0.07	0.32	0.21	1.80
		Hornbaek	UHSLC	56.1	12.5	110	0.08	0.35	0.21	1.59
		Hornbaek	DMI	56.1	12.5	117	0.08	0.35	0.21	1.75
	France	Brest	UHSLC	48.4	-4.5	152	0.64	3.88	0.13	1.37
		Brest	REFMAR	48.4	-4.5	153	0.64	3.89	0.13	1.38
		Brest 60Minute	CMEMS	48.4	-4.5	104	0.64	3.89	0.13	1.37
	Sweden	Furuogrund	SMHI	64.9	21.2	105	0.12	0.48	0.22	1.23
		Furuogrund	CMEMS	64.9	21.2	105	0.11	0.48	0.22	1.23
		Kungsholmsfort	SMHI	56.1	15.6	131	0.08	0.35	0.17	1.20
		Kungsholmsfort	CMEMS	56.1	15.6	131	0.08	0.35	0.17	1.20
		Landsort	SMHI	58.8	17.9	118	0.10	0.38	0.16	0.85
		Landsort1	CMEMS	58.8	17.9	118	0.10	0.37	0.16	0.85
		Ratan	SMHI	64.0	20.9	129	0.11	0.46	0.21	1.24
		Ratan	CMEMS	64.0	20.9	129	0.11	0.46	0.21	1.24
		Smogen	SMHI	58.4	11.2	109	0.08	0.40	0.18	1.47
		Smogen	CMEMS	58.4	11.2	109	0.08	0.39	0.18	1.47
		Stockholm	UHSLC	59.3	18.1	125	0.10	0.39	0.16	1.05
		Stockholm	SMHI	59.3	18.1	132	0.10	0.39	0.16	1.05
		Stockholm	CMEMS	59.3	18.1	132	0.10	0.39	0.16	1.05
		Ystad	SMHI	55.4	13.8	100	0.08	0.34	0.18	1.66
		Ystad	CMEMS	55.4	13.8	100	0.08	0.34	0.18	1.65
Oceania	Australia	Fremantle	BOM	-32.1	115.7	106	0.14	0.67	0.14	1.01
		Sydney Fort Denison	BOM	-33.9	151.2	102	0.21	1.16	0.08	0.54

https://doi.org/10.5194/egusphere-2026-2118
Preprint. Discussion started: 22 April 2026

These records, despite originating from the same location, can vary in length. Our results at these locations were generally similar across alternative tide gauge records. Therefore, for each location, we present results based on the record with the most calendar years of observations. In case of ties, results based on records provided by the National Oceanic and Atmospheric Administration (NOAA) and the Swedish Meteorological and Hydrological Institute (SMHI) are presented.

Hourly storm tide observations were obtained by subtracting the annual mean of the hourly tide gauge observations from the tide gauge records (Rasmussen et al., 2018). Hourly tidal observations for each year were derived using the *MATLAB* tidal harmonic-analysis package *UTide* (Codiga, 2011). To prevent double counting, a tidal cycle was defined as the interval between one local minimum in the hourly tidal observations and the observation preceding the next local minimum (Koh et al., 2024). Within each tidal cycle, the maximum tide and storm tide were identified, and their difference was defined as the skew surge (Batstone et al., 2013). The standard deviations of peak tide and skew surge observations at each tide gauge record are tabulated in Table 1, along with their maximum recorded values.

2.2 Background on Joint Probability Methods

In general, JPMs require the joint probability distribution f_{XY} of tide X and surge Y (Pugh and Vassie, 1978). In the Skew Surge JPM (SSJPM), X and Y represent peak tide and skew surge respectively. Storm tide Z is treated as a random variable and assumed to be the sum of random variables X and Y, i.e. $Z = X + Y$, with the cumulative distribution function (CDF) of Z calculated as

$$F_Z(z) = \int_{-\infty}^{\infty} \int_{-\infty}^{z-x} f_{XY}(x, y)\, dy\, dx. \quad (1)$$

With sufficient observations of X, a good estimator of $F_Z(z)$ is (Tawn, 1992)

$$F_Z(z) = \left(\prod_{t=1}^{N} F_{Y|X}(z - X_t | x = X_t) \right)^{\frac{1}{N}}, \quad (2)$$

where X_t is the tth observation of X when ordered chronologically while the index t numbers tidal cycles from the earliest to the most recent. N is the number of observations of X, and $F_{Y|X}$ is the conditional CDF of Y given X. However, due to the difficulties associated with obtaining f_{XY} and $F_{Y|X}$, X and Y are often assumed to be independent random variables. Their assumed independence implies that $F_{Y|X}$ is equal to F_Y, the marginal CDF of Y, thereby replacing the challenging task of obtaining $F_{Y|X}$ with the simpler task of determining the univariate marginal distribution F_Y.

In the earliest JPM (Pugh and Vassie, 1978), one year of observations was considered sufficient to estimate the marginal distributions of X and Y. Since Tawn (1992), the preferred estimator of F_Y is the piecewise function

$$F_Y(y) = \begin{cases} \hat{F}_Y(y) & \text{for } y \le \mu, \\ G(y)\left(1 - \hat{F}_Y(\mu)\right) + \hat{F}_Y(\mu) & \text{for } y \ge \mu, \end{cases} \quad (3)$$

https://doi.org/10.5194/egusphere-2026-2118
Preprint. Discussion started: 22 April 2026

where the empirical CDF of surge $\hat{F}_Y$ is used below a threshold μ, while the upper tail (i.e. above the threshold μ) is modelled using an extreme value distribution G. Since Batstone et al. (2013), the preferred parametric distribution for G is the generalised Pareto distribution (GPD), defined by the CDF

$$G(y) = 1 - \left(1 + \xi \frac{y - \mu}{\sigma}\right)^{-\frac{1}{\xi}}, \quad (4)$$

where ξ is a shape parameter and σ is a scale parameter. In our implementation, μ was selected to be the level between the 97.5th (Batstone et al., 2013) and the 99.7th (Baranes et al., 2020) percentiles of skew surge observations—common thresholds in studies where ≥100 years of records are available—for which the GPD provided the best fit (according to the Kolmogorov-Smirnov test). A sensitivity test showed that the SSJPM and CJPM exhibit similar sensitivity to μ within this percentile range. Bootstrapping was used to obtain the confidence intervals of ξ and σ, which represent uncertainty in the height of extreme skew surges and their distribution. While only the empirical CDF of skew surge, $\hat{F}_Y$, is required for the SSJPM, we also define the empirical CDFs of peak tide, $\hat{F}_X$, and storm tide, $\hat{F}_Z$, as

$$\hat{F}_X(X_i) = \frac{i}{N+1}, \quad \hat{F}_Y(Y_j) = \frac{j}{M+1}, \quad \hat{F}_Z(Z_k) = \frac{k}{L+1}, \quad (5)$$

where X_i is the ith observation of X when ordered in ascending order, i is the rank of X_i, and N is the number of observations of X. Y_j, j, and M are defined analogously for skew surge Y, and Z_k, k, and L for storm tide Z. Note that the X_i's referenced here and the X_t's in Eq. (2) are the same collection of peak tides. Their difference is that X_i's are ordered by magnitude and X_t's by time. Equations (2)—(5) are combined to obtain F_Z. In practical applications, it is useful to convert F_Z into a function of return level z and return period RP (in years) via the transformation

$$RP(z) = \frac{1}{1 - F_Z^{T\theta(z)}(z)}, \quad (6)$$

where T is the average number of observations of Z per year and θ is the extremal index. The extremal index, θ, accounts for autocorrelation by reducing T to the effective average number of *independent* observations of Z per year (Leadbetter, 1983). Similarly, the empirical return level of storm tides can be estimated as $\widehat{RP}(Z_k) = \left[1 - \hat{F}_Z^{T\theta(Z_k)}(Z_k)\right]^{-1}$. Given L observations of Z over D years, T can be estimated using $T = L/D$. $1/\theta(z)$ can be interpreted as the mean cluster size—the average length of time that Z remains above a threshold z before subsiding below z. We used the estimator (Ferro and Segers, 2003; Baranes et al., 2020)

$$\frac{1}{\hat{\theta}(z)} = \frac{2\left[\sum_{k=1}^{E(z)-1}(I_k - 1)\right]^2}{(E(z)-1)\sum_{k=1}^{E(z)-1}[(I_k - 1)(I_k - 2)]}, \quad (7)$$

where $E(z)$ is the number of observations above z and I_k is the interexceedance time. The estimator $1/\hat{\theta}(z)$ was then fit to the functional form

$$\frac{1}{\theta(z)} = 1 + ae^{-bz}, \quad (8)$$

https://doi.org/10.5194/egusphere-2026-2118
Preprint. Discussion started: 22 April 2026

with $a, b \geq 0$.

2.3 Copula Joint Probability Method (CJPM)

Assuming statistical independence between X and Y may not be reasonable at certain locations (Santamaria-Aguilar and Vafeidis, 2018; Williams et al., 2016), potentially causing the SSJPM to be unsuitable at those locations. If the joint PDF f_{XY} can be obtained, then the assumption of independence between X and Y need not be made. To model f_{XY}, we propose using a copula. Introduced by Sklar (1959) (translated by van Vliet, 2023), Sklar's theorem states that, in the bivariate case, every joint CDF F_{XY} can be expressed in terms of its marginals and a copula C such that

$$F_{XY}(x, y) = C\big(F_X(x), F_Y(y)\big). \quad (9)$$

The copula C itself is a valid multivariate CDF that supports marginals uniform in the interval [0,1]. Taking the second-order mixed partial derivative of Eq. (9) yields

$$f_{XY}(x, y) = c\big(F_X(x), F_Y(y)\big) f_X(x) f_Y(y), \quad (10)$$

where $c = \frac{\partial^2 C}{\partial F_X \partial F_Y}$ is called the copula density of C. One interpretation is that the copula density $c\big(F_X(x), F_Y(y)\big)$ combines the marginal PDFs f_X and f_Y to form the joint PDF f_{XY} via the product of the three terms. The CJPM consists of three overall steps: fitting the marginals following the SSJPM using Eq. (3)–(5), fitting the joint peak tide–skew surge observations to a copula as described in the following paragraph, and finally combining the copula with Eq. (1) to obtain Eq. (11)–(12) and consequently $F_Z(z)$. These steps and those of the SSJPM are shown using a schematic diagram, where their differences can be visualised (Fig. 2).

To fit our observations to a copula, we require *joint observations* of peak tide and skew surge (X_t, Y_t)—observations of pairs of X and Y measured during the same tidal cycle t. After computing the marginals $\hat{F}_X$ and F_Y, we transformed the joint observations to pseudo observations—observations where marginals are uniformly distributed but may not be jointly uniform—using the probability integral transform $(U_t, V_t) = \left(\hat{F}_X(X_t), F_Y(Y_t)\right)$. Using the Python library *pyvinecopulib* (Nagler and Vatter, 2025), we fitted the Transformation local likelihood (TLL) copula (Geenens et al., 2017) to the pseudo observations (U_t, V_t) using maximum likelihood estimation. Bootstrapping was used to obtain the confidence intervals of the TLL copula's parameters, which represent uncertainty in peak tide–skew surge correlation. The TLL copula is a non-parametric local likelihood density estimator, adapting locally to the empirically observed distribution, making it suitable for representing both central and tail dependence structures without assuming a parametric form. Mathematical details of the TLL copula are provided in Appendix A. These properties make the TLL copula suitable for the CJPM for two reasons.

First, just like observations of Z, observations of (X_t, Y_t) and, consequently, (U_t, V_t) are generally significantly autocorrelated. In JPMs, we preserve the autocorrelation in (X_t, Y_t) and make corrections to Z using the extremal index θ. Local likelihood density estimators have shown some success at handling autocorrelated observations (Geenens and Wang, 2018).

https://doi.org/10.5194/egusphere-2026-2118
Preprint. Discussion started: 22 April 2026

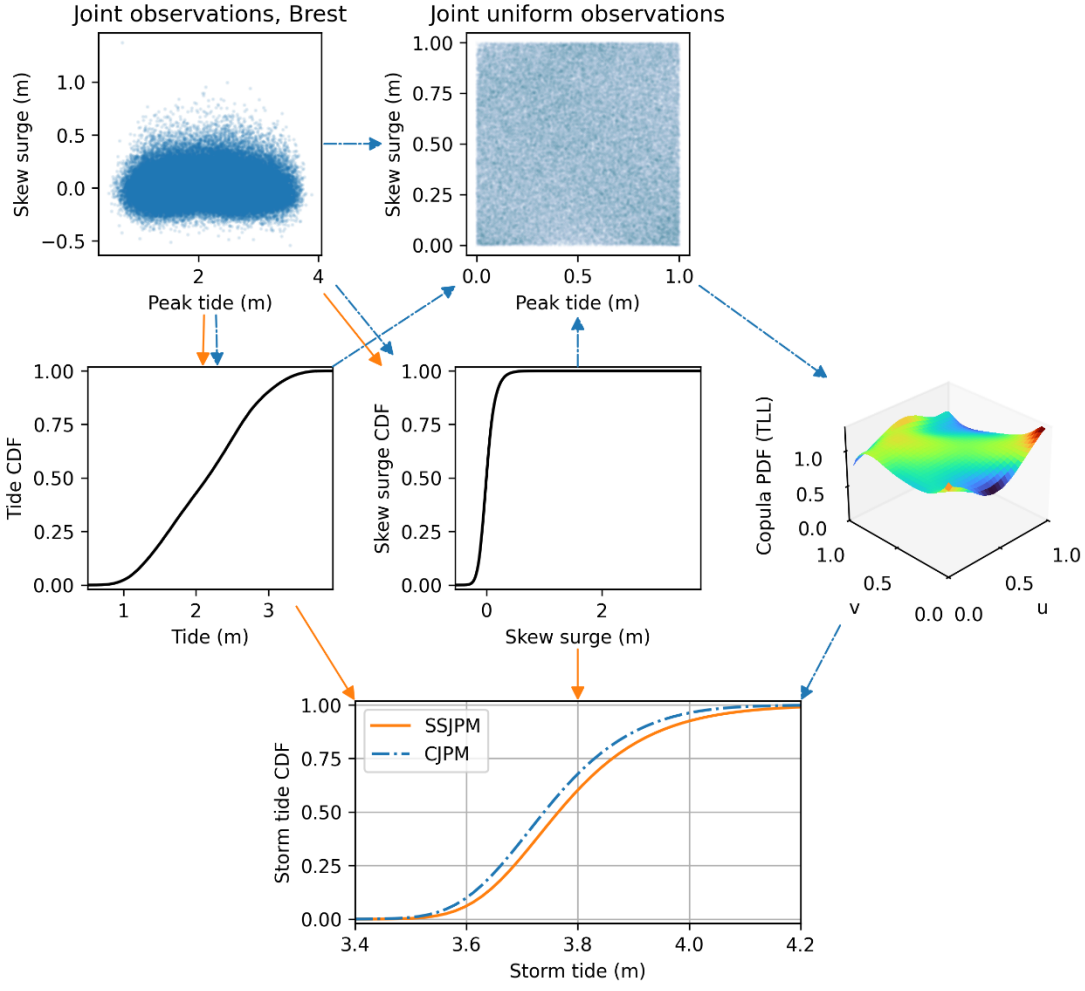

Figure 2. Schematic of the SSJPM (orange solid) and CJPM (blue dash-dotted), illustrated using tide gauge records and intermediate results from Brest, France. In the SSJPM, the marginal distributions of peak tide and skew surge are fitted to observations, then convoluted to obtain the storm tide distribution. In the CJPM, the marginal distributions are fitted using the same procedure as the SSJPM. Using the marginal distributions, the joint observations are transformed to joint uniform observations to which the TLL copula is fitted. The storm tide distribution is subsequently obtained using Eq. (12).

Second, the methodology of JPMs requires the full PDF f_{XY}, although we are most interested in the extremes of Z (e.g. return periods of at least 10 years). Certain studies into the correlation between tides and surges address this by conditioning on large values of Z (Arns et al., 2020; Ragno et al., 2024). However, we suspect this introduces collider bias into analyses. We further explain collider bias and discuss its implications in the discussion section. Hence, we require a copula such as the TLL that can represent the bulk of the observations while also capturing the characteristics in the upper tails. Local likelihood density estimators have also been shown to outperform conventional kernel density methods at the copula boundaries and the tails of the joint distributions (Geenens, 2014; Geenens et al., 2017; Geenens and Wang, 2018).

https://doi.org/10.5194/egusphere-2026-2118
Preprint. Discussion started: 22 April 2026

The requirement for a copula to fit the full PDF of f_{XY} despite our primary interest being in the extremes of Z also complicates the use of traditional measures to evaluate the copula's goodness-of-fit. This is because those measures are dominated by the bulk of the observations in the central tendency, while extremes make up a small proportion of these observations. Even in cases where the TLL was deemed a worse overall fit to observations than one- and two-parameter parametric copulas, our testing showed that the TLL outperformed parametric copulas at estimating return levels (see the discussion in Sect. 4 relating to Fig. 4 and Table 2).

Combining Eq. (1) and (10) and making the change of variables $u = F_X(x)$ and $v = F_Y(y)$, we obtain

$$F_Z(z) = \int_0^1 \int_0^{F_Y\left(z - F_X^{-1}(u)\right)} c(u, v)\, dv\, du. \quad (11)$$

With sufficient observations of X, a good estimator of Eq. (11) is

$$F_Z(z) = \left(\prod_{t=1}^{N} C_{V|U}\left(F_Y(z - X_t) | u = \hat{F}_X(X_t)\right)\right)^{\frac{1}{N}}, \quad (12)$$

where $C_{V|U}$ is the conditional CDF of V given U, and N is the number of observations of X. Note the similarities between Eq. (2) and (12), where $F_{Y|X}$ is replaced by $C_{V|U}$, Y by $F_Y(Y)$, and X by $\hat{F}_X(X)$. Equation (12) can be expressed in terms of return levels and return periods by combining Eq. (12) with Eq. (6)–(8).

2.4 Cross-validation

As the CJPM has more parameters than the SSJPM, it is important to assess whether the CJPM exhibits overfitting. Cross-validation evaluates the performance of the CJPM and SSJPM on their ability to describe the data on which they were trained and their ability to predict unseen data. In this study, we employed 4-fold cross-validation.

In the 4-fold cross-validation, each tide gauge record was shuffled and divided into four equal subsets. In each iteration, three subsets were used for training (i.e. the CJPM and SSJPM were fitted to these subsets), and the remaining subset was used for validation. During training, the parameters a and b in Eq. (8) were fixed at specified values because shuffling the tide gauge data destroys the autocorrelation represented by θ. Hence, a and b were set to the values obtained when fitting the JPMs to the full tide gauge records.

The training error for each JPM was calculated by first determining the empirical return periods of the 20 largest observations in the training set. Model-predicted return levels for these return periods were then obtained. The mean absolute difference between the 20 observations and the model-predicted return levels were defined as the training error. The validation error was calculated analogously using the 20 largest observations from the validation set instead of the training set.

This procedure was repeated for four iterations such that each subset served as the validation set once. The performance of each JPM was evaluated by aggregating the training and validation errors across all folds.

https://doi.org/10.5194/egusphere-2026-2118
Preprint. Discussion started: 22 April 2026

However, shuffling a tide gauge record breaks any seasonality that may be present in the record. To preserve the effect that seasonality may have on the performance of the CJPM and the SSJPM, we also employed an *unshuffled 4-fold cross-validation* where each tide gauge record was divided into four equal contiguous subsets without shuffling.

3 Results

We illustrate the impact of incorporating a copula into the SSJPM by highlighting results from three locations (Fig. 3). We plot return levels as a function of return period, i.e. the inverse of Eq. (6), which represents the average number of years between two instances of storm tide Z exceeding a given height. For instance, the CJPM curve passes through the coordinates (50 years, 4.05 m) at Brest (Fig. 3a), indicating that the CJPM estimates Z to exceed 4.05 m on average once every 50 years. As the estimator $1/\theta(z)$ is only valid at large values of z, we limit our discussion to return periods above 10 years.

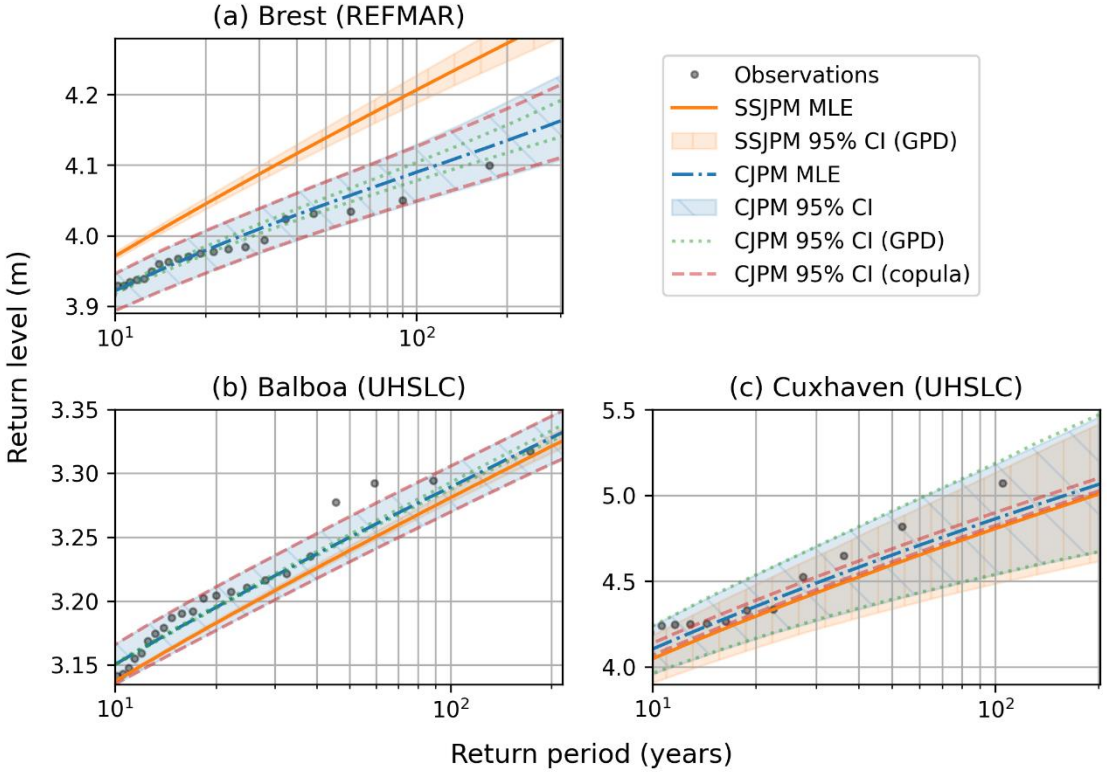

Figure 3. Return levels as a function of return period at three tide gauge locations. Grey points denote empirical return levels. The SSJPM's central estimate and its 95 % confidence interval are represented by the orange solid line and the vertically hatched shaded region respectively. The corresponding CJPM results are shown using the blue dash-dotted line and the diagonally hatched shaded region. The green dotted lines indicate the CJPM's confidence interval due to uncertainty in GPD parameters (with copula parameters held constant), while the red dashed lines indicate the CJPM's confidence interval due to uncertainty in copula parameters (with GPD parameters held constant).

https://doi.org/10.5194/egusphere-2026-2118
Preprint. Discussion started: 22 April 2026

The longest available tide gauge record is from Brest, France, with a duration of 153 years, provided by the Réseaux de référence des observations marégraphiques (REFMAR). Here, there is a clear difference between the estimates of the CJPM and the SSJPM (Fig. 3a). When compared to the empirical return levels, the SSJPM appears to overestimate ESLs at Brest. Pirazzoli and Tomasin (2007) also found that tide–surge interaction and seasonality caused the JPM to overestimate return levels at Brest, though their results were based on tide and non-tidal residuals instead of peak tide and skew surge. The SSJPM's central estimate is 10 cm higher than that of the CJPM's at a return period of roughly 60 years, further diverging at larger return periods. A change of the same order of magnitude was also reported by D'Arcy et al. (2023) when incorporating peak tide–skew surge interaction into the SSJPM and applying it to tide gauge records of Sheerness, England. Compared to the SSJPM, the CJPM shows closer visual and numerical agreement with the observed historical ESLs of Brest, with the CJPM's confidence intervals encompassing Brest's empirical return levels. In the CJPM, the uncertainty in peak tide–skew surge correlation and the uncertainty in the height of extreme skew surges contribute to uncertainty in estimated return levels. The contributions from each source of uncertainty are shown by the red dashed and green dotted lines respectively, while the blue shaded region represents their combined effect. Of the two sources of uncertainty, correlation appears to be the larger contributor to the uncertainty in Brest's estimated return levels, as indicated by the wider confidence interval. The uncertainty associated with correlation alone, and the resulting confidence interval, is sufficient to encompass the empirical return levels and thus capture the observed spread in empirical return levels. On the other hand, uncertainty in the height of extreme skew surges alone does not appear to fully account for the observed variability in Brest's empirical return levels. At a return period of 100 years, the confidence interval of the CJPM has a width of 8 cm when correlation uncertainty is included, and 3 cm otherwise. By accounting for the uncertainty in peak tide–skew surge correlation, the CJPM produces a more wider confidence interval which better agrees with empirical return levels. We found similar results at Honolulu, San Diego, San Francisco, and Thunder Bay, where correlation contributes more to uncertainty than the height of extreme skew surges when estimating return levels and the CJPM shows closer agreement with empirical return levels than the SSJPM.

Our results at Balboa, Panama, illustrate a scenario where the central estimates of the CJPM and the SSJPM do not diverge drastically, but the CJPM shows somewhat closer agreement with the empirical return levels by accounting for uncertainty in peak tide–skew surge correlation (Fig. 3b). The difference in their central estimates is less than 1.4 cm between return periods of 10 and 100 years. This difference is small when compared to the CJPM's confidence interval, which has a width of more than 3 cm and fully encompasses the SSJPM's estimates. This indicates that the SSJPM's estimates fall within the range of plausible outcomes represented by the CJPM. Similar to Brest, uncertainty in the CJPM's return level estimates is largely driven by uncertainty in peak tide–skew surge correlation, with relatively small contribution from uncertainty in the height of extreme skew surges. With the exception of two outliers, the uncertainty associated with correlation encompasses the empirical return levels, which the uncertainty associated with extreme skew surge distribution does not capture. Other locations where the CJPM and SSJPM had similar central estimates but the uncertainty in correlation was found to be at least as impactful as the uncertainty in extreme skew surge heights are Landsort, Seattle, Sydney Fort, and Victoria BC.

https://doi.org/10.5194/egusphere-2026-2118
Preprint. Discussion started: 22 April 2026

At Cuxhaven, Germany, uncertainty in the height of extreme skew surges is the dominant contributor to uncertainty in return level estimates, while uncertainty in peak tide–skew surge correlation has a smaller contribution (Fig. 3c). Here, the uncertainty associated with extreme skew surge distribution is sufficient to capture most of the variability in empirical return levels. However, with a central estimate approximately 6 cm lower than that of the CJPM, the SSJPM appears to slightly underestimate ESLs when compared against empirical return levels. Meanwhile, the empirical return levels lie fully within the CJPM's confidence intervals. Nonetheless, this difference between the CJPM and SSJPM is small relative to their confidence intervals of about 65 cm wide at the 100-year return period mark. Fremantle and Furuogrund are the other locations where peak tide–skew surge correlation does not contribute substantially to uncertainty in estimated return levels, but leads to small differences in central estimates when accounted for.

At the remaining 10 locations, the CJPM produced estimates that do not meaningfully differ from those of the SSJPM. The central estimates and confidence intervals of both the CJPM and the SSJPM were similar, and the uncertainty in peak tide–skew surge correlation had negligible influence on the uncertainty of estimated return levels.

In the 4-fold cross-validation, the CJPM and SSJPM had an overall training error of 3.4 cm and 5.1 cm respectively. This indicates that the CJPM fits the observed data better, consistent with the results shown in Fig. 3 and other tide gauge locations. The CJPM also performed better than the SSJPM in out-of-sample evaluation, with an overall testing error of 4.6 cm compared to the SSJPM's 5.7 cm. The increase in error between the training and testing datasets of 1.3 cm for the CJPM is larger than the SSJPM's 0.6 cm, suggesting a slightly greater tendency for overfitting in the CJPM compared to the SSJPM. This is expected as the CJPM is a generalisation of the SSJPM and has more parameters, with the SSJPM being a special case of the CJPM where $c(u, v) = 1$. Nonetheless, the CJPM's lower testing error, together with existing evidence of peak tide–skew surge correlation (D'Arcy et al., 2023; Santamaria-Aguilar and Vafeidis, 2018; Williams et al., 2016), suggests that the CJPM's increased complexity could be warranted. Furthermore, the cross-validation errors were computed using only the central estimates of return levels without considering that the CJPM distinguishes correlation-associated uncertainty. Hence, the cross-validation results do not fully reflect the additional insights the CJPM provides relative to the SSJPM.

Results of the unshuffled 4-fold cross-validation are consistent with these findings. In the unshuffled cross-validation, the CJPM and SSJPM had an overall training error of 3.5 cm and 5.3 cm respectively, close to those of the shuffled cross-validation. The overall testing error is 7.8 cm and 8.8 cm for the CJPM and SSJPM respectively. With an increase of 3.2 cm and 3.1 cm respectively over the CJPM's and SSJPM's testing error in the shuffled cross-validation, these results indicate that seasonality affects the performance of both the CJPM and SSJPM to a similar degree.

4 Discussion

The CJPM generalises the SSJPM by relaxing the SSJPM's assumption of independence between peak tide and skew surge, i.e. the SSJPM is a special case of the CJPM with the constraint $c(u, v) = 1$. Our results highlight two potential advantages of the CJPM over the SSJPM. First, by accounting for the correlation between peak tide and skew surge, the CJPM's central

https://doi.org/10.5194/egusphere-2026-2118
Preprint. Discussion started: 22 April 2026

estimates of return levels were found to be in closer visual and numerical agreement with empirical return levels than the SSJPM. Second, the CJPM accounts for the uncertainty in peak tide–skew surge correlation. The isolated impact of peak tide-skew surge correlation on the uncertainty in estimated return levels can be used to determine whether peak tide–skew surge correlation causes the central estimates of the CJPM and SSJPM to deviate significantly. Combining correlation uncertainty with other sources of uncertainty can yield a wider, more conservative confidence interval in return level estimates which could more fully reflect the total uncertainty in ESL estimates.

Sources of correlation between peak tide and skew surge may include tide–surge interaction (Santamaria-Aguilar and Vafeidis, 2018; Olbert et al., 2013; Dixon and Tawn, 1994) and confounding variables such as seasonality (D'Arcy et al., 2023; Williams et al., 2016). Tide–surge interaction has been observed to occur in shallow (Idier et al., 2019; Mawdsley and Haigh, 2016) and enclosed (Costa et al., 2023; Yang et al., 2023) water bodies. These geographical characteristics are found where the Brest and Thunder Bay tide gauges are located (GEBCO Compilation Group, 2024; Kastrisios et al., 2023)—two locations where the CJPM and the SSJPM differed by more than 10 cm in their central estimates. However, the CJPM neither identifies the causes of correlation nor distinguishes between different drivers of correlation. The CJPM only leverages observed correlation to produce estimates of return levels that account for correlation. Further research will be needed to identify the presence of tide–surge or peak tide–skew surge interaction at these locations, understand the underlying dynamics, and quantify their effects. Pirazzoli and Tomasin (2007) also found indications that tide–surge interaction and seasonal correlation between tides and non-tidal residuals are present at Brest. Artefacts may also manifest in observations of tide due to errors in harmonic analysis (Horsburgh and Wilson, 2007), which can potentially induce correlation. In cases where a correlation structure between peak tide and skew surge is absent or negligible, the copula density $c \approx 1$ and the CJPM converges to the SSJPM. Additionally, if hourly tides and non-tidal residuals are independent, we expect the SSJPM and CJPM to converge to the Revised JPM of Tawn (1992). Conversely, copulas could also be applied to the Revised JPM to account for correlation between hourly tides and non-tidal residuals.

We also note that the specific value of μ, restricted to values between the 97.5th and 99.7th percentile of skew surges, did not have a statistically significant impact on return levels estimated by either the CJPM and the SSJPM. We chose this percentile range as it is typical of the thresholds used in similar studies. However, lower percentiles can also be tested and, if viable, chosen, potentially reducing the uncertainty in the distribution of extreme skew surges. Identifying an appropriate value for μ is challenging, and exploratory analyses of the underlying sea level records remain the most reliable method to do so. To our knowledge, there is no generalised or global "one-size-fits-all" algorithm for selecting μ that can be regarded as a preferred default across variants of SSJPM. This remains an important question under active investigation and several approaches have been proposed (Collings et al., 2025; Murphy et al., 2025; Belzile et al., 2023). Incorporating these methods into the CJPM, or the SSJPM in general, could be explored in future studies. When using the CJPM for local-scale analyses, we recommend choosing μ based on the aforementioned exploratory analysis. However, this approach is most likely impractical for global-scale analyses.

https://doi.org/10.5194/egusphere-2026-2118
Preprint. Discussion started: 22 April 2026

Our cross-validation results indicate that the CJPM, despite having more parameters than the SSJPM, does not substantially overfit to training data; relative to the SSJPM, the CJPM maintains comparable generalisation performance despite its increased model complexity. In the unshuffled cross-validation, results indicate that seasonality affects the performance of the CJPM and the SSJPM to a similar extent. Other studies have accounted for seasonality in JPMs by isolating the winter and summer seasons (Baranes et al., 2020), adding a seasonal covariate to the GPD parameters (D'Arcy et al., 2023), or using a mixture distribution (Palmer et al., 2024). Similar approaches could be considered for future iterations of the CJPM.

Unfortunately, our methodology and analyses require ample observations for validation. Our criterion for tide gauge records to be at least 100 years long precludes many regions from our analysis, especially those with shorter records that may be the most vulnerable to ESLs. Many more tide gauge records could have been included in this study if we reduced our requirements to, for example, 50 years. However, even our threshold of 100 years leaves us with few observations for cross-validation. With the shortest possible testing sets of 25 years, only two observations have empirical return periods above 10 years. The third to twentieth largest observations have empirical return periods between 1.25 and 8.33 years. Hence, long sea level records are needed for robust cross-validation. However, the use of century-scale tide gauge records also raises potential issues related to non-stationarity. Although annual de-meaning is applied to remove changes in mean sea level, cases where extremes are non-stationary over such timescales have been observed (Calafat et al., 2021; D'Arcy et al., 2023; Palmer et al., 2024). In addition, stationarity of the correlation between peak tide and skew surge could be of particular interest. Extending the CJPM to account for non-stationarity via approaches such as time-varying GPD and copula parameters could be an interesting direction for future research.

Quantifying the differences between the results of the SSJPM and the CJPM may be of scientific interest. The discovery of substantial and statistically significant differences could motivate further investigation into the underlying causes of the correlation between peak tide and skew surge. Practitioners can evaluate the copula-associated uncertainty and the differences in return level estimates at a location to determine whether the additional complexity of the CJPM is warranted, or if the SSJPM is adequate for their application.

The benefits of using a copula to model the correlation structure between peak tide and skew surge can be challenging to quantify. Ragno et al. (2024) also utilised copulas when estimating quantiles of storm tide in the Adriatic and Tyrrhenian Seas. However, following Arns et al. (2020), they selected joint peak tide–skew surge observations by conditioning on a storm tide threshold. Due to collider bias, also known as Berkson's paradox, such conditioning can create a false correlation between peak tide and skew surge (McElreath, 2020). Collider bias occurs when a response variable (the collider) is causally influenced by two or more explanatory variables that may not necessarily be mutually dependent, as is the case of tide, surge, and their sum, i.e. storm tide. Conditioning on a collider, such as storm tide, causes a false correlation between the explanatory variables. This false correlation can cause the benefits of using a copula to be overstated if comparisons are not done fairly and carefully. When conditioning on large storm tides, the conditioned observations are likely to have a strong negative linear correlation

https://doi.org/10.5194/egusphere-2026-2118
Preprint. Discussion started: 22 April 2026

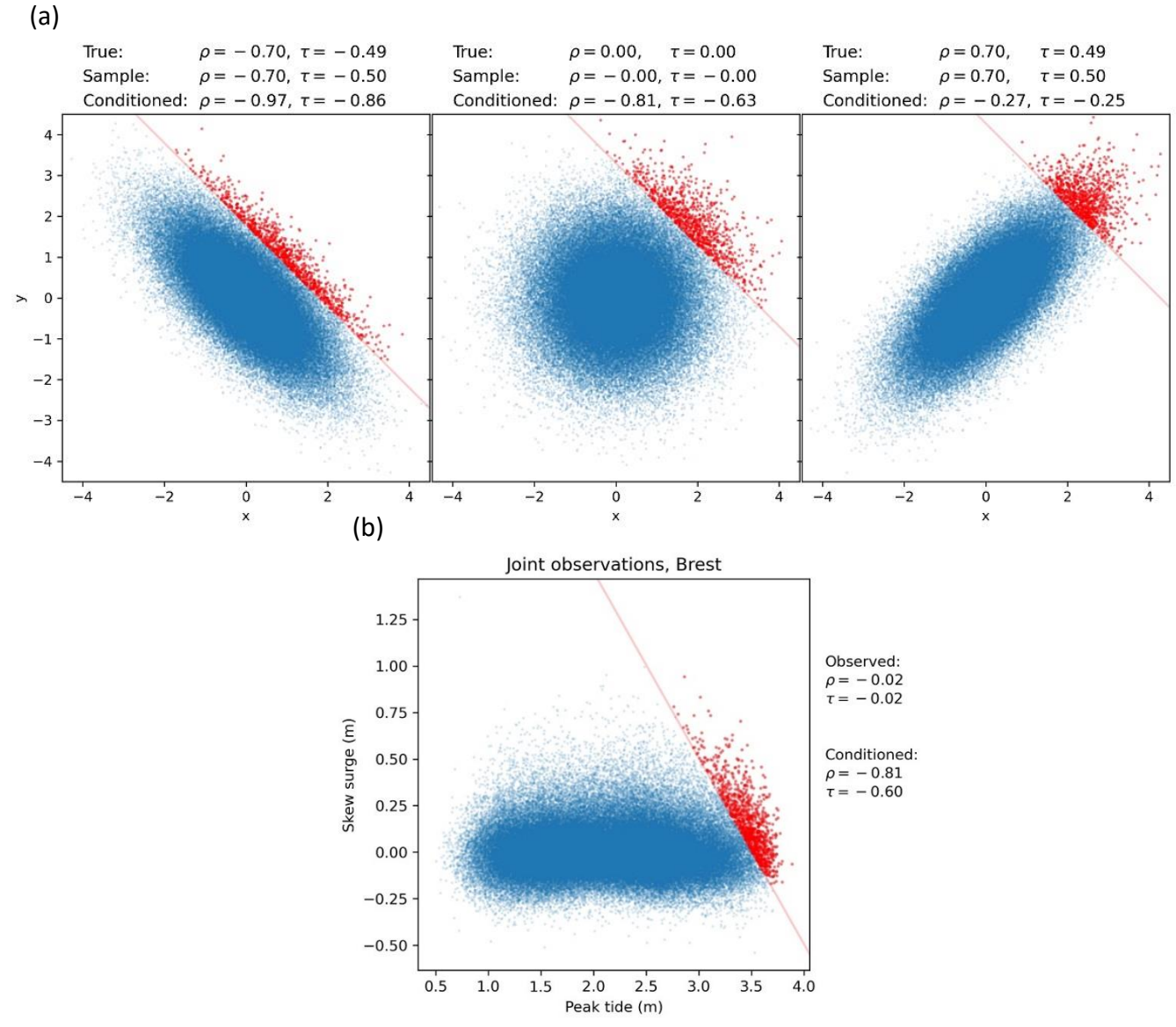

Figure 4. (a) Illustration of collider bias or Berkson's paradox using randomly generated Gaussian samples with specified correlation coefficients. Three samples, each of size 100,000, are generated from bivariate Normal distributions with standard Normal marginals and correlation coefficients of -0.7, 0, and 0.7. For each (x, y) pair, we compute $z = x + y$ and identify the 99th percentile of z within each sample. (x, y) pairs for which z is greater than or equal to the 99th percentile (i.e. conditioned data) are highlighted in red. The conditioned data can exhibit a strong negative correlation even when the underlying data do not. When working with conditioned data in isolation, care should be taken when making conclusions about the correlation between x and y. Conversely, the correlation measures of the full sample do not necessarily reflect the correlation structure at extreme values of z. (b) Illustration of collider bias or Berkson's paradox using tide gauge records from Brest, France.

regardless of the correlation (or lack thereof) in the original unconditioned observations (Fig. 4). Hence, if the SSJPM is applied to both the conditioned and unconditioned observations, return levels based on the conditioned observations are likely to be

https://doi.org/10.5194/egusphere-2026-2118
Preprint. Discussion started: 22 April 2026

overstated compared to results based on unconditioned observations. The SSJPM is not equipped to handle correlation and not designed to be applied on observations conditioned this way. For this reason, comparisons to the SSJPM should be done against results obtained from applying the SSJPM to unconditioned observations, as we have done in this paper. On the other hand, applying the CJPM to either the unconditioned or conditioned observations should yield similar results as the CJPM can account for the correlation (linear and non-linear) that may be present in either set of observations.

Collider bias may also cause measures of linear correlation between large storm tide-conditioned peak tide and skew surge to be misunderstood. Such measures, which indicate that peak tide and skew surge (or tides and surges in general) are negatively correlated at large storm tides, may be misunderstood to imply that extreme surges are more common during lower high tides (or lower tide levels) and less common during higher high tides (or higher tide levels), even if this is not necessarily the case (Fig. 4). To date, the most comprehensive set of peak tide–skew surge correlation measures is presented in Arns et al. (2020). Supplementary Figure 2 of Arns et al. (2020) also helps illustrate how collider bias occurs due to conditioning on storm tides: at the six locations shown, the Kendall rank correlation coefficient between peak tide and skew surge is close to 0 without any conditioning but ranges from -0.72 to -0.55 when conditioned on large storm tides.

Linear correlation between peak tide and skew surge in the full length of tide gauge observations tends to be close to 0 as most observations (e.g. the central tendency) are not correlated. Nevertheless, significant and genuine (not arising from collider bias) correlation structure, both linear or non-linear, may exist in some subset of the observations' domain. The existence of such correlation structure may statistically significantly influence the SSJPM, as evidenced by our results at Brest, or potentially JPMs in general. For this reason, we found that measures of linear correlation may be insufficient for assessing the importance of introducing a copula to the SSJPM for a given location or sea level record. Instead, a popular statistical method for identifying and quantifying tide–surge interaction uses timing-based measures (Costa et al., 2023; Horsburgh and Wilson, 2007), although they are not useful for modelling peak tide–skew surge correlation in JPMs. This also reflects the difficulty of defining useful measures of correlation between the magnitudes of tides and surges in the context of JPMs.

To illustrate the limitations of linear correlation measures, we provide a comparison of the TLL copula to parametric copulas using the tide gauge records from Brest. The parametric copulas used are tabulated in Table 2, where various goodness-of-fit scores and the Kendall's τ of each copula are also tabulated. The return levels and return periods of each CJPM and the SSJPM are plotted in Fig. 5. Relative to empirical return levels with return periods of at least 10 years, the TLL-based CJPM achieved a mean absolute error of 0.9 cm while the parametric copula-based CJPM had an average mean absolute error of 6.3 cm (Fig. 5). This is despite the TLL copula not necessarily being the best fit to the observations (Table 2). This may be due to uncorrelated central tendency having significant weight in the copulas' goodness-of-fit measures. This indicates that a copula deemed to fit the data better may not produce the best estimates of return levels, hence our decision to use the TLL copula. We also find that the Kendall's τ of all fitted copulas are similar and small in magnitude (Table 2) despite producing very different return levels and return periods (Fig. 5). This indicates that the Kendall's τ of peak tide and skew surge would likely be a poor indicator of the presence of any non-linear correlation structure and whether the copula-based CJPM should be adopted over the SSJPM. A more helpful assessment of whether a copula is important for a given location could be to apply both the CJPM

https://doi.org/10.5194/egusphere-2026-2118
Preprint. Discussion started: 22 April 2026

Table 2. Goodness-of-fit scores and Kendall's τ of copula families fitted to the Brest (REFMAR) tide gauge record. AIC, BIC, and mBIC refer to the Akaike, Bayesian, and modified Bayesian information criterion respectively. The best-fitting copula based on each criterion is emphasised in bold.

Copula	τ	mBIC	BIC	AIC	Log likelihood
Independence	0	4.61	0	0	0
Gaussian	-0.01	-42.60	-42.81	-52.39	27.20
Student's t	-0.02	-68.54	-68.76	-87.92	45.96
Clayton	-0.02	-218.39	-218.60	-228.18	115.09
Gumbel	-0.02	-156.77	-156.98	-166.57	84.28
Frank	-0.02	-91.05	-91.26	-100.85	51.42
Joe	-0.02	-177.38	-177.59	-187.17	94.58
Clayton-Gumbel (BB1)	-0.02	-206.80	-207.01	-226.17	115.08
Joe-Gumbel (BB6)	-0.02	-165.80	-166.01	-185.17	94.58
Joe-Clayton (BB7)	-0.02	-206.80	-207.01	-226.17	115.08
Joe-Frank (BB8)	-0.03	**-259.38**	**-259.59**	-278.75	141.38
TLL	-0.02	-118.03	-118.24	**-803.44**	**473.24**

and the SSJPM to a sea level record and determine whether the estimates of the SSJPM lie within the CJPM's confidence interval associated with peak tide–skew surge correlation. A statistically significant difference in estimates implies that the original observations were correlated, at least to a degree that can statistically significantly influence the SSJPM's results. The mean absolute error of TLL- and parametric-based CJPMs were also computed for the remaining 22 locations and compared in the same manner as at Brest. We found that the TLL-based CJPM performs similarly to the parametric-based CJPMs, with noticeable improvements at some locations including San Francisco, Seattle, Cuxhaven, and Fremantle. This suggests that while the flexibility and additional complexity of the TLL copula may not be necessary in all cases, it provides a robust and conservative option in the absence of prior knowledge of the underlying dependence structure, consistent with its qualitative properties discussed in Sect. 2.3. One potential downside is that the TLL copula can be more data-hungry than its parametric counterparts. If this is a concern, strategies such as defaulting to independence (i.e. the SSJPM) in the absence of clear evidence of dependence, selecting copula families via cross-validation with tail-focused scoring, or comparing a small set of standard parametric copulas (e.g. Gaussian, Student's t, Clayton, Gumbel, or Frank) can be considered.

https://doi.org/10.5194/egusphere-2026-2118
Preprint. Discussion started: 22 April 2026

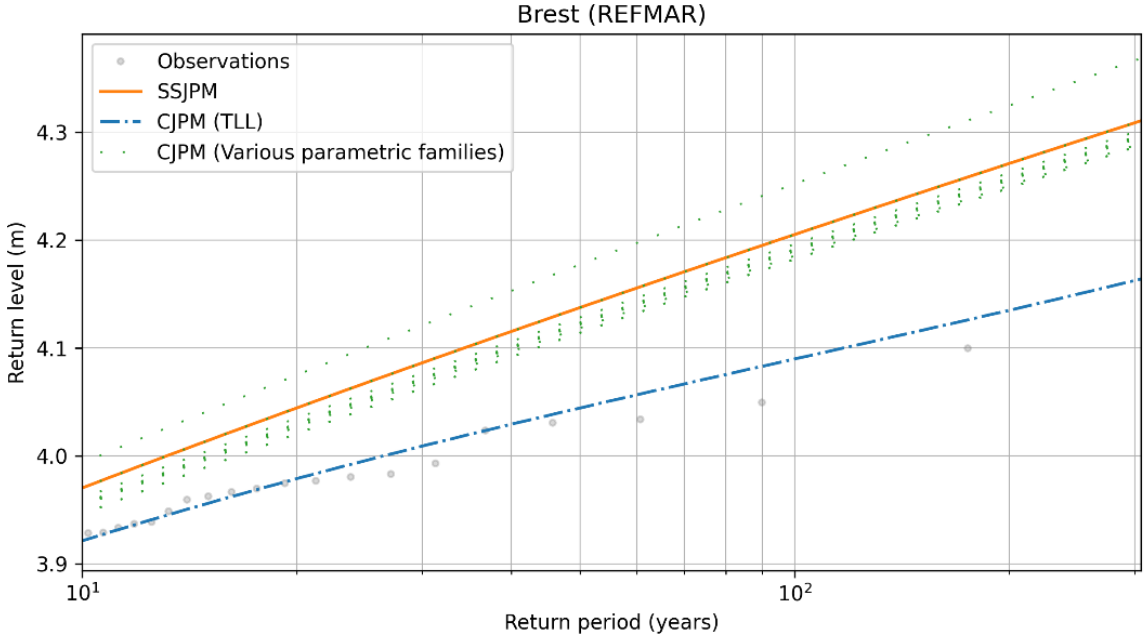

Figure 5. Return level estimates from the parametric copula-based CJPMs (green dotted), the TLL copula-based CJPM (blue dash-dotted), and the SSJPM (orange solid) at Brest. The parametric copulas used are listed in Table 2.

The CJPM relaxes the assumption that the peak tide and skew surge distributions are independent, making it more general than the SSJPM and suitable at more locations. We showed that using a copula to account for the correlation structure between peak tide and skew surge can lead to central estimates different from the SSJPM's while enabling quantification of correlation-associated uncertainty. Furthermore, the CJPM is agnostic to the underlying causes of the correlation, and fits a copula purely based on the empirical correlation structure of the underlying observations. For instance, the CJPM showed closer agreement with observations at Brest relative to the SSJPM without prior knowledge of whether the correlation between peak tide and skew surge is due to tide–surge interaction or seasonality. The agnostic nature of the CJPM makes it a good option for sites with insufficient information on the drivers of potential correlation between peak tide–skew surge and the correlation structure of their joint distribution. One such region is Southeast Asia, where tide–surge interaction is likely to be present at many coastal locations, but relatively few studies have been done to investigate its causes and how it may impact the height of storm tides and ESLs in this region (Koh et al., 2024). However, the sparse spatial distribution of tide gauges and the limited length of their records pose an additional challenge when estimating ESLs in this region. The performance of the CJPM under such conditions could be investigated in future studies. The CJPM contrasts with methods such as those proposed by Baranes et al. (2020), D'Arcy et al. (2023), and Palmer et al. (2024) that account for correlation due to seasonal relationships. On the other hand, this makes the CJPM unsuitable for attribution studies. Nevertheless, given the acceptable increase in computational cost of the CJPM over the SSJPM, the practical downsides of choosing the CJPM over the SSJPM appear limited when the objective is the estimation of return levels and return periods.

https://doi.org/10.5194/egusphere-2026-2118
Preprint. Discussion started: 22 April 2026

5 Conclusions

Correlation between peak tide and skew surge can measurably affect the estimation of extreme sea levels using joint probability methods, but quantifying the correlation structure and attributing correlation to its underlying physical causes is not trivial. We propose the Copula Joint Probability Method (CJPM), which builds on the Skew Surge Joint Probability Method (SSJPM) by introducing a copula to quantify the correlation without requiring prior understanding of its physical causes. Differing results between the SSJPM and CJPM may indicate the presence of non-negligible correlation and can motivate further investigation into underlying causes or drivers of the correlation, although the CJPM does not directly identify these drivers. In addition, the CJPM allows the contribution of correlation-associated uncertainty to estimated extreme sea levels to be quantified, providing additional insight beyond that available from the SSJPM. By not assuming independence between peak tide and skew surge, the CJPM is more general and may be applicable across a wider range of locations than the SSJPM. Further studies could evaluate the performance of the CJPM when applied to shorter tide gauge records (e.g., < 30 years) to assess its suitability at locations with limited data availability, such as Southeast Asia. Incorporating newer and improved methods for threshold selection of extreme skew surges could also be explored and may further improve the robustness of the CJPM.

Appendix A: Transformation local likelihood (TLL) copula

The TLL copula is obtained by first transforming joint observations (U_t, V_t) from the *copula space* to a *Gaussian layer*. Then, a local likelihood density estimator is used to estimate the PDF in the Gaussian layer before transforming the PDF back to the copula space to obtain the copula density. The TLL copula is defined as follows. We obtain joint observations in the Gaussian layer via $(A_t, B_t) = \left(\Phi^{-1}(U_t), \Phi^{-1}(V_t)\right)$, where Φ is the CDF of the standard normal distribution. The marginals of the joint observations (A_t, B_t) are standard normally distributed, i.e. $A \sim N(0,1)$ and $B \sim N(0,1)$, and we denote their joint PDF as f_{AB}. According to Eq. (10), the copula density can be written as

$$c(u,v) = \frac{f_{AB}\left(\Phi^{-1}(u), \Phi^{-1}(v)\right)}{\phi\left(\Phi^{-1}(u)\right)\phi\left(\Phi^{-1}(v)\right)}, \quad \text{(A1)}$$

where ϕ is the PDF of the standard normal distribution. Hence, c can be determined from f_{AB}. $\log f_{AB}$ is approximated using a second-order polynomial

$$\begin{aligned}\log f_{AB}(a', b') &\approx P\left(a' - a, b' - b \mid \boldsymbol{\alpha}_{a,b}\right) \\ &= \alpha_{0;(a,b)} + \alpha_{1,0;(a,b)}(a' - a) + \alpha_{0,1;(a,b)}(b' - b) \\ &+ \alpha_{2,0;(a,b)}(a' - a)^2 + \alpha_{1,1;(a,b)}(a' - a)(b' - b) + \alpha_{0,2;(a,b)}(b' - b)^2, \quad \text{(A2)}\end{aligned}$$

where the polynomial coefficients

$$\boldsymbol{\alpha}_{a,b} = [\alpha_{0;(a,b)} \quad \alpha_{1,0;(a,b)} \quad \alpha_{0,1;(a,b)} \quad \alpha_{2,0;(a,b)} \quad \alpha_{1,1;(a,b)} \quad \alpha_{0,2;(a,b)}]^{\mathrm{T}} \quad \text{(A3)}$$

https://doi.org/10.5194/egusphere-2026-2118
Preprint. Discussion started: 22 April 2026

are functions of a and b and are fitted to (A_t, B_t). The best fitting polynomial coefficients are obtained by solving the weighted maximum likelihood problem

$$\widehat{\boldsymbol{\alpha}}_{a,b} = \arg\max_{\boldsymbol{\alpha}_{a,b}} \left\{ \sum_{t=1}^{N} K\left(\begin{bmatrix} a \\ b \end{bmatrix} \mid \begin{bmatrix} A_t \\ B_t \end{bmatrix}, \mathbf{B}\mathbf{B}^{\mathrm{T}} \right) P\left(A_t - a, B_t - b \mid \boldsymbol{\alpha}_{a,b}\right) \right.$$

$$\left. -N \iint_{R^2} K\left(\begin{bmatrix} a \\ b \end{bmatrix} \mid \begin{bmatrix} a' \\ b' \end{bmatrix}, \mathbf{B}\mathbf{B}^{\mathrm{T}} \right) \exp\left(P\left(a' - a, b' - b \mid \boldsymbol{\alpha}_{a,b}\right) \right) da'db' \right\}, \quad \text{(A4)}$$

where

$$K(\mathbf{x} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{2\pi\sqrt{\det(\boldsymbol{\Sigma})}} \exp\left(-\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu}) \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right) \quad \text{(A5)}$$

is the Gaussian kernel (i.e. the PDF of the multivariate normal distribution) with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. $\mathbf{B}$ is a bandwidth matrix where $\mathbf{B}\mathbf{B}^{\mathrm{T}} = \boldsymbol{\Sigma}$. $\mathbf{B} = 3N^{-1/10}\widehat{\boldsymbol{\Sigma}}_{AB}^{1/2}$, where $\widehat{\boldsymbol{\Sigma}}_{AB}$ is the empirical covariance matrix of (A_t, B_t), has been found to produce good results and is used in *pyvinecopulib*'s implementation (Nagler and Vatter, 2025; Nagler, 2018). Finally, the approximation $f_{AB}(a, b) \approx \exp\left(\alpha_{0;(a,b)}\right)$ is made, and the TLL copula density is given by

$$c(u, v) = \frac{\exp\left(\alpha_{0;\left(\Phi^{-1}(u), \Phi^{-1}(v)\right)}\right)}{\phi\left(\Phi^{-1}(u)\right)\phi\left(\Phi^{-1}(v)\right)}. \quad \text{(A6)}$$

Renormalisation is applied to Eq. (A6) to ensure that it is a valid probability density that integrates to 1 (Geenens, 2017; Nagler, 2018; Nagler et al., 2017).

Code and data availability

The analysis code used to produce the figures and tables can be downloaded from https://doi.org/10.5281/zenodo.19312731 (Koh, 2026). The GESLA 3 sea level data set can be downloaded from https://gesla787883612.wordpress.com (Caldwell et al., 2015; Haigh et al., 2023; Woodworth et al., 2017). The UTide MATLAB functions can be downloaded from https://www.mathworks.com/matlabcentral/fileexchange/46523 (Codiga, 2011). The pyvinecopulib Python library can be downloaded from https://doi.org/10.5281/zenodo.14841456 (Nagler and Vatter, 2025).

https://doi.org/10.5194/egusphere-2026-2118
Preprint. Discussion started: 22 April 2026

Author contributions

ZYK: conceptualisation, data curation, formal analysis, investigation, methodology, software, validation, visualisation, writing (original draft preparation), and writing (review and editing). BSG: conceptualisation, investigation, methodology, and writing (review and editing). JD: conceptualisation, funding acquisition, methodology, and writing (review and editing). BPH: conceptualisation, funding acquisition, and writing (review and editing). LYC: conceptualisation, funding acquisition, methodology, project administration, resources, supervision, and writing (review and editing).

Competing interests

The contact author has declared that none of the authors has any competing interests.

Disclaimer

Any opinions, findings, conclusions, or recommendations expressed in this material are those of the author(s) and do not reflect the views of the National Research Foundation, Singapore, and the National Environment Agency, Singapore. Copernicus Publications remains neutral with regard to jurisdictional claims made in the text, published maps, institutional affiliations, or any other geographical representation in this paper. While Copernicus Publications makes every effort to include appropriate place names, the final responsibility lies with the authors. Views expressed in the text are those of the authors and do not necessarily reflect the views of the publisher.

Acknowledgements

We acknowledge the use of data from the Global Extreme Sea Level Analysis (GESLA) dataset, available at https://gesla787883612.wordpress.com/downloads. We thank the data providers and institutions that compiled and maintain the GESLA dataset.

Financial support

This research is supported by the National Research Foundation, Singapore, and the National Environment Agency, Singapore, under the National Sea-Level Programme Funding Initiative (Award no. USS-IF-2020-3).

https://doi.org/10.5194/egusphere-2026-2118
Preprint. Discussion started: 22 April 2026

References

Arns, A., Wahl, T., Wolff, C., Vafeidis, A. T., Haigh, I. D., Woodworth, P., Niehüser, S., and Jensen, J.: Non-linear interaction modulates global extreme sea levels, coastal flood exposure, and impacts, Nat. Commun., 11, 1918, https://doi.org/10.1038/s41467-020-15752-5, 2020.

Baranes, H. E., Woodruff, J. D., Talke, S. A., Kopp, R. E., Ray, R. D., and DeConto, R. M.: Tidally driven interannual variation in extreme sea level frequencies in the Gulf of Maine, J. Geophys. Res.-Oceans, 125, e2020JC016291, https://doi.org/10.1029/2020JC016291, 2020.

Batstone, C., Lawless, M., Tawn, J., Horsburgh, K., Blackman, D., McMillan, A., Worth, D., Laeger, S., and Hunt, T.: A UK best-practice approach for extreme sea-level analysis along complex topographic coastlines, Ocean Eng., 71, 28–39, https://doi.org/10.1016/j.oceaneng.2013.02.003, 2013.

Belzile, L., Dutang, C., Northrop, P., and Opitz, T.: A modeler's guide to extreme value software, Extremes, 26, 1–44, https://doi.org/10.1007/s10687-023-00475-9, 2023.

Buchanan, M. K., Kopp, R. E., Oppenheimer, M., and Tebaldi, C.: Allowances for evolving coastal flood risk under uncertain local sea-level rise, Clim. Change, 137, 347–362, https://doi.org/10.1007/s10584-016-1664-7, 2016.

Calafat, F. M., Wahl, T., Tadesse, M. G., and Sparrow, S. N.: Trends in Europe storm surge extremes match the rate of sea-level rise, Nature, 603, 841–845, https://doi.org/10.1038/s41586-022-04426-5, 2022.

Caldwell, P. C., Merrifield, M. A., and Thompson, P. R.: Sea level measured by tide gauges from global oceans as part of the Joint Archive for Sea Level (JASL) since 1846, NOAA National Centers for Environmental Information, [data set], https://doi.org/10.7289/V5V40S7W, 2015.

Caruso, M. F. and Marani, M.: Extreme-coastal-water-level estimation and projection: a comparison of statistical methods, Nat. Hazards Earth Syst. Sci., 22, 1109–1128, https://doi.org/10.5194/nhess-22-1109-2022, 2022.

Codiga, D. L.: UTide Unified Tidal Analysis and Prediction Functions (Version 1.0.0.0), MATLAB Central File Exchange [code], https://www.mathworks.com/matlabcentral/fileexchange/46523, 2011.

Coles, S.: An introduction to statistical modeling of extreme values, in: Springer Series in Statistics, Springer, London, UK, ISBN 978-1-4471-3675-0, https://doi.org/10.1007/978-1-4471-3675-0, 2001.

Collings, T. P., Murphy-Barltrop, C. J. R., Murphy, C., Haigh, I. D., Bates, P. D., and Quinn, N. D.: Automated tail-informed threshold selection for extreme coastal sea levels, Nat. Hazards Earth Syst. Sci., 25, 4545–4562, https://doi.org/10.5194/nhess-25-4545-2025, 2025.

Cosby, A. G., Lebakula, V., Smith, C. N., Wanik, D. W., Bergene, K., Rose, A. N., Swanson, D., and Bloom, D. E.: Accelerating growth of human coastal populations at the global and continent levels: 2000–2018, Sci. Rep., 14, 22489, https://doi.org/10.1038/s41598-024-73287-x, 2024.

https://doi.org/10.5194/egusphere-2026-2118
Preprint. Discussion started: 22 April 2026

584 Costa, W., Bryan, K. R., Stephens, S. A., and Coco, G.: A regional analysis of tide-surge interactions during extreme water
585 levels in complex coastal systems of Aotearoa New Zealand, Front. Mar. Sci., 10, 1170756,
586 https://doi.org/10.3389/fmars.2023.1170756, 2023.
587 D'Arcy, E., Tawn, J. A., Joly, A., and Sifnioti, D. E.: Accounting for seasonality in extreme sea-level estimation, Ann. Appl.
588 Stat., 17, 3500–3525, https://doi.org/10.1214/23-AOAS1773, 2023.
589 Diaz, D. B.: Estimating global damages from sea level rise with the Coastal Impact and Adaptation Model (CIAM), Clim.
590 Change, 137, 143–156, https://doi.org/10.1007/s10584-016-1675-4, 2016.
591 Dixon, M. J. and Tawn, J. A.: Extreme sea-levels at the UK A-class sites: site-by-site analyses, Proudman Oceanographic
592 Laboratory (Internal Document No. 65), https://ntslf.org/sites/ntslf/files/pdf/other_reports/id65.pdf (last access: 31 March
593 2026), 1994.
594 Dixon, M. J. and Tawn, J. A.: The effect of non-stationarity on extreme sea-level estimation, J. R. Stat. Soc. C-Appl. Stat., 48,
595 135–151, https://doi.org/10.1111/1467-9876.00145, 1999.
596 Ferro, C. A. T. and Segers, J.: Inference for clusters of extreme values, J. R. Stat. Soc. B-Stat. Methodol., 65, 545–
597 556, https://doi.org/10.1111/1467-9868.00401, 2003.
598 Galiatsatou, P., Makris, C., Prinos, P., and Kokkinos, D.: Non-stationary joint probability analysis of extreme marine variables
599 to assess design water levels at the shoreline in a changing climate, Nat. Hazards, 98, 1051–
600 1089, https://doi.org/10.1007/s11069-019-03645-w, 2019.
601 GEBCO Compilation Group: The GEBCO_2024 Grid – a continuous terrain model of the global oceans and land, NERC EDS
602 British Oceanographic Data Centre NOC [data set], https://doi.org/10.5285/1c44ce99-0a0d-5f4f-e063-7086abc0ea0f, 2024.
603 Geenens, G.: Probit transformation for kernel density estimation on the unit interval, J. Am. Stat. Assoc., 109, 346–358,
604 https://doi.org/10.1080/01621459.2013.842173, 2014.
605 Geenens, G., Charpentier, A., and Paindaveine, D.: Probit transformation for nonparametric kernel estimation of the copula
606 density, Bernoulli, 23, 1848–1873. https://doi.org/10.3150/15-BEJ798, 2017.
607 Geenens, G. and Wang, C.: Local-likelihood transformation kernel density estimation for positive random variables. J.
608 Comput. Graph. Stat., 27, 822–835. https://doi.org/10.1080/10618600.2018.1424636, 2018.
609 Gregory, J. M., Griffies, S. M., Hughes, C. W., Lowe, J. A., Church, J. A., Fukimori, I., Gomez, N., Kopp, R. E., Landerer, F.,
610 Cozannet, G. L., Ponte, R. M., Stammer, D., Tamisiea, M. E., and van de Wal, R. S. W.: Concepts and Terminology for Sea
611 Level: Mean, Variability and Change, Both Local and Global, Surv. Geophys., 40, 1251–1289, https://doi.org/10.1007/s10712-
612 019-09525-z, 2019.
613 Haasnoot, M., Winter, G., Brown, S., Dawson, R. J., Ward, P. J., and Eilander, D.: Long-term sea-level rise necessitates a
614 commitment to adaptation: A first order assessment, Clim. Risk Manag., 34, 100355,
615 https://doi.org/10.1016/j.crm.2021.100355, 2021.

https://doi.org/10.5194/egusphere-2026-2118
Preprint. Discussion started: 22 April 2026

Haigh, I. D., Marcos, M., Talke, S. A., Woodworth, P. L., Hunter, J. R., Hague, B. S., Arns, A., Bradshaw, E., and Thompson, P.: GESLA Version 3: A major update to the global higher-frequency sea-level dataset, Geosci. Data J., 10, 293–314, https://doi.org/10.1002/gdj3.174, 2022.

Hinkel, J., Lincke, D., Vafeidis, A. T., Perrette, M., Nicholls, R. J., Tol, R. S. J., Marzeion, B., Fettweis, X., Ionescu, C., and Levermann, A.: Coastal flood damage and adaptation costs under 21st century sea-level rise, P. Natl. Acad. Sci. USA, 111, 3292–3297, https://doi.org/10.1073/PNAS.1222469111, 2014.

Horsburgh, K. J. and Wilson, C.: Tide-surge interaction and its role in the distribution of surge residuals in the North Sea, J. Geophys. Res.-Oceans, 112, C08003, https://doi.org/10.1029/2006JC004033, 2007.

Idier, D., Bertin, X., Thompson, P., and Pickering, M.: Interactions Between Mean Sea Level, Tide, Surge, Waves and Flooding: Mechanisms and Contributions to Sea Level Variations at the Coast, Surv. Geophys., 40, 1603–1630, https://doi.org/10.1007/s10712-019-09549-5, 2019.

Kastrisios, C., Dyer, N., Nada, T., Contarinis, S., and Cordero, J.: Increasing efficiency of nautical chart production and accessibility to marine environment data through an open-science compilation workflow. ISPRS Int. J. Geo-Inf., 12, 116. https://doi.org/10.3390/ijgi12030116, 2023.

Kirezci, E., Young, I. R., Ranasinghe, R., Muis, S., Nicholls, R. J., Lincke, D., and Hinkel, J.: Projections of global-scale extreme sea levels and resulting episodic coastal flooding over the 21st Century, Sci. Rep., 10, 11629, https://doi.org/10.1038/s41598-020-67736-6, 2020.

Koh, Z. Y.: Analysis code for "Estimating Extreme Sea Levels Using a Copula Joint Probability Method", Zenodo [code], https://doi.org/10.5281/zenodo.19312731, 2026

Koh, Z. Y., Grandey, B. S., Samanta, D., Switzer, A. D., Horton, B. P., Dauwels, J., and Chew, L. Y.: Tide–surge interaction observed at Singapore and the east coast of Peninsular Malaysia using a semi-empirical model, Ocean Sci., 20, 1495–1511, https://doi.org/10.5194/os-20-1495-2024, 2024.

Kopp, R. E., Garner, G. G., Hermans, T. H. J., Jha, S., Kumar, P., Reedy, A., Slangen, A. B. A., Turilli, M., Edwards, T. L., Gregory, J. M., Koubbe, G., Levermann, A., Merzky, A., Nowicki, S., Palmer, M. D., and Smith, C.: The Framework for Assessing Changes To Sea-level (FACTS) v1.0: a platform for characterizing parametric and structural uncertainty in future global, relative, and extreme sea-level change, Geosci. Model Dev., 16, 7461–7489, https://doi.org/10.5194/gmd-16-7461-2023, 2023.

Kulp, S. A. and Strauss, B. H.: New elevation data triple estimates of global vulnerability to sea-level rise and coastal flooding, Nat. Commun., 10, 1–12, https://doi.org/10.1038/s41467-019-12808-z, 2019.

Leadbetter, M. R.: Extremes and local dependence in stationary sequences, Z. Wahrscheinlichkeitstheorie verw. Gebiete, 65, 291–306, https://doi.org/10.1007/BF00532484, 1983.

Lee, U.-J., Cho, H.-Y., Lee, B. W., and Ko, D.-H.: Joint probability distribution of significant wave height and peak wave period using Gaussian copula method, J. Coast. Res., 116(spl), 96–100, https://doi.org/10.2112/JCR-SI116-020.1, 2024.

https://doi.org/10.5194/egusphere-2026-2118
Preprint. Discussion started: 22 April 2026

649 Li, J., Pan, S., Chen, Y., Gan, M.: The performance of the copulas in estimating the joint probability of extreme waves and
650 surges along east coasts of the mainland China, Ocean Eng. 237, 109581, https://doi.org/10.1016/j.oceaneng.2021.109581,
651 2021.
652 Masina, M., Lamberti, A., and Archetti, R.: Coastal flooding: A copula based approach for estimating the joint probability of
653 water levels and waves, Coast. Eng., 97, 37–52, https://doi.org/10.1016/j.coastaleng.2014.12.010, 2015.
654 Mawdsley, R. J. and Haigh, I. D.: Spatial and temporal variability and long-term trends in skew surges globally, Front. Mar.
655 Sci., 3, 29, https://doi.org/10.3389/fmars.2016.00029, 2016.
656 McElreath, R.: The Haunted DAG & The Causal Terror, in: Statistical Rethinking: A Bayesian Course with Examples in R
657 and STAN (2nd Edition), Chapman and Hall/CRC, New York, NY, USA, 165–194, ISBN 9780429029608,
658 https://doi.org/10.1201/9780429029608, 2020.
659 Mesbahzadeh, T., Miglietta, M. M., Mirakbari, M., Soleimani Sardoo, F., and Abdolhoseini, M.: Joint modeling of
660 precipitation and temperature using copula theory for current and future prediction under climate change scenarios in arid
661 lands (Case study, Kerman Province, Iran), Adv. Meteorol., https://doi.org/10.1155/2019/6848049, 2019.
662 Murphy, C., Tawn, J. A., and Varty, Z.: Automated threshold selection and associated inference uncertainty for univariate
663 extremes, Technometrics, 67, 215–224, https://doi.org/10.1080/00401706.2024.2421744, 2025.
664 Nagler, T., Schellhase, C., and Czado, C.: Nonparametric estimation of simplified vine copula models: comparison of methods,
665 Depend. Model., 5, 99–120, https://doi.org/10.1515/demo-2017-0007, 2017.
666 Nagler, T.: kdecopula: An R package for the kernel estimation of bivariate copula densities, J. Stat. Softw., 84, 1–22.
667 https://doi.org/10.18637/jss.v084.i07, 2018.
668 Nagler, T. and Vatter, T.: pyvinecopulib (Version v0.7.1), Zenodo [code], https://doi.org/10.5281/zenodo.14841456, 2025.
669 Olbert, A. I., Nash, S., Cunnane, C., and Hartnett, M.: Tide–surge interactions and their effects on total sea levels in Irish
670 coastal waters, Ocean Dyn., 63, 599–614, https://doi.org/10.1007/s10236-013-0618-0, 2013.
671 Oppenheimer, M., Glavovic, B. C., Hinkel, J., van de Wal, R., Magnan, A. K., Abd-Elgawad, A., Cai, R., Cifuentes-Jara, M.,
672 Deconto, R. M., Ghosh, T., Hay, J., Isla, F., Marzeion, B., Meyssignac, B., and Sebesvari, Z.: Sea level rise and implications
673 for low lying islands, coasts and communities, in: IPCC Special Report on the Ocean and Cryosphere in a Changing Climate,
674 edited by: Pörtner, H.-O., Roberts, D. C., Masson-Delmotte, V., Zhai, P., Tignor, M., Poloczanska, E., Mintenbeck, K.,
675 Alegréia, A., Nicolai, M., Okem, A., Petzold, J., Rama, B., and Weyer, N. M., Cambridge University Press, Cambridge, UK
676 and New York, NY, USA, 321–445, https://doi.org/10.1017/9781009157964.006, 2019.
677 Palmer, K., Watson, C. S., Power, H. E., and Hunter, J. R.: Quantifying the mean sea level, tide, and surge contributions to
678 changing coastal high water levels, J. Geophys. Res.-Oceans, 129, e2023JC020737. https://doi.org/10.1029/2023JC020737,
679 2024.
680 Pirazzoli, P. A. and Tomasin, A.: Estimation of return periods for extreme sea levels: a simplified empirical correction of the
681 joint probabilities method with examples from the French Atlantic coast and three ports in the southwest of the UK, Ocean
682 Dyn., 57, 91–107, https://doi.org/10.1007/s10236-006-0096-8, 2007.

https://doi.org/10.5194/egusphere-2026-2118
Preprint. Discussion started: 22 April 2026

Pugh, D. T. and Vassie, J. M.: Extreme sea levels from tide and surge probability, Coast. Eng. Proc., 1, 52, https://doi.org/10.9753/icce.v16.52, 1978.

Pugh, D. T. and Woodworth, P.: Introduction, in: Sea-Level Science: Understanding Tides, Surges, Tsunamis and Mean Sea-Level Changes, Cambridge University Press, Cambridge, UK, 1–16, https://doi.org/10.1017/CBO9781139235778.004, 2014a.

Pugh, D. T. and Woodworth, P.: Tidal analysis and prediction, in: Sea-Level Science: Understanding Tides, Surges, Tsunamis and Mean Sea-Level Changes, Cambridge University Press, Cambridge, UK, 60–96, https://doi.org/10.1017/CBO9781139235778.007, 2014b.

Ragno, E., Antonini, A., and Pasquali, D.: Investigating extreme sea level components and their interactions in the Adriatic and Tyrrhenian Seas. Weather Clim. Extremes, 41, 100590. https://doi.org/10.1016/j.wace.2023.100590, 2023.

Rasmussen, D. J., Bittermann, K., Buchanan, M. K., Kulp, S., Strauss, B. H., Kopp, R. E., and Oppenheimer, M.: Extreme sea level implications of 1.5 °C, 2.0 °C, and 2.5 °C temperature stabilization targets in the 21st and 22nd centuries, Environ. Res. Lett., 13, 034040, https://doi.org/10.1088/1748-9326/aaac87, 2018.

Santamaria-Aguilar, S. and Vafeidis, A. T.: Are Extreme Skew Surges Independent of High Water Levels in a Mixed Semidiurnal Tidal Regime?, J. Geophys. Res.-Oceans, 123, 8877–8886, https://doi.org/10.1029/2018JC014282, 2018.

Sklar, M.: Fonctions de répartition à n dimensions et leurs marges. Publications de l'Institut de statistique de l'Université de Paris, 8, 229–231, https://hal.science/hal-04094463v1 (last access: 6 April 2026), 1959.

Tavakol, A., Rahmani, V., and Harrington Jr., J.: Probability of compound climate extremes in a changing climate: A copula-based study of hot, dry, and windy events in the central United States, Environ. Res. Lett., 15, 104058, https://doi.org/10.1088/1748-9326/abb1ef, 2020.

Tawn, J. A.: Estimating probabilities of extreme sea-levels, J. R. Stat. Soc. C-Appl. Stat., 41, 77–93, https://doi.org/10.2307/2347619, 1992.

Tebaldi, C., Ranasinghe, R., Vousdoukas, M., Rasmussen, D. J., Vega-Westhoff, B., Kirezci, E., Kopp, R. E., Sriver, R., and Mentaschi, L.: Extreme sea levels at different global warming levels, Nat. Clim. Change, 11, 746–751, https://doi.org/10.1038/s41558-021-01127-1, 2021.

Van Vliet, B.: Abe Sklar's "FONCTIONS DE REPARTITION A N DIMENSIONS ET LEURS MARGES": The Original Document and an English Translation, SSRN, https://doi.org/10.2139/ssrn.4198458 (last access: 6 April 2026), 2023.

Williams, J., Horsburgh, K. J., Williams, J. A., and Proctor, R. N. F.: Tide and skew surge independence: New insights for flood risk, Geophys. Res. Lett., 43, 6410–6417, https://doi.org/10.1002/2016GL069522, 2016.

Williams, J., Matthews, A., and Jevrejeva, S.: Development of an automatic tide gauge processing system, National Oceanography Centre (Research & Consultancy Report No. 64), https://psmsl.org/cme/NOC_RandC_64_Final.pdf (last access: 31 March 2026), 2019.

Woodworth, P. L., Hunter, J. R., Marcos, M., Caldwell, P., Menéndez, M., and Haigh, I.: Towards a global higher-frequency sea level dataset, Geosci. Data J., 3, 50–59, https://doi.org/10.1002/gdj3.42, 2016.

https://doi.org/10.5194/egusphere-2026-2118
Preprint. Discussion started: 22 April 2026

716 Yang, S., Sheng, J., Ohashi, K., Yang, B., Chen, S., Xing, J.: Non-linear interactions between tides and storm surges during
717 extreme weather events over the eastern Canadian shelf, Ocean Dyn., 73, 279–301, https://doi.org/10.1007/s10236-023-01556-
718 w, 2023.
719 Yavuzdoğan, A. and Tanır Kayıkçı, E.: A copula approach for sea level anomaly prediction: a case study for the Black Sea.
720 Surv. Rev., 53, 436–446. https://doi.org/10.1080/00396265.2020.1816314, 2021.
721
722
723