A spatially-dependent synthetic global dataset of extreme sea level events

Huazhi Li [a,*], Toon Haer [a], Anaïs Couasnon [a,c], Alejandra R. Enríquez [a,b], Sanne Muis [a,c], Philip J. Ward [a]

[a] *Institute for Environmental Studies (IVM), Vrije Universiteit Amsterdam, Amsterdam, the Netherlands*
[b] *Department of Civil, Environmental and Construction Engineering, National Center for Integrated Coastal Research, University of Central Florida, Orlando, FL, USA*
[c] *Deltares, Delft, the Netherlands*

A B S T R A C T

Current coastal flood risk assessments fail to capture flood spatial dependence at large scales. In this paper, we develop the first global synthetic dataset of spatially-dependent extreme sea level events, by applying an existing conditional multivariate statistical model to 40-year global reanalysis sea levels. The resulting dataset contains 10,000 years of extreme events with realistic spatial dependence under current climate conditions. The benchmarking against reanalysis data demonstrates a high agreement, with a coefficient of determination (R^2) of 0.96 for the mean event footprint sizes and a mean bias of -0.04 m for 1 in 50-year water levels. By comparing well-known historic events, we show that our approach can produce events with similar spatial characteristics. Our dataset enables the future development of spatially-dependent flood hazard maps for deriving accurate large-scale risk profiles, which can help yield new insights into the spatial patterns of coastal flooding and support coastal communities in devising effective management plans.

1. Introduction

Coastal flooding is among the world's most devastating natural hazards (CRED, 2018). In the last two decades (2002–2022), reported global economic losses due to coastal floods exceeded US$ 1,900 billion, and over 200,000 people lost their lives (CRED, 2023). These impacts can be particularly high when a flood event affects a large spatial area, that is, a widespread event. For example, the most damaging weather and climate disaster between 1980 and 2021 in the U.S. is Hurricane Katrina. In late August 2005, it led to widespread flooding in nine states and caused approximately US$ 180 billion losses and more than 1,800 deaths (NOAA, 2022). In a warming future, with accelerating sea level rise (Pörtner et al., 2022), and possible changes in storm surges (Little et al., 2015), wind waves (Perez et al., 2015) and tides (Mawdsley et al., 2015), high sea levels will occur more frequently which is expected to intensify coastal flood hazard and risk (Tiggeloven et al., 2020; Vitousek et al., 2017; Vousdoukas et al., 2018). Without adaptation, 123 cm of global mean sea level rise is projected to escalate the annual economic damage of coastal flooding to 9.3% of global gross domestic product in 2100 (Hinkel et al., 2014). Therefore, to better protect coastal communities from flood impacts, it is necessary to reliably assess extreme sea levels (i.e. probability of occurrence), their associated flood hazards (i.e.

inundation extent and depth), and the consequences of these floods (e.g. risk).

Despite the high impact potential of widespread extreme sea levels and associated coastal flooding, current risk assessments are typically based on an at-site univariate flood frequency analysis (Vorogushyn et al., 2018). This approach assumes "complete dependence" scenarios, in which return periods of water levels are spatially constant across all locations (e.g. Hallegatte et al., 2013; Hinkel et al., 2014; Kirezci et al., 2020; Tiggeloven et al., 2020). Water levels and their associated return periods during actual events, however, vary strongly across space. Analyses of historic river floods demonstrate that their spatial patterns show asymptotic dependence between sites (Keef et al., 2009, 2013b): during an extreme event, proximal locations may experience similar extremes and the extremity is likely to decrease as the distance between sites increases. As with river floods, extreme sea level events also show spatial dependence between sites (Lamb et al., 2010; Paprotny et al., 2016) and the dependence tends to increase with the increasing event extremity (Wyncoll et al., 2016). Therefore, the at-site univariate assumption is an unrealistic representation of spatial patterns of extreme sea level events. Although at the local scale this approach may be acceptable (as water levels may be strongly correlated at different locations along a short coastline), this assumption becomes problematic as

---

the spatial scale increases. Besides, coastal flood events (driven by high sea levels) of the same severity are found to occur at different frequencies at different locations, as shown for the U.S. (Sweet and Park, 2014) and Australia (Hague et al., 2022). In this work, we aim to address the spatially-constant assumption of the past studies by looking into the spatial heterogeneity of return periods of sea level events.

Capturing spatial dependence of extreme sea level events is paramount for risk management and decision-making (Jongman et al., 2014; Vorogushyn et al., 2018). For example, international communities and governments (e.g. the United Nations Office for Disaster Risk Reduction and the European Union) need to know the spatially-dependent probability distribution of annual losses to develop (trans)national financing plans robust to present and future risks, such as the EU solidarity fund (Ward et al., 2015). This information is also required in the (re)insurance industry to quantify for pricing and reassure market regulators that they can maintain solvency in high loss years. For emergency services, accurate estimates of the occurrence probability and potentially affected area of a widespread sea level event can be used to arrange temporary measures and mobilize sufficient resources in good time and at the right locations (Quinn et al., 2019). Lastly, comprehensive knowledge of spatial dependence of extreme sea levels can help coastal planners to design more efficient and science-based disaster preparedness plans. This is especially the case for the spatial planning of critical infrastructures. For example, extreme sea level events can cause large-scale failures of power systems, as witnessed during Hurricanes Maria, Harvey, and Florence (Feng et al., 2022).

These concerns have driven research to improved flood assessments by explicitly accounting for spatial dependence, in particular for river floods (e.g. Jongman et al., 2014; Keef et al., 2013b, 2009; Metin et al., 2020; Nguyen et al., 2020). However, only a limited number of studies have addressed the spatial dependence of extreme sea levels. By analyzing historic events using tide gauge records, two studies looked into the spatial patterns of extreme sea level and skew surge events on the coasts of the UK (Haigh et al., 2016) and New Zealand (Stephens et al., 2020). Recently, Enríquez et al. (2020) used both reanalysis and observed data to assess the spatial footprints of storm surges at the global scale. These studies have improved our understanding of spatial dependence of extreme sea levels; however, they were limited by the use of short-term forcing data, resulting in a relatively small sample of extreme events. Although using such an event set can capture the historic spatial dependence, it is not sufficient to represent all possible sea level patterns that may occur (Serinaldi and Kilsby, 2017). One notable exception is the work of Rueda et al. (2016) who generated synthetic events through a weather-type and statistical downscaling emulator to characterize the joint probability of extreme waves and storm surges on the northern coast of Spain. In addition, Vousdoukas et al. (2020) used a probabilistic approach to produce realizations of spatially-dependent extreme sea levels for assessing the effectiveness of raising dyke in Europe. To date, a global-scale reliable assessment of spatial dependence of extreme sea levels is still lacking. Here we address this gap by looking into spatial patterns of events driven by meteorological and oceanographic forcing such as storm surges from tropical and extra tropical cyclones.

The approaches used for capturing spatial dependence in river flooding could be beneficial and transferrable for simulating the spatial patterns of sea level events. In general, two methods exist to model the spatial dependence in river flood peaks, namely: (1) continuous hydrological simulation (e.g. Falter et al., 2016, 2015; Metin et al., 2020; Winter et al., 2019); and (2) event-based statistical approach (e.g. Jongman et al., 2014; Nguyen et al., 2020; Quinn et al., 2019; Wing et al., 2020). Compared to the continuous hydrological approach, the event-based method is much less computationally expensive and allows the stochastic generation of a large number of extreme events from the statistical model, including those not seen in the records. This approach therefore enables the future development of flood hazard maps that explicitly account for the spatial dependence (especially of the upper tail

of the joint distribution), which can then form the basis of reliable estimates of the risk distribution across all probabilities. To date, two main classes of statistical techniques have been used to model the spatial dependence in the event-based approach: copulas (e.g. Brunner et al., 2019; Couasnon et al., 2018; Jongman et al., 2014; Nguyen et al., 2020) and the conditional multivariate exceedance model (e.g. Keef et al., 2013b, 2009; Quinn et al., 2019; Wing et al., 2020). Both approaches are capable of calculating the spatial dependence. However, standard copula models are typically constrained to only one type of extremal dependence (Heffernan, 2001), while the conditional exceedance model offers more flexibility to handle a range of extremal dependence classes but at the cost of higher sensitivity (Towe et al., 2019). As a result, event-based multivariate extreme models may be a better choice for understanding spatial dependence of extreme sea level events at large spatial scales.

In this paper, we leverage these insights from hydrological flood modelling to develop the first global synthetic dataset of spatially-dependent extreme sea level events using an existing event-based conditional multivariate exceedance approach. We focus on sea level events driven by meteorological and oceanographic forcing, and therefore our method is not appropriate for examining minor or 'nuisance' events caused by high tides alone. We use 40-year global extreme sea level (mean sea levels, storm surges, and tides) time series (Muis et al., 2022) simulated by Global Tide and Surge Model (GTSM) version 3.0 forced with ERA5 climate reanalysis (Hersbach et al., 2020), hereafter GTSM-ERA5 dataset. The dynamic contribution of waves is not considered here as the breaking of waves in the nearshore is a local process that requires high-resolution data and models to simulate. For the sake of simplicity and because we do not foresee extreme sea level events impacting areas at quasi-continental scale or larger, we first divide the globe into 10 large conterminous regions (see Fig. S1 and Table S1). For each region, we perform a clustering of the GTSM-ERA5 stations based on Pearson correlations between sea levels. Next, we use the conditional multivariate exceedance model (Heffernan and Tawn, 2004) to calculate the pairwise dependence within each cluster. The resulting dependence patterns are then used to simulate 10,000 years of synthetic extreme sea level events. Lastly, we benchmark the synthetic dataset with the input GTSM-ERA5 reanalysis dataset by comparing event spatial footprint sizes and return levels. We further evaluate this dataset by showing its capability of reproducing several well-known historic events. We focus on comparisons of very rare events here because such events are relatively difficult to capture (normally only a few resembling events can be generated) compared to more frequent but less extreme events (many similar events can be generated) using our approach. Besides, including such low-probability but potentially high impact events with accurate spatial dependence is useful for broad-scale impact and adaptation analyses especially in highly developed and well-protected regions.

2. Materials and methods

Fig. 1 summarizes the statistical framework for developing the global synthetic dataset of spatially-dependent extreme sea levels. The framework consists of four main steps:

1. Preprocess the time series of global GTSM-ERA5 total water levels;
2. Calculate the pairwise spatial dependence between GTSM-ERA5 stations where spatiotemporal limits are first set to ensure independent and identically distributed events;
3. Generate the spatially-dependent synthetic event set by simulating 10,000 years of synthetic events of spatially-dependent water levels;
4. Benchmark the event set based on event footprints and return levels and evaluate the event set by examining its capability of reproducing several well-known historic events.
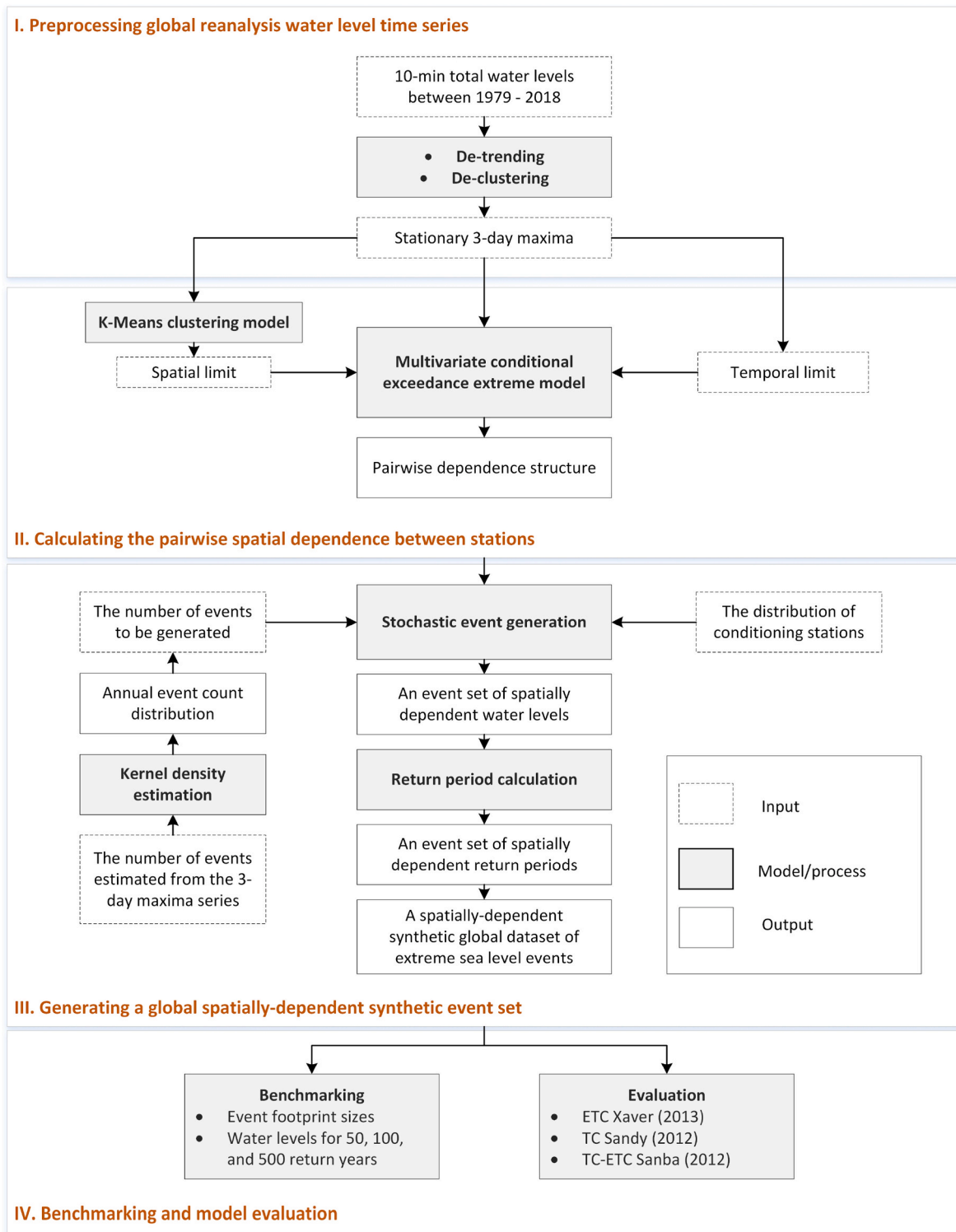
I. Preprocessing global reanalysis water level time series

10-min total water levels between 1979 - 2018

- **De-trending**
- **De-clustering**

Stationary 3-day maxima

K-Means clustering model

Spatial limit

Multivariate conditional exceedance extreme model

Temporal limit

Pairwise dependence structure

II. Calculating the pairwise spatial dependence between stations

The number of events to be generated

Stochastic event generation

The distribution of conditioning stations

Annual event count distribution

An event set of spatially dependent water levels

Kernel density estimation

Return period calculation

The number of events estimated from the 3-day maxima series

An event set of spatially dependent return periods

A spatially-dependent synthetic global dataset of extreme sea level events

Input

Model/process

Output

III. Generating a global spatially-dependent synthetic event set

Benchmarking
- Event footprint sizes
- Water levels for 50, 100, and 500 return years

Evaluation
- ETC Xaver (2013)
- TC Sandy (2012)
- TC-ETC Sanba (2012)

IV. Benchmarking and model evaluation

Fig. 1. Flowchart of four main steps in the methodological framework.

2.1. Preprocessing global reanalysis water level time series

We use 40-year global extreme sea level time series (Muis et al., 2022) simulated by the third generation of the Global Tide and Surge Model (GTSM) forced with 10-m wind speed and atmospheric pressure from the ERA5 climate reanalysis (Hersbach et al., 2020). The simulation of GTSM-ERA5 reanalysis water levels also considers a spatially-varying sea level rise field derived from the Coupled Model Intercomparison Project - Phase 5 (CMIP5) experiments (Taylor et al., 2012). These water levels are 10-min total water level time series consisting of mean sea levels, tides, and storm surges for the historic period 1979–2018. The time series are stored at 43,119 output points including coastal output locations, tide gauge stations, and deep ocean points. The spatial resolution of these output stations increases globally from the deep ocean toward the coast, which is 5.0°, 2.5°, and 1.0° for stations with a distance of >500 km, 100–500 km, and <100 km from land,

respectively. For Europe, the corresponding resolution is 1.0°, 0.5°, and 0.25° due to the use of higher-resolution regional climate simulations. The GTSM-ERA5 dataset has been previously validated against observations, which has shown a good agreement between modeled and observed water levels (Dullaart et al., 2020; Muis et al., 2020).

In this study, we restrict our analysis to 17,394 coastal stations covering the global coastline, by removing ocean stations as well as stations in the Arctic and Antarctica. The water levels at selected stations are then linearly detrended by removing the annual mean sea levels to adapt to the present climate (Read and Vogel, 2015). Here water levels are defined relative to the 1979–2018 mean. We further convert these water levels into 3-day maxima to account for the surge traveling time (Wyncoll et al., 2016) and ensure independent and identically distributed events (Enríquez et al., 2020). These preprocessing steps prepare quasi-stationary time series for the spatial dependence calculation.

2.2. Calculating the pairwise spatial dependence between stations

2.2.1. Spatiotemporal limits for defining independent events

First, we specify spatiotemporal limits for defining independent and identically distributed events. As independence between extreme sea levels is well captured by 3-day maxima series (Vousdoukas et al., 2018), we do not set additional temporal limits to the 3-day water level maxima series. A spatial limit is used to account for the physical limits to the maximum extent of sea level events because extreme water levels at distant locations are less likely to be driven by the same weather event. In this paper, we focus primarily on large spatial footprints of extreme events and therefore do not perform our analysis in small areas (i.e. at the scale of bays, cities, and states). On the other hand, the spatial footprints of such events are less likely to extend beyond a quasi-continental scale. High tide events may have spatial footprints extending on a global scale as the spring and neap tides occur at (approximately) the same time everywhere around the world; however these events require different methods to characterize (e.g. Hague et al., 2022; Sweet and Park, 2014) and thus are beyond the scope of this study. As a trade-off between these considerations, we follow the approach of Enríquez et al. (2020) to divide the globe into 10 large conterminous coastal regions and conduct our analysis separately in each region. The 10 regions are (1) Northwest Pacific, (2) Southwest Pacific, (3) Northwest Atlantic, (4) Southwest Atlantic, (5) Northeast Atlantic, (6) Southeast Atlantic, (7) Indian Ocean, (8) North Asia, (9) South Asia, and (10) Oceania, see Fig. S1 and Table S1.

To further estimate the spatial limit in each region, extreme sea levels are first identified at each station by applying the peak-over-threshold method to the 3-day water level maxima using the 95th percentile, following the approach of Enríquez et al. (2020). Next, the K-Means algorithm (Hastie et al., 2009) is applied to the extreme sea level series to cluster stations in each region into subgroups, in which the Pearson correlation is used to measure the similarity between stations. To identify the number of clusters (NoC) that can best represent the spatial extent of extreme sea level events in each region, we run the K-Means model with a range of NoC values from 5 to 30. It can be expected that with higher NoC, the statistical correlation of clustered stations becomes stronger; however, setting the maximum NoC to 30 prevents the identification of small-scale spatial footprints. Note that we select 30 as the upper limit instead of 20 used by Enríquez et al. (2020) since this research uses a finer-resolution input dataset in which adopting higher NoC values allows for capturing more possible unobserved spatial patterns. The optimal NoC obtained for each region is the one resulting in the highest ratio of mean correlation and average standard deviation following the method of Enríquez et al. (2020). Table S1 summarizes the final NoC value and the number of stations in each region. Detailed clustering information including cluster size and station locations can be found in Figs. S2–11 in the Supplementary Materials. For simplicity, we allow the dependence to be only calculated from a given station to any other stations within each cluster, while stations outside this cluster are assumed to have no dependence on the given station.

2.2.2. Multivariate conditional exceedance extreme model

Within each cluster, we consider a set Δ of d stations (i.e. the number of stations identified by the K-Means clustering algorithm) and the water level component X_i (i.e. the 3-day maxima water level series) at station i. The joint distribution of water levels within a cluster is then defined as $X = \{X_i, i \in \Delta, \Delta = \{1, ..., d\}\}$, in which each X_i has a continuous distribution. This joint distribution can be separated into independent marginal distributions and joint distributions on a common marginal distribution (i.e. the dependence structure). To this end, the conditional multivariate exceedance model of Heffernan and Tawn (2004) consists of two main steps: (1) defining marginal distribution independently at each station, and (2) calculating the pairwise dependence between stations based on regression functions. Full model details can be found in Heffernan and Tawn (2004) and Keef et al. (2013b, 2009).

To convert X onto a marginal scale, the conditional exceedance model initially used the Gumbel marginal distribution. However, the Laplace distribution is adopted here because it provides better estimations of the linear relationship between extremes (Keef et al., 2013a). Following this, we denote the set of water level component $X_i, i \in \Delta, \Delta = \{1, ..., d\}$ where X_i is the 3-day water level maxima at station i and achieve the transformation by:

$$Y_i = \begin{cases} log\{2F_i(X_i)\}, X_i < F_i^{-1}(0.5) \\ -log\{2[1 - F_i(X_i)]\}, X_i \geq F_i^{-1}(0.5) \end{cases} \tag{1}$$

where F_i is the marginal distribution of X_i.

Y then has Laplace margins with the probability function:

$$Pr\,(Y_i < y) = \begin{cases} \dfrac{1}{2}e^y, y < 0 \\ 1 - \dfrac{1}{2}e^{-y}, y \geq 0 \end{cases} \tag{2}$$

In this way, both lower and upper tails of Y_i are exactly exponential. Therefore, for any $u > 0$ where u is a threshold, the distributions of $Y_i - u|Y_i > u$ and $-(Y_i + u)|Y_i < -u$ are both exponential with a mean of 1. These marginal distributions define the probability that a certain water level is exceeded at each station. These functions are semi-parametric as the water levels above the threshold (i.e. u) are fitted to a generalized Pareto distribution while an empirical distribution is used for data below the threshold. As suggested by prior studies (Keef et al., 2013b; Quinn et al., 2019; Wing et al., 2020), this step requires careful consideration in threshold selection given that the quantified threshold should be high enough to avoid misestimation of extreme dependence, but low enough to ensure sufficient samples with which to fit the generalized Pareto model. We therefore use the cross-validatory method (Northrop et al., 2017) to assist in the threshold diagnosis process. Instead of using a unified threshold, we quantify the best-fit threshold for each station and use this set of station-wise thresholds as we believe this can best represent the extreme distribution pattern for a particular cluster.

The second step is to calculate the pairwise dependence between stations. The dependence model follows the equation below:

$$Y_{-i}|Y_i = aY_i + Y_i^b Z_{-i} \quad \text{for} \quad Y_i > v \tag{3}$$

where Y_{-i} is a vector of marginal distributions excluding Y_i, v is the threshold above which the dependence is estimated, a is a vector of parameters describing the overall strength of the dependence ($-1 \leq a \leq 1$, with $-1 \leq a < 0$ and $0 < a \leq 1$ corresponding respectively to negative and positive dependence), b is a vector of parameters describing how the dependence changes ($b < 1$, with positive values meaning the variance increases as y increases), Z_{-i} is a vector of $d - 1$ residuals with non-zero means and independent of Y_i. The residuals do not follow any simple class of parametric distributions (Keef et al.,

2013a, 2013b). For this, a quantile-based inference approach is used to exploit properties of the model itself for estimating the residuals. Full details can be found in Keef et al. (2013b). The equation (3) is used for calculating the dependence between two stations and this calculation is therefore repeated for all possible pairs within a cluster of interest. That is, given a cluster of d stations, in total $d^2 - 1$ pairs of dependence structure are modeled, whereby each pair characterizes distinct parameters $a_{j|i}$, $b_{j|i}$ and residuals $Z_{j|i}$ for the j th station.

2.3. Generating a global spatially-dependent synthetic event set

2.3.1. Stochastic event generation

In this study, the conditional multivariate exceedance model estimates the conditional distribution of water levels within any cluster of a set of stations when the water level at a given station is extreme. This information can be used to develop an event set of a large number of spatially-dependent events, whereby for every individual event at least one station has an extreme water level. The event set is therefore expressed by $E = \{y \in \mathbb{R}^d : \exists i \in \Delta, \Delta = \{1, ..., d\}, y_i > u\}$, where u is a high threshold. We follow the approach of Keef et al. (2013b) to further divide the full set into subsets where events are simulated given each station is the conditioning station: $E_i = \{y \in \mathbb{R}^d : y_i > u \text{ and } y_i = \max(y)\}$ for $i \in \Delta$. The event set can be quantified by a multinomial distribution with n_s events (i.e. the total number of events to be generated) and a probability vector $Pr(Y \in E_i)/Pr(Y \in E)$ for $i \in \Delta$.

Following this, n_s is first estimated across any given cluster. We use the 99th quantile threshold to identify extreme events from the input 3-day maxima series to empirically calculate the number of events per year for any given cluster. Discrete kernel density estimation (Wansouwé et al., 2016) is then applied to generate an annual event count distribution. In this way, the annual event number is allowed to be extrapolated to higher values to a reasonable extent (i.e. there are likely to be more events expected to occur in the simulation than in the observations). For generating a 10,000-year event set, we estimate the total number of events by summing up 10,000 values of the number of events per year which are randomly sampled from the annual event count distribution.

The next step is to estimate the portion of events for which a given station is the conditioning station. We use the empirical data to calculate the probability that each station has the largest water level (on the Laplace margin) out of all stations during any given event in the cluster. To generate E_i, we repeat the following simulation steps until the desired number of events is obtained:

1. Sample Y_i from the fitted generalized Pareto distribution at station i, conditional on $Y_i > u$;
2. Independently select a joint residual Z_i;
3. Calculate Y_{-i} based on the dependence model with the estimated regression parameters a_i, b_i, using Eq. (3);
4. Reject the sample Y if Y_i is not maximum over all stations (on the marginal scale).

We use the 99th quantile as the prediction/event sampling threshold u. This leads to the development of a 10,000-year event set of spatially-dependent events, in which for every individual event at least one station has a water level higher than its 99th quantile. Compared to the 95th quantile used in the clustering analysis, we apply the 99th quantile throughout the event generation process. By using this higher threshold, we aim to reduce the excessive generation of events (i.e. avoid generating events with very low return periods) and further minimize the computational costs.

2.3.2. Return period calculation

From the 10,000-year of spatially-dependent extreme sea level events, we empirically estimate the return periods of water levels for each station. We do not calculate the return periods statistically (i.e. fitting extreme value distributions to water levels), as is commonly adopted in extreme sea level studies (e.g. Wahl et al., 2017). A limitation of the statistical approach is that the fitted return periods are strongly influenced by the selected distribution, especially for higher return periods (Esteves, 2013). By using an empirical approach, no shape parameter needs to be estimated which is a particularly sensitive parameter of great influence for the upper tail. Besides, these return periods can be calculated without extrapolation of the data. Here we use Weibull's plotting formula because this approach has been proven to give the best empirical estimation of return periods compared to other methods (Makkonen, 2006). Weibull's formula is given as:

$$P_{exc}(wl) = \frac{i}{n+1} \bullet \frac{n}{m} \tag{4.1}$$

$$T(wl) = 1 / P_{exc}(wl) \tag{4.2}$$

where $P_{exc}(wl)$ is the exceedance probability of a given water level, i is the rank of this water level in the calculated extreme distribution with 1 referring to the highest water level, n is the number of events in the set, m is the temporal length of the set in years (here, $m = 10,000$). The return period $T(wl)$ can then be calculated as the inverse of the exceedance probability $P_{exc}(wl)$. The generated event set is representative of 10,000-year extreme sea level events. Given that the return period calculation is dominated by the rank i, a more uncertain result can be expected as the return period becomes higher. Therefore, we only estimate the return period up to 1,000 years. Table S2 summarizes the range of lowest estimated water level return periods in each region.

2.4. Benchmarking and model evaluation

To benchmark the 10,000-year global synthetic dataset of extreme sea level events, we compare the simulated events with extreme events from the input GTSM-ERA5 dataset. The extreme events of the GTSM-ERA5 dataset are extracted using a 99th quantile threshold (the same as the prediction/conditioning threshold) under the same spatiotemporal limits. To evaluate the performance of the synthetic set in capturing the spatial dependence, we first compare the mean event footprint sizes of the synthetic set against those of the GTSM-ERA5 dataset. As both datasets have the same output stations, the event footprint is approximated by the number of stations impacted by a certain event (i.e. the number of stations experiencing water level greater than the 99th quantile). The mean is calculated from the spatial footprints of all events that affect a given station. We also carry out an additional analysis to assess the sensitivity of benchmarking results on event spatial footprints to the number of clusters chosen for region South Asia. The benchmarking results are evaluated by the coefficient of determination R^2 of mean event spatial footprints. In addition to the event footprint comparison, we also compare water levels of these two datasets for a range of return periods. In contrast with the empirical return period calculation for the synthetic set, we fit a generalized Pareto distribution to the GTSM-ERA5 water levels using a peak-over-threshold approach to estimate the return levels at each station. We use the same spatially-varying thresholds (selected by the cross-validatory model, see above) to identify extremes and estimate the distribution parameters using the L-moments method. We also calculate the 90% confidence interval (i.e. 5th and 95th percentiles) for the estimated return levels based on Monte Carlo simulation with 1,000 replications. We benchmark the synthetic water levels against the GTSM-ERA5 water levels for the return period of 50, 100, and 500 years. The comparison is shown using three statistical measures, namely (1) coefficient of determination R^2; 2) mean bias (MB); and 3) mean absolute percentage error (MAPE).

Next to comparing the spatial characteristics, we evaluate the event set by assessing its performance in reproducing well-known historic

extreme sea level events as captured in the GTSM-ERA5 dataset. Such events are of most interest for risk management as they occur with low probability but often affect a large spatial area with high impacts. We consider three historic events driven by two types of storms, tropical cyclone (TC) and extratropical cyclone (ETC), in different regions. The three events are: (1) ETC Xaver (2013) in northwestern Europe, (2) TC Sandy (2012) on the northeastern coast of US, and (3) TC-ETC transitioning Sanba (2012) in Japan and South Korea. We first extract these events from the GTSM-ERA5 dataset and identify the respective conditioning stations (i.e. the station with the highest return period during the event). Based on this information, we inspect a number of candidate events that have similar spatial characteristics in the synthetic event set. We then compare the return periods as well as the corresponding water levels station-by-station for these events and fit a linear regression model to the water levels. Lastly, we select the events with the highest R^2 value of the regression model as the most resembling event.

3. Results

3.1. Benchmarking the synthetic dataset on the basis of event spatial footprint sizes

Fig. 2 (A-J) plots the mean sizes of event spatial footprints in the synthetic event set and the GTSM-ERA5 dataset in each. In general, we find a close agreement between two datasets, with an average R^2 of 0.96 across the 10 regions. High agreements are found in Northwest Pacific, Southwest Pacific, Southwest Atlantic, Southeast Atlantic, Indian Ocean, and Oceania, with $R^2 \geq 0.97$; R^2 for Northwest Atlantic and Northeast Atlantic are 0.95 and 0.92, respectively. However, in the South Asia region, the synthetic dataset shows less good performance ($R^2 = 0.85$) and tends to overestimate the mean event footprint at a number of stations, especially for footprint size greater than 150 stations. This is likely the result of using a smaller number of clusters in this region (only 6 clusters were identified), whilst the numbers of clusters used in other regions are substantially higher (ranging from 20 to 30), see Table S1. Potentially unobserved spatial patterns can be explored by using higher cluster numbers. This is proved by the figure of sensitivity analysis result
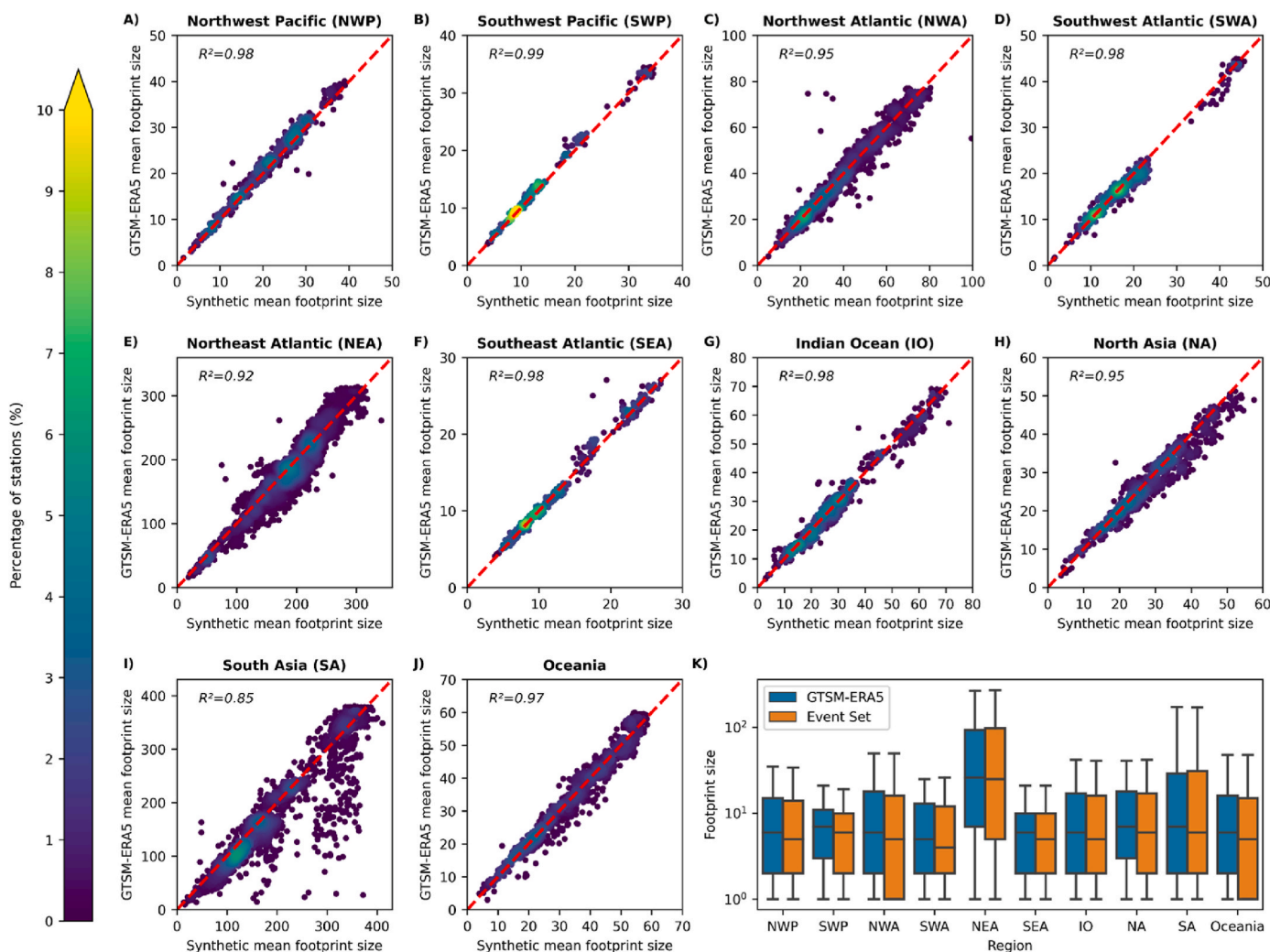

Fig. 2. Comparison of event spatial footprint sizes between the synthetic event set and the GTSM-ERA5 dataset. The panel shows **A-J)** scatter density plots of the mean footprint sizes for each of the 10 regions where each point in the subplots represents a station, and **K)** boxplot of all footprint sizes for the 10 regions. The density of mean footprint sizes is calculated using the two-dimensional histogram method. Light yellow indicates dense footprint sizes while dark purple shows less frequent footprint sizes, with the color bar indicating the percentage of stations (in a region) featuring a given footprint size. The red dashed line represents the perfect fit line. In the boxplot, black lines represent the median event footprint sizes. The box edges correspond to the 25th and 75th percentiles, while the whiskers correspond to the 5th and 95th percentiles. Outliers are not shown. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

(Fig. S12) which shows that with an increasing number of clusters, the agreement becomes increasingly closer (i.e. higher R^2). In addition, we note that even for regions showing a good agreement, there are minor differences in the mean event footprint sizes between the two datasets. This is because the synthetic dataset is much longer (10,000 years) and therefore contains a considerably larger sample of possible events at each station than the GTSM-ERA5 dataset (40 years). This is reflected in

the comparison of all event footprint sizes (Fig. 2K). Overall, the comparison of spatial footprint sizes demonstrates that our approach can produce extreme sea level events with observed footprints from the reanalysis dataset as well as many new events with realistic spatial footprints.

Fig. 2 also shows the range of mean event footprint sizes for each region. We observe relatively small mean event footprints in five regions including Northwest Pacific, Southwest Pacific, Southwest Atlantic,
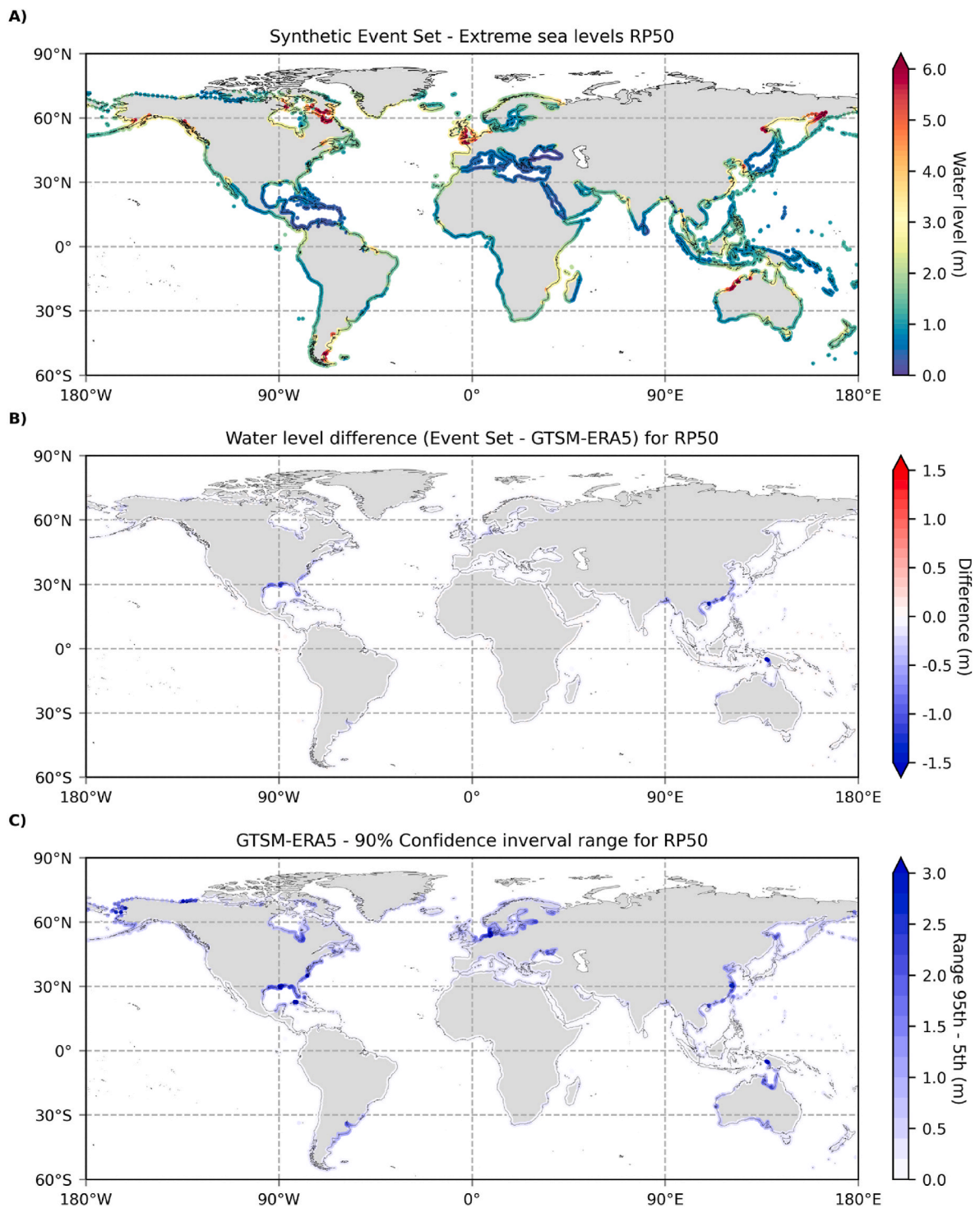
Fig. 3. Global maps of extreme sea levels with a 50-year return period (RP50). The panel shows **A)** extreme sea levels (m) derived from the synthetic event set, **B)** difference (m) in water levels between the synthetic set and GTSM-ERA5, and **C)** the 90% confidence interval range (m) of GTSM-ERA5. The water level difference is calculated by subtracting GTSM-ERA5 return levels from those of the synthetic set. The confidence interval range is given by the difference between the 5th and 95th percentiles.

Southeast Atlantic, and North Asia, with the maxima footprints being 27–52 stations. Three regions (Northwest Atlantic, Indian Ocean, and Oceania) show higher mean footprint sizes with the maxima of 78, 69, and 60 stations, respectively. The maximum mean footprint size is highly correlated to the number of stations in each cluster as the most extreme events are likely to impact nearly all locations within a cluster. Therefore, regions with a larger number of stations in each cluster tend to characterize notably larger footprint sizes; this is especially the case in Northeast Atlantic and South Asia. It is worth noting that there are gaps in mean event footprint sizes for some regions (Fig. 2A, B, D). This is caused by the different numbers of stations in each cluster. Coastlines with complex geographies require more clusters to identify the spatial patterns; therefore, such clusters have a relatively smaller spatial extent and therefore involve a smaller number of stations. One exception is the region South Asia where its coastlines are complex (e.g. many small islands); however, only 6 clusters are identified for this region. In addition, these gaps can arise from various oceanographic drivers (e.g. astronomical tides, TC- and ETC-induced surge levels) and their interactions as different drivers can cause extreme events of distinct spatial footprints.

In addition to the mean footprint size, Fig. 2 shows the mean footprint density. The density is defined as the percentage of stations experiencing a particular footprint size. We find that each region is characterized by distinct patterns in the density distribution of mean footprint sizes. For example, Southwest Pacific tends to experience events with smaller footprint sizes more often, where two frequent mean footprints of approximately 10 and 14 stations can be observed with density exceeding 10% and 8%, respectively. Similar patterns can be found in Southwest Atlantic and Southeast Atlantic, while extreme events impacting larger areas (i.e. more stations affected) are more frequent in Northeast Atlantic and South Asia with the highest density of around 4% (footprint size 150–200 stations) and 6% (footprint size 100–150 stations), respectively. In other regions including North Asia and Oceania, the mean event footprint sizes are nearly evenly distributed with the highest density being around 4% (16–24 stations) and 3% (15–25 stations).

3.2. Benchmarking the synthetic dataset on the basis of return levels

Fig. 3A shows the global map of the extreme sea levels of the synthetic event set for the 50-year return period (RP50). The RP50 water levels exceed 5.0 m in areas with a shallow and wide continental shelf and a stormy climate. Such regions are Gulf of Alaska (West Canada), Hudson Strait (Northeast Canada), Argentina Sea (Southeast Argentina), North Sea (Northwest Europe), the northwestern coast of Australia, Yellow Sea (Northeast China), and Sea of Okhotsk (Northeast Russia). Low water levels (<1.0 m) are found particularly in equatorial areas with a steep ocean topography, such as the Congo basin in Africa and the East Indies (from Sumatra to New Guinea) in Asia. Several other regions also experience low water levels, including the Caribbean area, the Mediterranean region, and Sea of Japan.

The difference in RP50 water levels between the synthetic event set and GTSM-ERA5 is shown in Fig. 3B. The synthetic set shows an overall high agreement with the GTSM-ERA5 dataset, with R^2 of 0.99 (see

Table 1
Model performance of the water levels between the synthetic event set and GTSM-ERA5 for 50-, 100-, and 500-year return periods. The model performance is assessed by three statistical measures which are coefficient of determination R^2, mean bias (MB), and mean absolute percentage error (MAPE). The standard deviations (SDs) of MB and MAPE are shown in the brackets.

Return period (year)	R^2	MB (m)	MAPE (%)
50	0.99	−0.04 (0.12)	2.62 (3.88)
100	0.98	−0.05 (0.16)	3.11 (4.77)
500	0.88	−0.08 (0.55)	4.22 (7.08)

Table 1). Averaged over all the stations, the mean bias (MB) is −0.04 m with a standard deviation (SD) of 0.12 m, while the mean absolute percentage error (MAPE) is 2.62% (SD 3.88%). This indicates that compared to GTSM-ERA5, the synthetic set slightly underestimates the water levels for RP50. At around 93% of the locations, the water level differences are smaller than 0.10 m. Large differences (>1.0 m) are found in the Gulf Coast of US, the coastline of Bangladesh, southern China, and the coast of southern Papua in Indonesia. However, the ranges of the 90% confidence intervals for these regions are also larger than 1.0 m (Fig. 3C). Extreme sea level events in these regions (except southern Papua) are typically caused by tropical cyclones (TCs). From the 40-year GTSM-ERA5 input data, we can only capture too few (or many) or too weak (or extreme) TC-induced events because of the stochastic nature of TCs. This later results in an incomplete sample and misestimation of such events in the synthetic set.

The synthetic event set shows less good performance in water levels for higher return periods. For the RP100 water levels, the MB is −0.05 m (SD 0.16 m) and the MAPE is 3.11% (SD 4.77%) while the MB and MAPE of RP500 water levels are −0.08m (SD 0.55m) and 4.22% (SD 7.08%), respectively (see Table 1). Underestimations of return levels are found in the same hotspots as those identified in the RP50 water levels, but with considerably larger differences (Figs. S13–14). As in previous case, the 90% confidence interval ranges increase with higher return periods and are larger than the water level differences for these regions. This suggests that at least part of the increased differences can be caused by the uncertainties from the extreme value analysis to derive return levels for the GTSM-ERA5 dataset. With higher return periods, fitting a continuous GPD distribution to short-term data (in this study, 40-year water level time series) typically leads to more biased (higher) water levels.

3.3. The capability of reproducing historic events

Fig. 4 compares the return periods and respective water levels between these historic events in the GTSM-ERA5 dataset and the resembling events with the most similar spatial characteristics in the synthetic set. Visual inspection shows that the spatial patterns of the events in the two datasets are similar but the synthetic events are more extreme according to the return periods. For ETC Xaver (Fig. 4A), large return periods are found on the west coast of UK with the highest return period of 525 years in the Wash, while a similar pattern is observed in the synthetic event (Fig. 4B) but with a much higher return period of over 1,000 years at the same location. Similarly, the synthetic event shows higher return periods along the coast of northeast and south UK and parts of the French coastline. In contrast, the Belgian and Dutch coasts experience lower return-period water levels in the synthetic event. When looking at the station-by-station comparison of water levels (Fig. 4C), we find that the synthetic event shows good performance ($R^2 = 0.91$) with a mean bias of −0.05 m and a standard deviation of 0.33 m.

For TC Sandy, the water levels show lower agreement, with a R^2 of 0.77 (mean bias: 0.01m, standard deviation: 0.29 m). However, the spatial patterns of return periods show a much better comparison. Both events depict large return periods (>100 years) on the coastline of New Jersey (NJ), Long Island and the Connecticut (CT). As in the previous case, higher return periods in the synthetic event are found. For example, only one station experiences a return period of over 1,000 years in the GTSM-ERA5 event while the return periods at 18 stations exceed 1,000 years in the synthetic event. Further south, along the coasts of Delaware (DE), Maryland (MA), and Virginia (VA), the return periods are lower (<20 years) in both events.

The spatial comparison of TC-ETC Sanba is shown in Fig. 4 (H-J). For water levels, the synthetic event shows the best agreement (R^2: 0.93) among the three events. The mean bias and the standard deviation are 0.03 m and 0.12 m, respectively. For both events, the northwestern coast of Kyushu shows high return periods (most stations show >500-year return periods). Along the coastline of Chugoku, Shikoku, and the south
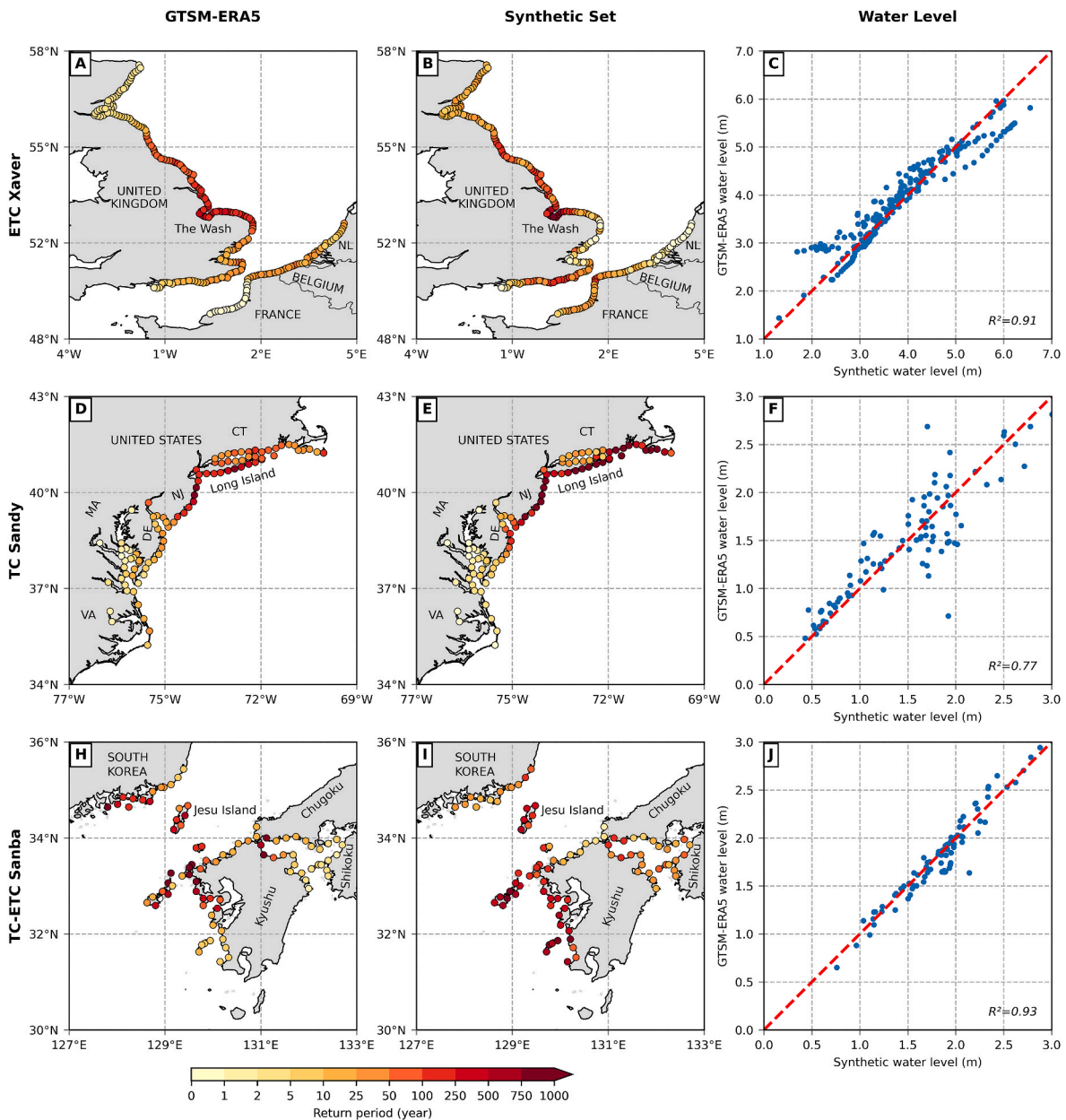
Fig. 4. Comparison between the GTSM-ERA5 historic events and simulated events with similar spatial characteristics from the synthetic event set. Three historic floods are: ETC Xaver (**A-C**), TC Sandy (**D-F**), and TC-ETC Sanba (**H-J**). Return periods of these events are plotted for the GTSM-ERA5 dataset (first column) and the synthetic event set (central column), while the comparison of corresponding water levels is shown in the right column.

coast of South Korea, the return periods in the synthetic event are slightly overestimated compared to the GTSM-ERA5 event. On Jesu Island, return periods are in good agreement.

This comparison demonstrates that the spatial dependence of historic extreme sea level events driven by different types of storms is well captured in our dataset. Besides, we find that extreme events can affect not only contiguous but unconnected stretches of coast, and sometimes even several countries, as observed in ETC Xaver and ETC-TC Sanba. We also exemplify that the return periods of widespread events are spatially heterogeneous where the largest difference in return periods of different locations within an event can be more than 1,000 years.

4. Discussion and conclusion

We have presented the first global synthetic dataset of spatially-dependent extreme sea level events for the current climate, using an

event-based conditional multivariate exceedance model. This dataset is comprised of 10,000 years of individual extreme events with realistic spatial dependence along the global coastline. Using such a dataset can help overcome the inaccurate spatial dependence assumption (i.e. the at-site univariate approach) and the sampling issue (i.e. an incomplete sample of extreme events derived from the short-term input data) which have limited past large-scale extreme sea level and coastal flood studies. The benchmarking of the synthetic event set against the GTSM-ERA5 reanalysis dataset demonstrates that our approach can well capture the spatial characteristics in extreme sea level events. The mean event footprints have an average R^2 of 0.96 across the globe. The benchmarking on water levels also shows a good performance with a global mean bias of -0.04 m for the 50-year return period, although differences are relatively large in the US Gulf Coast, the Bangladesh coastline, southern China, and southern Papua in Indonesia. The comparison of historic extreme sea level events indicates that the synthetic set can

produce events with similar spatial patterns of widespread historic events but with synthetic events to be more extreme.

To generate the synthetic dataset, we first apply the K-Means clustering model to estimate the maximum spatial extent of extreme sea level events for dependence calculation. The spatial footprints were previously examined in several studies (Enríquez et al., 2020; Haigh et al., 2016; Stephens et al., 2020). For the UK coast, four broad types of spatial footprints of extreme sea level events were identified (Haigh et al., 2016). All these four footprints are reflected in the clusters we obtain for the Northeast Atlantic region (Fig. S6) but with our clusters more extended. For example, the category one footprint in the study of Haigh et al. (2016) is captured in cluster 10 in our study which extends to the coastline of northeastern France. We also demonstrate the benefit of using our synthetic dataset over using a reanalysis dataset of 40 years for calculating return periods. Compared to the empirically derived return periods, fitting a continuous GPD distribution to short-timespan data typically leads to more uncertain/biased water levels (especially for low-probability events), although a few exceptions are found in several TC-prone regions. With the comparison of historic sea level events, we show the importance of accounting for spatial dependence in large-scale extreme sea level assessments. We find that extreme sea level events can affect a large spatial area; not only contiguous but unconnected stretches of coast, and sometimes even several countries can be impacted by the same event. Besides, the return periods of water levels at different locations within an event are spatially heterogeneous where the largest difference in return periods can be more than 1,000 years. Hence, using the at-site univariate assumption (i.e. spatially-constant return periods) would strongly misestimate the flood hazard and risk, as documented in several river flood risk studies (Jongman et al., 2014; Metin et al., 2020). In addition, this misestimation is expected to become larger as the spatial scale increases (Nguyen et al., 2020).

In our analysis, we assume that the spatial dependence only exists within the clusters and therefore stations outside these clusters have zero dependence. However, it can be expected that there will be some events which show different spatial footprints that extend beyond the clustering boundaries. An improvement could be made by also considering the dependence between adjacent clusters as previously done in studies for river flooding (Quinn et al., 2019; Wing et al., 2020). Similarly, we consider a 3-day window for ensuring independence although this time lag can vary for specific locations. Future work is recommended to use different lags such as 3-day maxima in a 9-day window and to further assess the sensitivity to these assumptions. In addition, we only evaluate the synthetic dataset by benchmarking against the modeled GTSM-ERA5 dataset. The evaluation of the synthetic dataset could be extended to validation against observational data at tidal gauges such as the GESLA Version 3 (Haigh et al., 2022). For locations where gauge records are scarce or too short, satellite altimetry (Cazenave et al., 2022; Tellman et al., 2021) becomes an alternative. By comparing event spatial footprint sizes, we find that there are gaps in footprints for some regions. Future analysis can investigate the relation of different footprints to the complexity of the coastal geography and the oceanographic drivers (Rueda et al., 2017). Benchmarking on return levels shows that the synthetic dataset has a lower performance in TC-prone regions because of the stochastic nature of TCs and corresponding incomplete TC-induced events sampled from the input dataset. Future work could therefore consider using additional synthetic datasets for a long time series of TC activity, such as the Synthetic Tropical cyclone geneRation Model (STORM) (Bloemendaal et al., 2020), and the corresponding return periods from the COastal dAtaset of Storm Tide Return Periods (COAST-RP) dataset (Dullaart et al., 2021). Waves are also an important component contributing to extreme sea levels and associated spatial dependence (Idier et al., 2019; Santos et al., 2017). We therefore recommend to include wave setup in the future analyses. Although we seek to assess the spatial dependence of all possible sea level events, our approach, based on extreme value theory, cannot characterize sea levels that drive chronic tidal flood events (Ghanbari

et al., 2019). With relative sea level rise, these events are occurring more frequently (Hague et al., 2022; Sweet and Park, 2014) and their impacts are increasing in some coastal regions (Moftakhari et al., 2018). An interesting future research direction would be to design a framework that can capture the spatial dependence of tidal water levels, and to incorporate frequent but minor flood events in large-scale hazard and risk assessment, as is already done for local case studies (e.g. Habel et al., 2020; Hino et al., 2019).

Our global synthetic dataset provides a first step toward improving global-scale coastal flood impact assessments by explicitly accounting for the spatial dependence of extreme sea levels. For example, this dataset can be used to develop coastal flood hazard maps with realistic spatial dependence for deriving improved large-scale spatially-dependent risk profiles than simply the average annual loss that have been applied in current risk studies. Such information is especially useful for international communities and organizations (such as the European Union) where the flood management responsibilities are shared between member states. Furthermore, this dataset can be used to examine how the spatial dependence can affect the risk estimate of coastal flooding, as previously done at local scales (Lamb et al., 2010) and for river flooding (Jongman et al., 2014; Metin et al., 2020; Nguyen et al., 2020).

CRediT authorship contribution statement

Huazhi Li: Conceptualization, Methodology, Software, Formal analysis, Investigation, Data Curation, Writing - Original Draft, Writing - Review & Editing, Visualization. Toon Haer: Conceptualization, Methodology, Investigation, Writing - Review & Editing, Supervision. Anaïs Couasnon: Methodology, Investigation, Writing - Review & Editing. Alejandra Enríquez: Conceptualization, Methodology, Software, Writing - Review & Editing. Sanne Muis: Conceptualization, Methodology, Investigation, Writing - Review & Editing. Philip Ward: Conceptualization, Methodology, Investigation, Writing - Review & Editing, Supervision.

Data and materials availability

All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. The GTSM-ERA5 reanalysis 10-min total water level time series (1979–2018) are available via https://cds.climate.copernicus.eu/cdsapp#!/dataset/sis-water-level-change-timeseries. The spatially-dependent global synthetic dataset of extreme sea level events is publicly available at SURFsara repository: https://surfdrive.surf.nl/files/index.php/s/MUtcaAOn20Z MgX9. The scripts to generate this dataset are available on GitHub: https://github.com/Huazhi-Li/Spatial_dependence. Other codes used to

produce results presented in the paper and central to its main claims are available upon request from the authors.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data and codes are made publicly available.

Acknowledgments

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.wace.2023.100596.

References

Bloemendaal, N., Haigh, I.D., de Moel, H., Muis, S., Haarsma, R.J., Aerts, J.C.J.H., 2020. Generation of a global synthetic tropical cyclone hazard dataset using STORM. Sci. Data 7 (1 7), 1–12. https://doi.org/10.1038/s41597-020-0381-2, 2020.

Brunner, M.I., Furrer, R., Favre, A.C., 2019. Modeling the spatial dependence of floods using the Fisher copula. Hydrol. Earth Syst. Sci. 23, 107–124. https://doi.org/10.5194/hess-23-107-2019.

Cazenave, A., Gouzenes, Y., Birol, F., Leger, F., Passaro, M., Calafat, F.M., Shaw, A., Nino, F., Legeais, J.F., Oelsmann, J., Restano, M., Benveniste, J., 2022. Sea level along the world's coastlines can be measured by a network of virtual altimetry stations. Commun. Earth Environ. 3 (1 3), 1–9. https://doi.org/10.1038/s43247-022-00448-z, 2022.

Couasnon, A., Sebastian, A., Morales-Nápoles, O., 2018. A Copula-Based Bayesian Network for Modeling Compound Flood Hazard from Riverine and Coastal Interactions at the Catchment Scale: an Application to the Houston Ship Channel, vol. 10. Texas. Water, Switzerland. https://doi.org/10.3390/w10091190.

CRED, 2023. 2022 disaster in numbers [WWW Document]. URL https://cred.be/sites/default/files/2022_EMDAT_report.pdf.

CRED, 2018. Natural disasters 2018 [WWW Document]. Emergency Events Database. URL. https://emdat.be/%0Asites/default/files/adsr_2018.pdf. (Accessed 24 February 2023).

Dullaart, J.C.M., Muis, S., Bloemendaal, N., Aerts, J.C.J.H., 2020. Advancing global storm surge modelling using the new ERA5 climate reanalysis. Clim. Dynam. 54, 1007–1021. https://doi.org/10.1007/S00382-019-05044-0/FIGURES/6.

Dullaart, J.C.M., Muis, S., Bloemendaal, N., Chertova, M.V., Couasnon, A., Aerts, J.C.J.H., 2021. Accounting for tropical cyclones more than doubles the global population exposed to low-probability coastal flooding. Commun. Earth Environ. 2 (1 2), 1–11. https://doi.org/10.1038/s43247-021-00204-9, 2021.

Enríquez, A.R., Wahl, T., Marcos, M., Haigh, I.D., 2020. Spatial footprints of storm surges along the global coastlines. J Geophys Res Oceans 125. https://doi.org/10.1029/2020JC016367.

Esteves, L.S., 2013. Consequences to flood management of using different probability distributions to estimate extreme rainfall. J. Environ. Manag. 115, 98–105. https://doi.org/10.1016/j.jenvman.2012.11.013.

Falter, D., Dung, N.v., Vorogushyn, S., Schröter, K., Hundecha, Y., Kreibich, H., Apel, H., Theiss*elmann, F., Merz, B., 2016. Continuous, large-scale simulation model for flood risk assessments: proof-of-concept. J Flood Risk Manag. https://doi.org/10.1111/jfr3.12105.

Falter, D., Schröter, K., Dung, N.V., Vorogushyn, S., Kreibich, H., Hundecha, Y., et al., 2015. Spatially coherent flood risk assessment based on long-term continuous simulation with a coupled model chain. J. Hydrol. (Amst.) 524, 182–193. https://doi.org/10.1016/j.jhydrol.2015.02.021.

Feng, K., Ouyang, M., Lin, N., 2022. Tropical cyclone-blackout-heatwave compound hazard resilience in a changing climate. Nat. Commun. 13 (1 13), 1–11. https://doi.org/10.1038/s41467-022-32018-4, 2022.

Ghanbari, M., Arabi, M., Obeysekera, J., Sweet, W., 2019. A coherent statistical model for coastal flood frequency analysis under nonstationary sea level conditions. Earth's Future 7, 162–177. https://doi.org/10.1029/2018EF001089.

Habel, S., Fletcher, C.H., Anderson, T.R., Thompson, P.R., 2020. Sea-level rise induced multi-mechanism flooding and contribution to urban infrastructure failure. Sci. Rep. 10, 1–12. https://doi.org/10.1038/s41598-020-60762-4.

Hague, B.S., Jones, D.A., Jakob, D., McGregor, S., Reef, R., 2022. Australian coastal flooding trends and forcing factors. Earth's Future 10. https://doi.org/10.1029/2021EF002483.

Haigh, I.D., Marcos, M., Talke, S.A., Woodworth, P.L., Hunter, J.R., Hague, B.S., Arns, A., Bradshaw, E., Thompson, P., 2022. GESLA Version 3: a major update to the global higher-frequency sea-level dataset. Geosci Data J. https://doi.org/10.1002/GDJ3.174.

Haigh, I.D., Wadey, M.P., Wahl, T., Ozsoy, O., Nicholls, R.J., Brown, J.M., Horsburgh, K., Gouldby, B., 2016. Spatial and temporal analysis of extreme sea level and storm surge events around the coastline of the UK. Sci. Data. https://doi.org/10.1038/sdata.2016.107.

Hallegatte, S., Green, C., Nicholls, R.J., Corfee-Morlot, J., 2013. Future flood losses in major coastal cities. Nat. Clim. Change. https://doi.org/10.1038/nclimate1979.

Hastie, T., Tibshirani, R., Friedman, J., 2009. The Elements of Statistical Learning: Data Mining, Inference, and Prediction, second ed. Springer Series in Statistics, New York, NY. https://doi.org/10.1007/978-0-387-84858-7. Springer New York.

Heffernan, J.E., 2001. A directory of coefficients of tail dependence. Extremes 3, 279–290. https://doi.org/10.1023/A:1011459127975.

Heffernan, J.E., Tawn, J.A., 2004. A conditional approach for multivariate extreme values (with discussion). J R Stat Soc Series B Stat Methodol 66, 497–546. https://doi.org/10.1111/j.1467-9868.2004.02050.x.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R.J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., Thépaut, J.N., 2020. The ERA5 global reanalysis. Q. J. R. Meteorol. Soc. 146, 1999–2049. https://doi.org/10.1002/qj.3803.

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

Hino, M., Belanger, S.T., Field, C.B., Davies, A.R., Mach, K.J., 2019. High-tide flooding disrupts local economic activity. Sci. Adv. 5 https://doi.org/10.1126/sciadv.aau2736.

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

Jongman, B., Hochrainer-Stigler, S., Feyen, L., Aerts, J.C.J.H., Mechler, R., Botzen, W.J.W., Bouwer, L.M., Pflug, G., Rojas, R., Ward, P.J., 2014. Increasing stress on disaster-risk finance due to large floods. Nat. Clim. Change 4, 264–268. https://doi.org/10.1038/nclimate2124.

Keef, C., Papasthopoulos, I., Tawn, J.A., 2013a. Estimation of the conditional distribution of a multivariate variable given that one of its components is large: additional constraints for the Heffernan and Tawn model. J. Multivariate Anal. 115, 396–404. https://doi.org/10.1016/j.jmva.2012.10.012.

Keef, C., Tawn, J.A., Lamb, R., 2013b. Estimating the probability of widespread flood events. Environmetrics 24, 13–21. https://doi.org/10.1002/env.2190.

Keef, C., Tawn, J.A., Svensson, C., 2009. Spatial risk assessment for extreme river flows. J R Stat Soc Ser C Appl Stat 58, 601–618. https://doi.org/10.1111/j.1467-9876.2009.00672.x.

Kirezci, E., Young, I.R., Ranasinghe, R., Muis, S., Nicholls, R.J., Lincke, D., Hinkel, J., 2020. Projections of global-scale extreme sea levels and resulting episodic coastal flooding over the 21st Century. Sci. Rep. 10 (1 10), 1–12. https://doi.org/10.1038/s41598-020-67736-6, 2020.

Lamb, R., Keef, C., Tawn, J., Laeger, S., Meadowcroft, I., Surendran, S., Dunning, P., Batstone, C., 2010. A new method to assess the risk of local and widespread flooding on rivers and coasts. J Flood Risk Manag 3, 323–336. https://doi.org/10.1111/j.1753-318X.2010.01081.x.

Little, C.M., Horton, R.M., Kopp, R.E., Oppenheimer, M., Vecchi, G.A., Villarini, G., 2015. Joint Projections of US East Coast Sea Level and Storm Surge. https://doi.org/10.1038/NCLIMATE2801.

Makkonen, L., 2006. Plotting positions in extreme value analysis. J. Appl. Meteorol. Climatol. 45, 334–340. https://doi.org/10.1175/JAM2349.1.

Mawdsley, R.J., Haigh, I.D., Wells, N.C., 2015. Global secular changes in different tidal high water, low water and range levels. Earth's Future 3, 66–81. https://doi.org/10.1002/2014EF000282.

Metin, A.D., Dung, N.V., Schröter, K., Vorogushyn, S., Guse, B., Kreibich, H., Merz, B., 2020. The role of spatial dependence for large-scale flood risk estimation. Nat. Hazards Earth Syst. Sci. 20, 967–979. https://doi.org/10.5194/nhess-20-967-2020.

Moftakhari, H.R., AghaKouchak, A., Sanders, B.F., Allaire, M., Matthew, R.A., 2018. What is nuisance flooding? Defining and monitoring an emerging challenge. Water Resour. Res. 54, 4218–4227. https://doi.org/10.1029/2018WR022828.

Muis, S., Aerts, J., Antolínez, J.A.Á., Dullaart, J., Duong, T.M., Erikson, L., Haarmsa, R., Apecechea, M.I., Mengel, M., Bars, D. Le, O'Neill, A., Ranasinghe, R., Roberts, M., Verlaan, M., Ward, P.J., Yan, K., 2022. Global Projections of Storm Surges Using High-Resolution CMIP6 Climate Models: Validation, Projected Changes, and Methodological Challenges. https://doi.org/10.1002/ESSOAR.10511919.1.

Muis, S., Apecechea, M.I., Dullaart, J., de Lima Rego, J., Madsen, K.S., Su, J., Yan, K., Verlaan, M., 2020. A high-resolution global dataset of extreme sea levels, tides, and storm surges, including future projections. Front. Mar. Sci. 7 https://doi.org/10.3389/fmars.2020.00263.

Nguyen, V.D., Metin, A.D., Alfieri, L., Vorogushyn, S., Merz, B., 2020. Biases in national and continental flood risk assessments by ignoring spatial dependence. Sci. Rep. 10, 1–8. https://doi.org/10.1038/s41598-020-76523-2.

NOAA, 2022. U . S . Billion-dollar weather & climate disasters 1980-2021 [WWW Document]. URL. https://www.ncei.noaa.gov/access/billions/events.pdf.

Northrop, P.J., Attalides, N., Jonathan, P., 2017. Cross-validatory extreme value threshold selection and uncertainty with application to ocean storm severity. J R Stat Soc Ser C Appl Stat 66, 93–120. https://doi.org/10.1111/rssc.12159.

Paprotny, D., Morales-Nápoles, O., Nikulin, G., 2016. Extreme sea levels under present and future climate: a pan-European database. In: E3S Web of Conferences. https://doi.org/10.1051/e3sconf/20160702001.

Perez, J., Menendez, M., Camus, P., Mendez, F.J., Losada, I.J., 2015. Statistical multi-model climate projections of surface ocean waves in Europe. Ocean Model. 96, 161–170. https://doi.org/10.1016/J.OCEMOD.2015.06.001.

Pörtner, H.-O., Roberts, D.C., Tignor, M., Poloczanska, E.S., Mintenbeck, K., Alegría, A., Craig, M., Langsdorf, S., Löschke, S., Möller, V., Okem, A., Rama, B. (Eds.), 2022. IPCC, 2022: Climate Change 2022: Impacts, Adaptation, and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change.

Quinn, N., Bates, P.D., Neal, J., Smith, A., Wing, O., Sampson, C., Smith, J., Heffernan, J., 2019. The spatial dependence of flood hazard and risk in the United States. Water Resour. Res. 55, 1890. https://doi.org/10.1029/2018WR024205. –1911.

Read, L.K., Vogel, R.M., 2015. Reliability, return periods, and risk under nonstationarity. Water Resour. Res. 51, 6381–6398. https://doi.org/10.1002/2015WR017089.

Rueda, A., Camus, P., Tomás, A., Vitousek, S., Méndez, F.J., 2016. A multivariate extreme wave and storm surge climate emulator based on weather patterns. Ocean Model. 104, 242–251. https://doi.org/10.1016/j.ocemod.2016.06.008.

Rueda, A., Vitousek, S., Camus, P., Tomás, A., Espejo, A., Losada, I.J., Barnard, P.L., Erikson, L.H., Ruggiero, P., Reguero, B.G., Mendez, F.J., 2017. A global classification of coastal flood hazard climates associated with large-scale oceanographic forcing/ 704/106/829/2737/704/4111/141/129 article. Sci. Rep. 7 https://doi.org/10.1038/s41598-017-05090-w.

Santos, V.M., Haigh, I.D., Wahl, T., 2017. Spatial and temporal clustering analysis of extreme wave events around the UK coastline. Journal of Marine Science and Engineering 2017 5, 28. https://doi.org/10.3390/JMSE5030028. Page 28 5.

Serinaldi, F., Kilsby, C.G., 2017. A blueprint for full collective flood risk estimation: demonstration for European river flooding. Risk Anal. 37, 1958–1976. https://doi.org/10.1111/risa.12747.

Stephens, S.A., Bell, R.G., Haigh, I.D., 2020. Spatial and temporal analysis of extreme storm-tide and skew-surge events around the coastline of New Zealand. Nat. Hazards Earth Syst. Sci. 20, 783–796, 10.5194/nhess-20-783-2020.

Sweet, W.V., Park, J., 2014. From the extreme to the mean: acceleration and tipping points of coastal inundation from sea level rise. Earth's Future 2, 579–600. https://doi.org/10.1002/2014ef000272.

Taylor, K.E., Stouffer, R.J., Meehl, G.A., 2012. An overview of CMIP5 and the experiment design. Bull. Am. Meteorol. Soc. https://doi.org/10.1175/BAMS-D-11-00094.1.

Tellman, B., Sullivan, J.A., Kuhn, C., Kettner, A.J., Doyle, C.S., Brakenridge, G.R., Erickson, T.A., Slayback, D.A., 2021. Satellite imaging reveals increased proportion

of population exposed to floods. Nature 596, 80–86. https://doi.org/10.1038/s41586-021-03695-w.

Tiggeloven, T., De Moel, H., Winsemius, H.C., Eilander, D., Erkens, G., Gebremedhin, E., Diaz Loaiza, A., Kuzma, S., Luo, T., Iceland, C., Bouwman, A., Van Huijstee, J., Ligtvoet, W., Ward, P.J., 2020. Global-scale benefit-cost analysis of coastal flood adaptation to different flood risk drivers using structural measures. Nat. Hazards Earth Syst. Sci. 20, 1025–1044. https://doi.org/10.5194/nhess-20-1025-2020.

Towe, R.P., Tawn, J.A., Lamb, R., Sherlock, C.G., 2019. Model-based inference of conditional extreme value distributions with hydrological applications. Environmetrics 30, e2575. https://doi.org/10.1002/ENV.2575.

Vitousek, S., Barnard, P.L., Fletcher, C.H., Frazer, N., Erikson, L., Storlazzi, C.D., 2017. Doubling of coastal flooding frequency within decades due to sea-level rise. Sci. Rep. 7, 1399. https://doi.org/10.1038/s41598-017-01362-7.

Vorogushyn, S., Bates, P.D., de Bruijn, K., Castellarin, A., Kreibich, H., Priest, S., Schröter, K., Bagli, S., Blöschl, G., Domeneghetti, A., Gouldby, B., Klijn, F., Lammersen, R., Neal, J.C., Ridder, N., Terink, W., Viavattene, C., Viglione, A., Zanardo, S., Merz, B., 2018. Evolutionary leap in large-scale flood risk assessment needed. WIREs Water 5, 1–7. https://doi.org/10.1002/wat2.1266.

Vousdoukas, M.I., Mentaschi, L., Hinkel, J., Ward, P.J., Mongelli, I., Ciscar, J.C., Feyen, L., 2020. Economic motivation for raising coastal flood defenses in Europe. Nat. Commun. 11, 1–11. https://doi.org/10.1038/s41467-020-15665-3.

Vousdoukas, M.I., Mentaschi, L., Voukouvalas, E., Verlaan, M., Jevrejeva, S., Jackson, L. P., Feyen, L., 2018. Global probabilistic projections of extreme sea levels show intensification of coastal flood hazard. Nat. Commun. 9 https://doi.org/10.1038/s41467-018-04692-w.

Wahl, T., Haigh, I.D., Nicholls, R.J., Arns, A., Dangendorf, S., Hinkel, J., Slangen, A.B.A., 2017. Understanding extreme sea levels for broad-scale coastal impact and adaptation analysis. Nat. Commun. 8, 1–12. https://doi.org/10.1038/ncomms16075.

Wansouwé, W.E., Somé, S.M., Kokonendji, C.C., 2016. Ake: an r package for discrete and continuous associated kernel estimations. R J. 8, 258–276. https://doi.org/10.32614/rj-2016-045.

Ward, P.J., Jongman, B., Salamon, P., Simpson, A., Bates, P.D., De Groeve, T., Muis, S., De Perez, E.C., Rudari, R., Trigg, M.A., Winsemius, H.C., 2015. Usefulness and limitations of global flood risk models, 2015 Nature Climate Change 5 (8 5), 712–715. https://doi.org/10.1038/nclimate2742.

Wing, O.E.J., Quinn, N., Bates, P.D., Neal, J.C., Smith, A.M., Sampson, C.C., Coxon, G., Yamazaki, D., Sutanudjaja, E.H., Alfieri, L., 2020. Toward global stochastic river flood modeling. Water Resour. Res. 56, e2020WR027692 https://doi.org/10.1029/2020WR027692.

Winter, B., Schneeberger, K., Dung, N.v., Huttenlau, M., Achleitner, S., Stötter, J., Merz, B., Vorogushyn, S., 2019. A continuous modelling approach for design flood estimation on sub-daily time scale. Hydrol. Sci. J. 64, 539–554. https://doi.org/10.1080/02626667.2019.1593419.

Wyncoll, D., Haigh, I., Gouldby, B., Hames, D., Laeger, S., Wall, A., et al., 2016. Spatial analysis and simulation of extreme coastal flooding scenarios for national-scale emergency planning. E3S Web of Conf. 7, 1–6. https://doi.org/10.1051/e3sconf/20160701001.