

# Assessing extreme total water levels across Europe for large-scale coastal flood analysis

Camila Cotrim<sup>1</sup>, Alexandra Toimil<sup>1</sup>, Iñigo J. Losada<sup>1</sup>, Melisa Menéndez<sup>1</sup>, Hector Lobeto<sup>1</sup>

<sup>1</sup>IHCantabria - Instituto de Hidráulica Ambiental de la Universidad de Cantabria, Santander, 39011, Spain

5 *Correspondence to:* Alexandra Toimil (toimila@unican.es)

**Abstract.** Coastal storm-induced flooding threatens millions of people and infrastructures, highlighting the need for comprehensive flood risk assessments. A key component of these assessments is the spatial characterization of total water level (TWL), the primary driver of coastal flooding. We propose a homogeneous methodology for developing large-scale TWL hindcasts to estimate extreme events, considering possible spatial variabilities in marine dynamics. This methodology is applied to the European coastline, integrating downscaled nearshore waves, storm surges, and tides. The resulting hourly time series of the TWL have a spatial resolution of 1 km and covers the period from 1985 to 2021. Spatial variability is considered in foreshore slopes and extreme value detection thresholds, addressing common simplifications in large-scale studies. In addition to a characterization of extreme events based on the relative contributions of TWL components, sensitivity analyses of the wave contribution, wave data resolution, foreshore slopes, and wave setup formulations are conducted. The tide-dominated Atlantic coast is where the TWL is most affected by the wave dataset. The storm surge-dominated Baltic region exhibits the lowest confidence in estimating TWL return levels, partially due to the data and methods used. The Mediterranean Sea, characterized by a mixed environment, is the most sensitive to the inclusion of wave contribution. A classification of TWL extremes revealed that no regions have extreme events dominated by wave setup when compared to the remaining components, while those dominated by tides show the highest return levels.

## 20 **1 Introduction**

The assessment of coastal impacts resulting from extreme coastal flooding events is essential for understanding and mitigating associated potential risks (Neumann et al., 2015). These risks include, among others, damage to coastal infrastructure and the built environment, as well as impacts on the population and ecosystems (Barnard et al., 2019; Rashid et al., 2021; Rasmussen et al., 2022). The total water level (TWL) is the most commonly used indicator in coastal flooding studies. It is frequently used as a forcing variable for flood models aimed at generating maps of flood extent and depth. Accurate coastal flood estimates require that TWL includes a proper spatial characterization of all relevant components (i.e., waves, storm surges, and astronomical tides) (Cabrita et al., 2024; Kirezci et al., 2020; Pugh, 1987). However, estimating extreme TWL at coastal locations becomes increasingly complex as the spatial scale of the study domain increases.

One key challenge lies in the method used to reconstruct the TWL, which requires the selection and characterization of relevant components at appropriate spatial and temporal scales. We consider TWL as a combination of astronomical tide, storm surge, and a contribution by ocean waves (usually the wave setup) (Pugh & Woodworth, 2014). Among these, assessing the contribution of waves is particularly critical, as it involves evaluating nearshore wave conditions and the associated wave setup, which depending on the formulation can be influenced by the foreshore slope. A second key challenge concerns the choice of an extreme value modeling approach to characterize extreme events in terms of probability of occurrence. Each stage in this process introduces uncertainty, from data selection and the formulation used to calculate wave contributions to TWL, to the definition of extreme events (e.g., threshold selection) and the fitting of statistical distributions to estimate return levels (Hinkel et al., 2021; Toimil et al., 2020).

Initial attempts at assessing coastal flood impacts at large scales relied on the DIVA model (Vafeidis et al., 2008), which provided TWL and return level scenarios. More recent efforts have focused on improving the spatial resolution and temporal extent of large-scale TWL hindcasts. For example, TWL has been globally reconstructed from 1993 to 2013 at a 111 km resolution (Vitousek et al., 2017), from 1979 to 2014 at 70 km resolution (Kirezci et al., 2020), and from 1993 to 2015 at 50 km resolution (Almar et al., 2021). For the European coastline, TWL from 1979 to 2014 was calculated with a 25 km resolution (Vousdoukas et al., 2016), and from 2010 to 2019 with 2.5 km resolution (Le Gal et al., 2023).

With respect to TWL components, storm surges and astronomical tides are commonly included in flood studies, whereas the wave component is often neglected despite its potential importance along many coasts (Muis et al., 2015, 2016; Paprotny et al., 2016, 2018). At large scales, few studies include wave effects, which may take different forms depending on the processes considered. These include: (1) static wave setup, defined as the superelevation of the coastal water surface due to wave breaking; (2) dynamic wave setup which adds the effect of the infragravity swash to the static wave setup; and (3) wave runup, which further includes the incident swash component (Stockdon et al., 2006). While the estimation of infragravity swash remains uncertain at large scales, incident swash generally does not sustain flooding capable of causing coastal damage (Hinkel et al., 2021). For two-dimensional flood modeling, the wave contribution component is usually represented by the static wave setup (hereafter referred to as wave setup). Several methods exist for estimating wave setup at large scales. The most widely used is the empirical approach proposed by Guza & Thornton (1981), which estimates wave setup as a fraction of the significant wave height ( $H_s$ ), namely  $0.2H_s$ . However, this method often leads to overestimation of wave setup (Hinkel et al., 2021). More accurate alternatives involve simplified parametrizations. For instance, the semi-empirical formulation by Stockdon et al. (2006) has been adopted globally (Rueda et al., 2017; Vitousek et al., 2017), as well as the Shore Protection Manual method (Kirezci et al., 2020; USACE, 1984). One way these parametrizations improve accuracy is by incorporating nearshore bathymetry, particularly the foreshore slope (Dodet et al., 2019; Gomes da Silva et al., 2020).

A limitation within the use of wave parametrizations lies in accurately characterizing the foreshore slope at large scales. Despite its importance, this parameter remains difficult to quantify, which limits its practical applicability in many coastal assessments. Consequently, some authors applied a constant slope of 1:30 (Kirezci et al., 2020), while others used a wave setup formulation designed for dissipative beaches (Vitousek et al., 2017), that does not require slope data (Stockdon et al., 2006). Alternatively,

a method developed by Sunamura (1984) allows for the estimation of spatially and temporally varying foreshore slopes based on wave conditions. Compared with locally observed slopes and globally constant values, this method accounts for morphological feedback, given that evolving beach morphology can influence wave contributions to sea level, as considered in Sunamura's formulation (Melet et al., 2020).

Another limitation in these studies is that offshore wave conditions are often used, despite their coarse resolution and limited suitability for accurately assessing coastal impacts. More precise estimates of coastal flooding require nearshore wave data, which better capture the relevant coastal dynamics, as wave setup is not an offshore process (Dodet et al., 2019). The use of nearshore information is however limited at large geographical scales due to modeling complexity and computational demands. Regarding the second key challenge, extreme value modeling can be approached through a variety of statistical frameworks, ranging from the well-established extreme value analysis (EVA), which provides the theoretical foundation for estimating return levels (Coles, 2001), to more refined methods designed to address different challenges. For example, dependence between variables can be treated with copulas (Bevacqua et al., 2020; MacPherson et al., 2019), spatial dependence can be included using Bayesian hierarchical models (Calafat & Marcos, 2020), and limited number of observations can be supported with regional frequency analysis (Collings et al., 2024). Yet, EVA remains the benchmark due to its low computational demands, relative simplicity, and long-standing use making it the baseline for comparison, especially in large-scale studies where reproducibility is essential.

EVA comprises two main steps: the identification of exceedances in a time series and a distribution family to which the identified sample is fitted to estimate return levels. The most commonly used approaches for identifying extreme events include the annual maxima (AM) and the peak-over-threshold (POT) methods, which differ primarily in sample size (Bezak et al., 2014). For an appropriately chosen threshold, POT generally yields larger samples of extremes than AM (Kirezci et al., 2020; Le Gal et al., 2023). However, threshold selection in POT is not straightforward. It must be high enough to exclude non-extreme events but low enough to ensure a sufficiently large sample for statistical analysis (Harley, 2017). In addition, the selected maxima peaks must represent independent extreme events, so a minimum time interval must be considered between two consecutive peaks. A stable threshold based on objective criteria, rather than arbitrary decisions, ensures consistency and more reliable results (Arns et al., 2013). Previous studies have tested different percentiles as the threshold and the 98<sup>th</sup> percentile (P98) of the TWL time series was used globally (Kirezci et al., 2020) while the 97<sup>th</sup> percentile (P97) was adopted for the European coast (Le Gal et al., 2023). Alternatively, a fixed TWL value can be used, although not suitable for large-scale studies where extreme magnitudes vary widely. To address spatial variability, Vousdoukas et al. (2016) adopted a threshold corresponding to an average number of independent events per year across the European coastline, while Vitousek et al. (2017) adjusted the selection of AM to the 3-largest events per year in a global study.

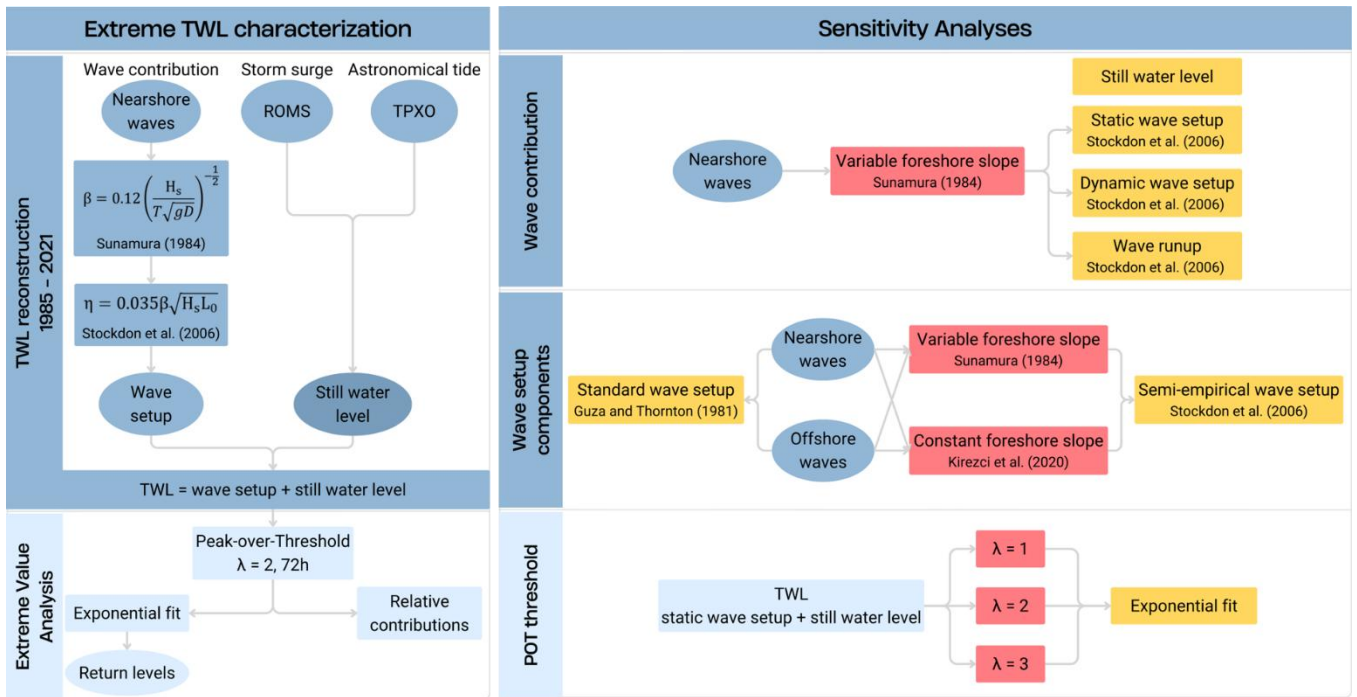
The choice of distribution families for estimating return levels also carries important implications that need to be carefully evaluated depending on regional and data characteristics. For example, Paprotny et al. (2016) applied a Gumbel distribution to the coast of Europe while Vitousek et al. (2017) applied a GEV, both based on a sample of extremes extracted with AM. Meanwhile, Kirezci et al. (2020) applied a Generalized Pareto Distribution (GPD) to a POT selected sample at the global scale.

Nevertheless, the key features of an EVA application go beyond the method adopted to sample extreme events or the statistical model used to fit the data. Depending on regional climatic characteristics, data available, and variables considered (e.g., individual sources, combined drivers, and outputs of coastal hazards), the most suitable approach may vary (Coles, 2001).

100 Accordingly, in this work, we address the above challenges by proposing a methodology for reconstructing large-scale TWL hindcasts and estimating extreme events. Basing on the hypothesis that a standardized strategy to estimate extreme TWL can effectively capture the diverse conditions governing coastal dynamics across large scales, the objective of this study was to create a flexible methodology which considers possible heterogeneous marine climate characteristics relevant in coastal storms. Focus is placed on the generation of the main input necessary for large-scale coastal flooding studies, the extreme TWL. We  
105 apply our method to the European coastline, producing hourly TWL time series at 1 km resolution spanning from 1985 to 2021. This represents the longest and highest spatial resolution TWL time series for Europe to date, incorporating nearshore wave information. Our TWL formulation includes wave setup, storm surges, and astronomical tides. The wave setup is computed semi-empirically using spatially and temporally varying foreshore slopes. Extreme events are identified using the POT method with a spatially variable threshold, and return levels are estimated with an exponential fit. To demonstrate the  
110 robustness of our approach, we conduct a series of sensitivity analyses of the effect different wave component options and POT thresholds have on the assessment of extreme events. For those, return levels extracted from EVA and flooded area derived from a static flood model are used as indicators of the effects certain decisions taken during the extreme TWL estimation process might have on coastal flooding.

## **2 Data and Methods**

115 The methodology developed to estimate extreme TWL at the large-scale is presented in Fig. 1. We reconstructed a TWL hindcast on an hourly basis by linearly summing time series of wave setup, storm surges, and astronomical tides. The wave setup was estimated with a semi-empirical formulation (Stockdon et al., 2006) considering spatially and temporally variable foreshore slopes (Sunamura, 1984). To identify extreme events in the TWL hindcast, we applied the POT method, using a spatially variable threshold. The selected extreme events were fitted to an exponential function to estimate return levels.  
120 Different steps of the methodology were evaluated and validated through sensitivity analyses focusing on: (1) wave contribution; (2) components of wave setup; and (3) the POT threshold.



**Figure 1: Methodology applied in the present study for TWL hindcast reconstruction and extreme value analysis (left). Sensitivity analyses applied to the wave contribution, wave setup components, and POT threshold (right).**

## 125 2.1 Study Area

The study area comprises the European coastline. Prior to reconstructing the TWL hindcast, coastal target points (CTP) were defined along the European coast. These points were used for downscaling wave conditions and reconstructing the TWL, with the objective of creating a continuous line of equally spaced points as close to the coast as possible. Due to the lack of high-resolution homogeneous bathymetry data at this scale, the CTPs were selected based on their relative depth ( $h/L$ ), where  $h$  represents the water depth and  $L$  the corresponding wavelength of the wave peak period with an exceedance probability of 0.1. In total, 51010 CTP with relative depth of approximately 0.1 were chosen, with an interval of 1 km between them and as close as possible to the coast.

## 130 2.2 Climate-related Datasets

The still water level includes the mean sea level, astronomical tide, and storm surge. Here, the local mean sea level is used as benchmark. The hourly time series of the astronomical tide and storm surge were reconstructed from the tidal constituents of the TPXO9 model (Egbert & Erofeeva, 2002) and from a European regional storm surge hindcast developed using the ROMs model (Shchepetkin & McWilliams, 2005), respectively. The astronomical tide provides data in coastal areas at a spatial resolution of  $1/30^\circ$  (approximately 3.5 km), whilst the European storm surge hindcast used provides storm surge information along the European coast at 5-11 km spatial resolution. The offshore waves were obtained from a European regional hindcast

140 developed with the WaveWatchIII numerical model (Tolman, 2009). This provides wave outputs along the European coastline  
 at a spatial resolution of  $1/8^\circ$  (10-15 km). The nearshore wave dataset was developed based on the DOW approach (Camus et  
 al., 2013), which combines numerical wave simulations using the SWAN model with advanced statistical techniques in order  
 to reach 1 km spatial resolution throughout the European coastline. Atmospheric forcings from the ERA5 reanalysis (Hersbach  
 et al., 2020) were used to generate these datasets consistently. A more detailed description of the models and their configuration  
 145 used, as well as an assessment of the data quality in coastal areas by comparison against coastal buoys and tide gauge records  
 is provided in the ‘Sea level and wave datasets’ section of the Supplementary Material and in IHCantabria (2024).

At each CTP, hourly time series were obtained for the astronomical tide, storm surge, offshore wave conditions and nearshore  
 wave conditions. As the offshore wave hindcast, storm surge hindcast, and TPXO databases do not share the same spatial  
 resolution as the nearshore wave dataset, the data from the nearest grid point was assigned to each CTP for those variables.

### 150 **2.3 Total Water Level**

The TWL time series was reconstructed from 1985 to 2021 on an hourly basis at each CTP by linearly adding the time series  
 of astronomical tide, storm surge, and wave setup (1).

$$\text{TWL} = \text{wave setup} + \text{SS} + \text{AT} \quad (1)$$

where SS is the storm surge and AT the astronomical tide.

Wave setup was used to represent the wave contribution to TWL. Using nearshore wave conditions, it was calculated with the  
 155 semi-empirical formulation, (2) and (3) (Stockdon et al., 2006) with a variable intertidal slope, (4) (Sunamura, 1984).

$$\eta = 0.035\beta\sqrt{H_s L_0} \quad (2)$$

$$L_0 = \frac{gT^2}{2\pi} \quad (3)$$

where  $\beta$  is the foreshore slope of the beach, between the low and high tide limits,  $H_s$  is the significant wave height,  $L_0$  is the  
 deep-water wavelength, and T is the wave period.

$$\beta = 0.12 \left( \frac{H_s}{T\sqrt{gD}} \right)^{-\frac{1}{2}} \quad (4)$$

where g is gravitational acceleration and D is the mean sediment grain size, set at 250  $\mu\text{m}$ , representative of fine to medium  
 grain size and following Rueda et al. (2017). The calculated foreshore slope time series for each CTP was capped at 0.20 and  
 160 normalized to ensure a mean close to 0.04, reflecting the global median beach slopes based on locally reported values (Barboza  
 & Defeo, 2015). The approach used to apply the approximation for the foreshore slope followed Melet et al. (2020) and has  
 been validated (see Supplementry Fig. S6).

Upon the TWL reconstruction, the stationarity of the time series was verified in each CTP and a validation of the hindcast was  
 performed by identifying historical coastal storms observed in the literature and comparing to the values estimated here. In  
 165 total, 36 events were validated for different events and across distinct areas of the European coastline. A summary of the

references and locations validated is presented in the Supplementary Material (Table S4). The validation was performed against still water level, static wave setup-based TWL, dynamic wave setup-based TWL, and wave-runup-based TWL, following Stockdon et al. (2006).

170 Since the European coastline reaches high latitudes, the presence of sea ice may affect the impact of the different coastal dynamics considered in this study (e.g., in the Baltic Sea). As all atmospheric-driven dynamics originate from a common forcing, ERA5, the sea ice cover provided by this reanalysis was used to account for the presence or absence of ice. This variable represents the fraction of each grid cell covered by ice. The temporal evolution of ice cover was extracted at the nodes closest to the coastal points. Ice presence was assumed when the coverage exceeded a fraction of 0.5, following the simple ice-blocking scheme used in wave propagation models such as WW3 (e.g., Kumar et al., 2025). Accordingly, when ice was  
175 present, the corresponding data at CTPs were excluded from the analysis, as neither waves nor tides effectively reach the coast under such conditions.

## 2.4 Model Validation Summary

The modeling chain used to reconstruct TWL combines independent, state-of-the-art datasets for tides, storm surges, and waves, each of which was validated against available observations prior to their integration. Table 1 presents a summary the  
180 validation results. Astronomical tides derived from TPXO9v5 were validated against 48 tide-gauge records distributed along the European coastline, showing high skill in reproducing tidal amplitudes and phases, with root mean square errors typically below 10 cm, biases close to zero, and Pearson correlation coefficients exceeding 0.9 at most locations (Table S1, Fig. S1). The storm surge hindcast generated with ROMS was validated against non-tidal residuals from the same tide-gauge network, yielding RMSE values generally between 6 and 15 cm, low bias, and correlation coefficients above 0.7 across most stations  
185 (Table S2, Fig. S2).

Offshore and nearshore wave hindcasts were validated against buoy observations, showing average biases below  $\pm 0.15$  m, RMSE values typically below 0.35 m, and correlation coefficients exceeding 0.8 for the majority of stations (Figs. S4–S5). The foreshore slope estimation used to parameterize wave setup was independently validated against high-resolution topobathymetric data and field-based observations across representative European coastal settings, confirming its suitability for  
190 large-scale applications (Fig. S6).

Direct validation of TWL as a composite variable remains challenging at continental scale due to the scarcity of observations capturing wave-induced water-level contributions. Consistent with previous large-scale studies, TWL reliability is therefore assessed through the validation of its individual components and by comparison with documented historical extreme events. This approach provides confidence that the reconstructed TWL hindcast is robust and appropriate for the large-scale flood  
195 hazard analyses targeted in this study.

**Table 1: Model validation summary obtained from the different validation analyses performed for the different databases composing TWL, including the foreshore slope.**

Component	Dataset / Model	Observations used	Key validation metrics (typical ranges)	Reference
<b>Astronomical tide</b>	TPXO9v5	48 tide gauges (Europe)	RMSE < 10 cm; Bias $\approx$ 0 cm; $r > 0.9$	Table S1, Fig. S1
<b>Storm surge</b>	ROMS (tide + meteo)	Non-tidal residuals from tide gauges	RMSE 6–15 cm; Bias < $\pm 2$ cm; $r > 0.7$	Table S2, Fig. S2
<b>Offshore waves</b>	WW3-based hindcast	84 wave buoys	Bias < $\pm 0.1$ m; RMSE $\approx$ 0.3 m; $r > 0.9$	Fig. S4
<b>Nearshore waves</b>	Downscaled wave hindcast	Coastal wave buoys	Bias < $\pm 0.15$ m; RMSE 0.2–0.35 m; $r > 0.8$	Fig. S5
<b>Foreshore slope</b>	Sunamura-based estimate	DEMs & field studies	Good agreement across sites; conservative estimates	Fig. S6
<b>Total Water Level</b>	Composite (TWL)	Historical extreme events	Qualitative agreement with documented extremes	Table S4

## 200 2.5 Extreme Value Analysis

The POT method was used to identify extreme events in the TWL hindcast. At each CTP, we selected a threshold that resulted in an average of two events per year, with a minimum interval of 72 hours between events. Next, we estimated the TWL return values (i.e., TWL values associated with return periods) by fitting the extreme sample to an exponential model.

To refine the methodology, we conducted several tests regarding the extreme event selection method, the threshold applied and the distribution fits. Preliminary tests applied were carried out on a set of points selected using clustering methods (Camus et al., 2011), as detailed in the Supplementary Material (Fig. S7). We compared the results of return levels and confidence intervals obtained from AM method using GEV and Gumbel distributions, and from the POT method using both exponential and GPD fits (Supplementary Fig. S9), as well as different thresholds options tested in POT (Supplementary Fig. S10). A stability check of the resulting parameters was assessed through a residual life plot (Coles, 2001; H. Liu et al., 2023) (Supplementary Fig. S11) and concluded that the optimal average amount of events per year ( $\lambda$ ) was two.

At the European scale, we conducted sensitivity analyses on the distribution fits and POT declustering time to ensure the selection of independent events. Regarding the distribution fits, around 13% of the coast exhibited a significant shape parameter, justifying the use of GPD instead of an exponential distribution (see Supplementary Fig. S12). Most of these points presented a negative shape parameter and half of them were located in tide-dominated areas, where the TWL is modulated by the astronomical tide. However, due to the narrow confidence bands in these regions, we opted to retain the exponential fit,

simplifying the method for continental-scale applications. Additionally, only 6.5% of the coast rejected the Anderson-Darling null-hypothesis which would justify the use of GPD (see Supplementary Fig. S13). Regarding the declustering interval, a 72h minimum interval resulted in the largest amount of CTP significantly fitting an exponential fit (93.5%), according to the Anderson-Darling test compared to 48h and 96h intervals (see Supplementary Fig. S8, sensitivity analysis on POT interval).

## 220 **2.6 Relative Contributions of TWL components**

To better understand the dominant conditions at each CTP, the relative contributions of each TWL component were computed (astronomical tide, storm surge, and wave setup). The contribution of each component was calculated by extracting its value from its respective time series at the corresponding time step and determining its proportion relative to the TWL. For mean conditions, we calculated this for each time step within the entire time series and then calculated the mean contribution for each component. For the extreme conditions, relative contributions of each TWL component were calculated for each peak identified by the POT.

## **2.7 Sensitivity Analyses**

Figure 1 also presents the three sensitivity analyses performed. The first two analyses refer to the inclusion of waves in a TWL study, while the third analysis examines the adopted threshold as part of EVA. Within these analyses, the EVA method (POT followed by exponential fit) was adopted in all cases. Two indicators were used to determine how sensitive results are to the different options tested in each analysis. On the one hand, the 100-yr TWL resulting from an exponential fit was analyzed. The variability of results was studied to determine the relative influence each option has on the outcome. On the other hand, the corresponding 100-yr TWL flooded area (FA) resulting from a static flood model applied to the European floodplain was evaluated. The flood model applied was a simple GIS-based bathtub approach, in which any area below a certain water level is considered flooded, provided it is hydraulically connected to the sea. This method was only used as a first-order approximation of the flooded area for comparative purposes in the sensitivity analyses. The floodplain was defined as the terrain below 15 m in elevation and hydraulically connected to the sea, based on the 25 m resolution Copernicus EU-DEM (Copernicus, 2019).

### **2.7.1 Sensitivity analysis: wave contribution**

As part of the validation of the TWL reconstruction, different options were tested in terms of the characterization of the wave contribution. Following the equations by Stockdon et al. (2006) with a variable intertidal slope, (4) (Sunamura, 1984) we tested: (i) static wave setup, following equation (2); (ii) dynamic wave setup, as in equations (5) and (6); and (iii) wave runup, equation (7). After each wave contribution was computed, the corresponding TWL was reconstructed by adding the storm surge and astronomical tide time series. The same resulting TWL variants used in this analysis, were the ones adopted in the TWL reconstruction validation.

$$S_{IG} = 0.06\sqrt{H_s L_0} \quad (5)$$

$$\eta_{dyn} = \eta + S_{IG} \quad (6)$$

$$R_2 = 1.1 \left( \eta + \frac{\sqrt{H_s L_0 (0.563\beta + 0.004)}}{2} \right) \quad (7)$$

### 2.7.2 Sensitivity analysis: wave setup components

Three elements were tested in the wave setup analysis: (1) the wave dataset (downscaled nearshore wave conditions and large-scale wave conditions, hereafter referred to as offshore wave conditions); (2) the foreshore slope approximation (variable and constant); and (3) the wave setup formulation (semi-empirical and standard). The combinations of these elements resulted in six TWL reconstruction approaches. While approach A shown in Table 2 represents the methodology developed in the present study, the remaining approaches were only used in this sensitivity analysis.

First, we calculated wave setup with the same formulation (2) but applying offshore wave conditions instead of nearshore downscaled conditions. Then, we applied the wave setup formulation by Guza & Thornton (1981) to both offshore and nearshore wave conditions (8).

$$\eta = 0.2H_s \quad (8)$$

For the comparison of foreshore slope approximations, we applied the variable slope approximation (Sunamura, 1984) to offshore wave conditions and also adopted a constant foreshore slope of 0.03 (Kirezci et al., 2020). Both foreshore slope options were used in conjunction with the semi-empirical wave setup formulation shown in equation (2) (Stockdon et al., 2006) and with both wave datasets. In total, six approaches were tested, as shown in Table 2. The same EVA method (POT with a threshold of  $\lambda=2$  fitted to an exponential fit) was applied to all approaches to obtain the 100-yr return levels.

**Table 2: Approaches used to reconstruct TWL in the sensitivity analysis of wave setup components, showing their respective combinations of wave datasets, foreshore slope approximations, and wave setup formulations adopted.**

Approach	Dataset	Foreshore slope	Wave setup
A	Nearshore	Sunamura (1984)	Stockdon et al. (2006)
B	Nearshore	Kirezci et al. (2020)	Stockdon et al. (2006)
C	Nearshore		Guza and Thornton (1981)
D	Offshore	Sunamura (1984)	Stockdon et al. (2006)
E	Offshore	Kirezci et al. (2020)	Stockdon et al. (2006)
F	Offshore		Guza and Thornton (1981)

### 2.7.3 Sensitivity analysis: POT threshold

265 Since the POT method is sensitive to the chosen threshold, which can introduce subjectivity and affect the results, a sensitivity analysis is needed to assess how robust and consistent the estimates are. Although the threshold selected in this study corresponds to a variable value corresponding to an average of two extremes events samples per year ( $\lambda=2$ ) in each CTP, other options were tested. The variability in the extreme TWL results was also analyzed for occurrence rates of  $\lambda=1$  and  $\lambda=3$ . The exponential distribution was fitted to the three resulting thresholds to obtain the 100-yr return levels. Outcomes were examined  
270 through the variability of results, Anderson-Darling tests, and the uncertainty represented by the 95<sup>th</sup> confidence intervals of 100-yr TWL.

## 3 Results

### 3.1 Wave and sea level datasets processing and total water level computation

Figure 2 presents the 2021 hindcast of all variables used in this study for four coastal target points (CTP) in which we  
275 reconstructed the TWL. The CTPs used as examples throughout this study were selected from a list of test points selected with K-means (see Supplementary Fig. S7). Point 1 is located on the Atlantic coast of France, point 2 in Estonia in the Baltic Sea, point 3 in Greece in the Mediterranean Sea, and point 4 in Norway (Fig. 2g). These points represent the centroids of the four most distinct clusters obtained from a K-means clustering applied to the relative contributions of the TWL components. Offshore wave conditions (Fig. 2a) were downscaled to the coast (hereinafter, nearshore wave conditions; see Fig. 2b), from  
280 which wave setup (Fig. 2c) was calculated to represent the wave component. Storm surge (Fig. 2d) and astronomical tide data from TPXO (Fig. 2e) were then added to the wave setup hindcast to obtain the TWL (Fig. 2f) time series. As a consequence, the variability of the TWL reflects the heterogeneity of the spatial distribution of its three components. The distinction between offshore and nearshore wave conditions is evident, particularly at points 3 and 4 where notable differences are observed. The results highlight that while the TWL time series of points 1 and 4 are primarily influenced by tides, point 2 appears to be  
285 dominated by storm surge. However, although point 4 shows smaller oscillations, higher wave setup values raise its TWL time series, unlike point 1, where the series remains centered around 0 m.

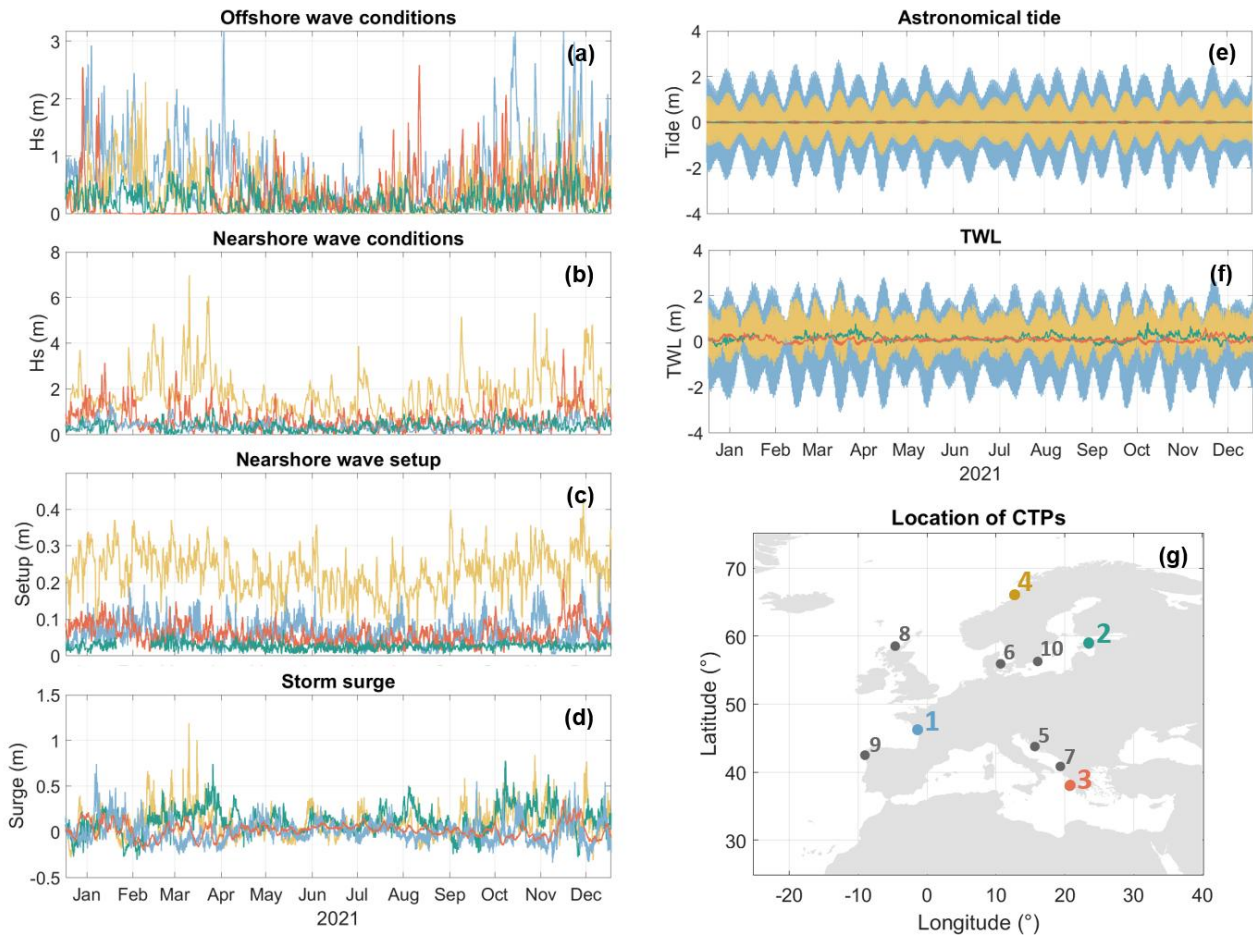


Figure 2: Hindcast time series referred to 2021 for the following variables: offshore significant wave height (a), nearshore significant wave height (b), nearshore wave setup (c), storm surge (d), astronomical tide (e), and TWL (f). Location of the CTPs used as examples (g). Time series are represented in the same colors as the corresponding CTP. Gray CTPs represent the remaining six test points selected with K-means.

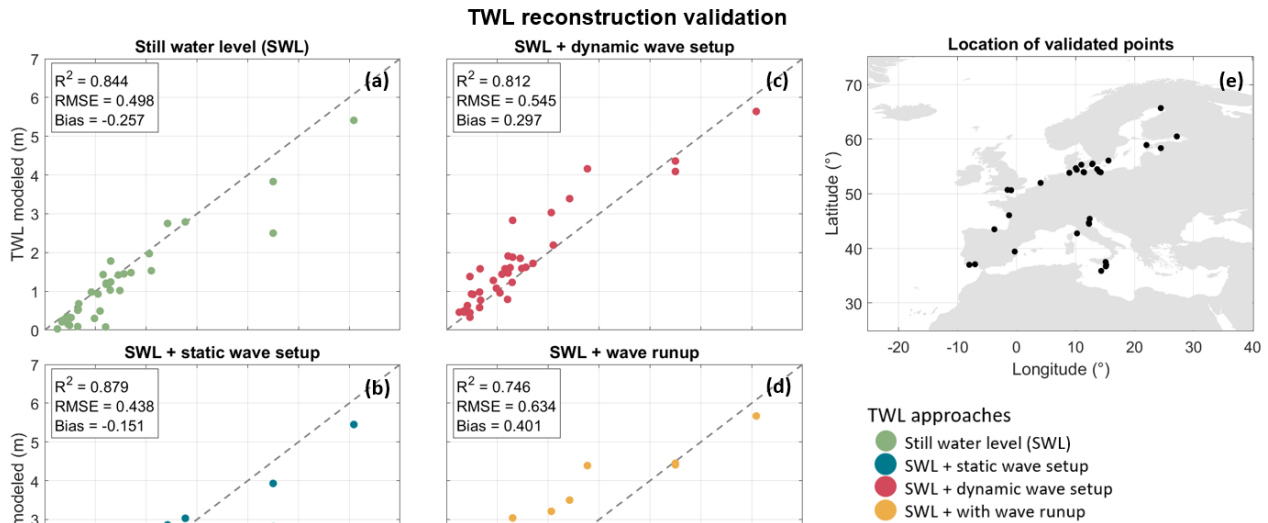
290

### 3.2 Validation of TWL reconstruction

Figure 3 presents the validation of the TWL hindcast. The validation was performed by comparing our estimated values against previously observed historical storms across the study area, based on previous studies. For more details on the references used in the validation of the TWL, please refer to the Supplementary Material (Table S4). In total, four TWL approaches were tested: still water level (SWL; astronomical tide and storm surge) (Fig. 3a), SWL combined with static wave setup (Fig. 3b), SWL with dynamic wave setup (Fig. 3c), and SWL with wave runup (Fig. 3d). Within the options tested for a standardized TWL reconstruction approach at the large-scale, the TWL based on SWL combined with static wave setup shows the best performance with 87.9% of the TWL variability being captured, an RMSE of 0.438 m indicating an average error of 7.5%,

300

and a bias of -0.151 m. Meanwhile, TWL based solely on still water level leads to a greater level of underestimation and the remaining two approaches, an even greater level of overestimation. These results highlight the importance that other aspects might have on the resulting TWL and, consequently, coastal flooding. Notably, the cases in which static wave setup-based TWL underestimates TWL the most are located along the Atlantic coast, such as Santander (Spain), Brouage (France), and the Elbe Estuary (Germany). This could be an indication that these regions are sensitive to infragravity waves, provided that dynamic wave setup offers a more accurate estimate of TWL.

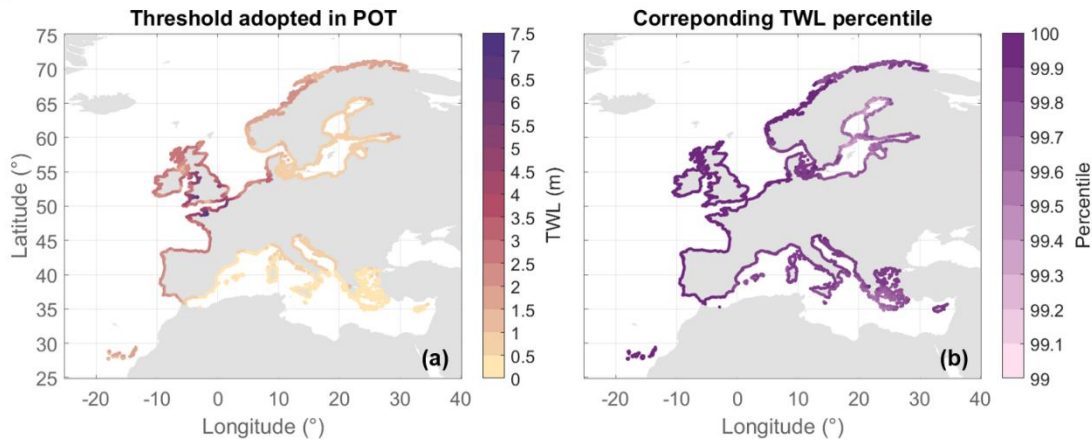


**Figure 3: TWL reconstruction validation based on TWL with still water level (a), static wave setup (b), dynamic wave setup (c), and wave runup (d) when comparing historical coastal storms observed in different points across the study area (e), according to previous studies.**

### 3.3 Extreme value analysis of the total water level

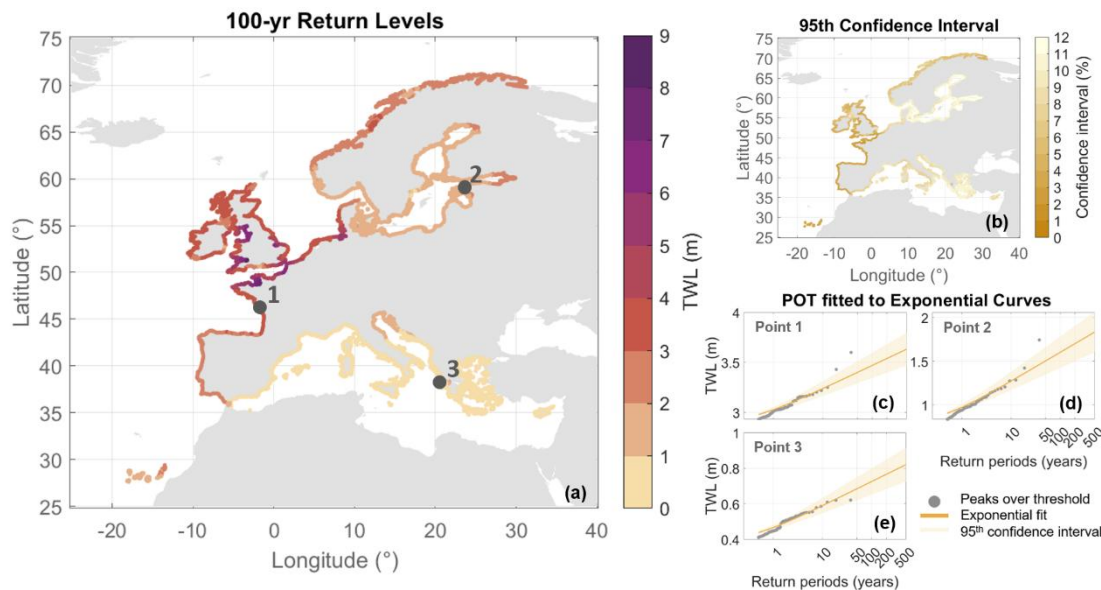
Figure 4 shows the TWL threshold used in the POT method to select the extreme events (Fig. 4a) as well as the corresponding TWL percentiles associated with these thresholds for each CTP (Fig. 4b–c). The results emphasize the importance of considering the spatial variability of thresholds as neither the TWL thresholds nor the corresponding percentiles are uniform across the study area. When adopting  $\lambda=2$ , the TWL threshold ranges from 0.3 m in the Messina Strait (Italy) to 7.2 m in the Bristol Channel (Great Britain), while corresponding percentiles range from P99.2 south of Stockholm (Sweden) to P99.9 in the Wash (Great Britain). The entire study area is characterized by percentiles above the P99, with 97.3% of the CTPs displaying a percentile above P99.5. When applying a constant percentile threshold across the study area,  $\lambda$  varies from 12.1

320 to 68.3 for P90 and 1.4 to 14.5 for P99.5 (see Supplementary Fig. S14). This suggests that using a constant threshold would lead to an excessive number of events in regions such as the Dutch coast and the Irish Sea, for example. Whereas in others regions, such as the Baltic Sea, the sample size would be insufficient, resulting in unstable results of EVA and return levels of TWL.



325 **Figure 4: TWL threshold adopted in each CTP for the POT analysis (a). TWL percentile that corresponds to the threshold adopted in each point (b).**

Figure 5 provides the spatial distribution of the 100-yr TWL for the coast of Europe and the associated 95<sup>th</sup> confidence interval values (Fig. 5b) at each CTP. This result is shown in percentage relative to the return level itself to facilitate the comparison of the results between different regions. For example, the 100-yr TWL in France is  $3.5 \pm 0.12$  m, or 0.3% (Fig. 5c). While in Estonia it is  $1.6 \pm 0.17$  m, or 10.8% (Fig. 5d) and in Greece, it is  $0.7 \pm 0.07$  m, or 9.8% (Fig. 5e). Higher confidence interval values indicate a broader confidence band and, consequently, greater uncertainty in the distribution fit. With TWL values reaching up to 8.5 m in the Bristol Channel, the highest 100-yr TWL values are observed around the British Isles and along the English Channel. This region also shows the narrowest confidence intervals, with the confidence band fluctuating by less than 4% for the 100-yr TWL. This may be linked to the dominant tidal regime in this area, which due to its high amplitude modulates both mean and extreme conditions. The Mediterranean Sea shows the lowest 100-yr TWLs, with values ranging from 0.49 m in the Messina Strait to 2 m in the Adriatic Sea. In this region, the 95<sup>th</sup> confidence interval increases from approximately 6.5% in the Gibraltar Strait to 11% in Greece. In the Baltic Sea, 100-yr TWL range from 0.91 m south of Stockholm to 2.52 m in St. Petersburg, increasing towards the regions of the inner gulfs. Additionally, this region displays the widest confidence intervals, ranging from 9% to 12% for the 100-yr TWL. Overall, these results suggest that return level estimates are more reliable in areas where TWL is dominated by the astronomical tide, such as the Bay of Biscay, where confidence intervals range from 2% to 5%. Similar behavior, albeit with slightly lower values, is observed for the 50-yr TWL (see Supplementary Fig. S15).



345 **Figure 5: Spatial distribution of the 100-yr return period TWL resulting from POT with an exponential fit (a). Spatial distribution of the 95th confidence interval relative to the 100-yr TWL event, in percentage (b). Examples of individual CTP with the exponential fits applied to the POT samples to extract return levels of TWL (c, d, e).**

### 3.4 Relative contributions of mean and extreme total water levels

350 Figure 6 displays the average storm conditions based on the relative contribution of TWL components at peak moment. Examples are shown for three points: point 1 on the Atlantic coast of France (Fig. 6a), point 2 in Estonia in the Baltic Sea (Fig. 6b), and point 3 in Greece in the Mediterranean Sea (Fig. 6c). The shape of the average storm at each point reflects the dominant TWL component during storms. We observe distinct behaviors: point 1 is tide-dominated, point 2 is storm surge-dominated, and point 3 reflects mixed-storm conditions. While point 1 shows semi-diurnal tidal oscillations, point 2 exhibits a single TWL peak, characteristic of a typical storm surge event. Point 3 shows a similar pattern to point 2, but also includes tidal noise throughout the average event. The spatial distribution of TWL component contributions is presented in Fig. 6d–f. Although diverse, it is possible to observe patterns across the study area and the larger European regions previously identified in the study: Atlantic coast, Baltic Sea, and Mediterranean Sea. Figure 6 also presents the dispersion of relative contributions of TWL components, aggregated per European basin. These help with the graphical representation of the uncertainty surrounding the

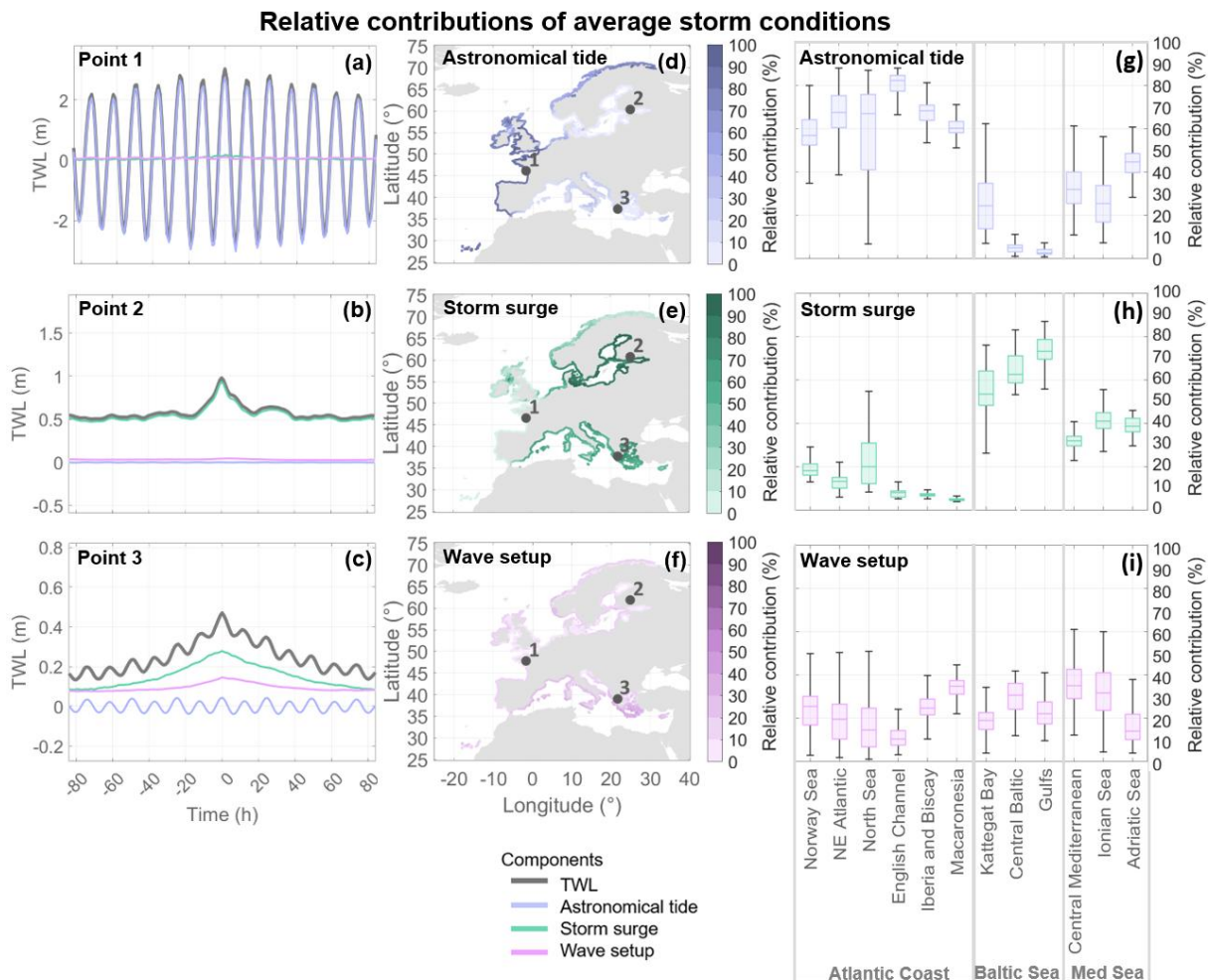
355

360 different patterns observed and the characterization of the three macro-regions (Fig. 6g–i).

First, the dominance of the astronomical tide along the Atlantic coast is evident and modulates storms regardless of their origin, although at different levels depending on the basin. Tidal contributions along the most exposed Atlantic coast can reach 100%, especially along the Atlantic Iberia and Bay of Biscay. The more northern basins tend to present less dominance of the astronomical tide and more presence of storm surge and wave setup. Notably, the North Sea is the most balanced basin in this

365 region in terms of the contributions of the three components. This is likely due to its shallow waters and extensive continental platform which enhance the action of storm surge. Second, the highest storm surge contributions found in the Baltic Sea also reach 100%. Storm surge tends to be higher in the Baltic Sea due to its wider continental platform. The transition role of the Kattegat Bay is reflected in the higher contribution of astronomical tide as opposed to the remaining two basins. Third, although the highest wave setup contributions are in the Mediterranean Sea, this region also experiences storm surge contributions above 50%. Within this region, the Adriatic Sea shows the highest contribution of storm surge likely because of its shallow water enabling the propagation of storm surges as well as its lower level of exposure to incoming waves. Meanwhile, the Ionian Sea presents the highest contributions of wave setup, likely because of its exposure levels and lower astronomical tide when compared to the Central Mediterranean basin.

370

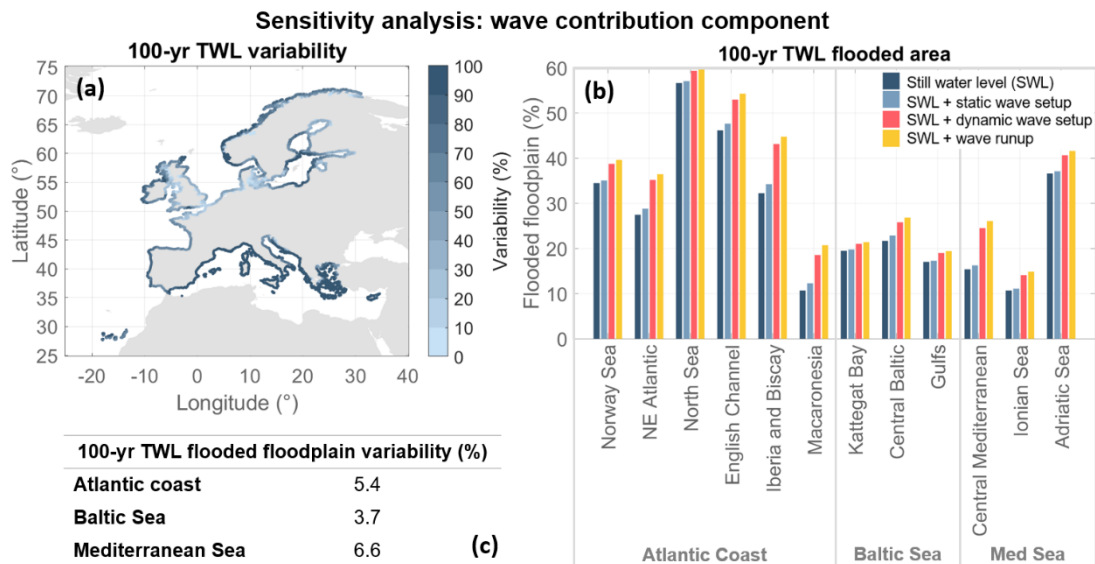


375 **Figure 6: Examples of individual CTP with average storm conditions per TWL component (a, b, c). Spatial distribution of mean relative contributions of astronomical tide (d), storm surge (e), and wave setup (f) under extreme conditions. Dispersion of relative contributions of TWL components per European basin, grouped per European region (g, h, i).**

### 3.5 Sensitivity analysis: wave contribution component

380 Figure 7a presents the variability of the 100-yr TWL for different options of the wave contribution component: still water level (SWL), SWL combined with static wave setup, SWL with dynamic wave setup, and SWL with wave runup. The highest levels of variability are observed in the Mediterranean Sea, more specifically in Central Mediterranean and Ionian Sea, which could be an indication of relatively larger contributions of waves to TWL compared to the remaining two components, or a response to low ranges of TWL, indicating that even a slight change in the wave contribution is reflected in the estimation of TWL extremes. Meanwhile, the lowest variabilities are found in the Baltic Sea, probably a response to its limited exposure to incoming waves.

390 Figure 7b presents the FA per European basin, relative to their respective floodplain area, under the 100-yr TWL when adopting different wave contribution components. Overall, the highest increase in the FA occurs when changing from a static wave setup based-TWL to a dynamic wave setup based-approach. The basins most affected include Central Mediterranean and Iberia and Biscay. The former is located in a region typically known for coastal wave storms (Lobeto et al., 2024), while the latter appears to be a region sensitive to infragravity waves, which had been observed in the TWL reconstruction validation as well. Meanwhile, the least affected basins are Kattegat Bay and the Gulfs, both located in the Baltic Sea. Besides presenting low values of TWL, these basins are also amongst the steepest floodplains. However, when looking at the macro EU regions, the most affected one is the Mediterranean Sea, followed by the Atlantic coast and the Baltic Sea (Figure 7c). These results highlight a clear spatial variability of the wave contribution to TWL. Finally, the European FA with SWL is 36.09%, when including static wave setup it increases to 36.82%, dynamic wave setup to 40.40%, and wave runup to 41.23%. These results show that while the large-scale results do not change dramatically, it is important to zoom in to smaller regions and basins to identify areas in which such decisions might have the greatest impact.



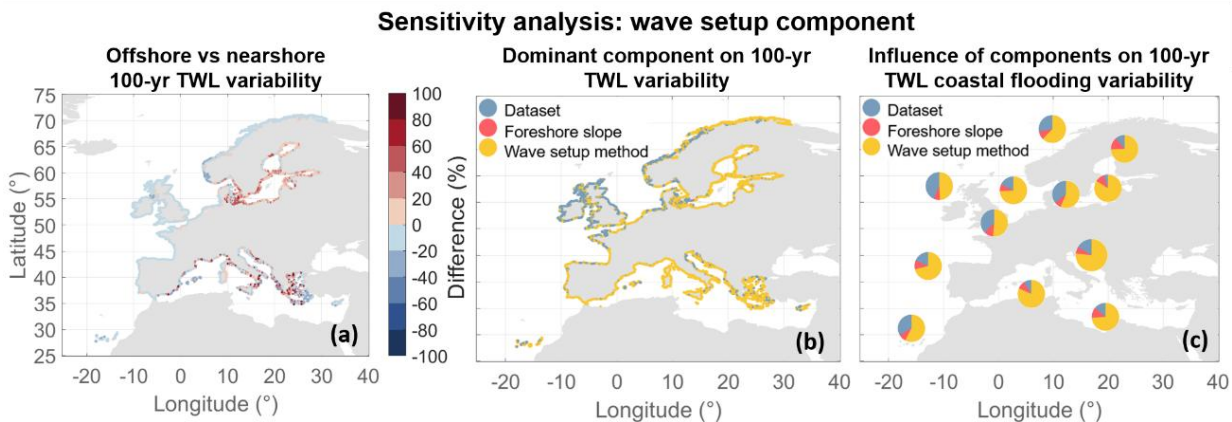
400 **Figure 7: Sensitivity analysis of the wave contribution considered in the TWL reconstruction: still water level (SWL) vs SWL combined with static wave setup vs SWL with dynamic wave setup vs SWL with wave runoff. (a) 100-yr TWL variability found in the results. (b) Results of maximum flooded area proportional to the floodplain in each European basin considering the 100-yr TWL. (c) Variability of 100-yr TWL flooded area relative to floodplain area per European region.**

405 **3.6 Sensitivity analysis: wave setup components**

Figure 8 presents the sensitivity analysis of the different TWL approaches tested, using the 100-yr TWL as the indicator. Figure 8a compares this indicator under offshore versus nearshore wave conditions. Positive values indicate that offshore TWL is higher than the nearshore TWL, and negative values indicate the opposite. The results highlight the spatial variability of the TWL outcomes. Generally, offshore data results in lower TWL along the Atlantic coast and higher TWL in the Baltic and Mediterranean Seas compared to nearshore data. This suggests a greater degree of wave transformation on the Atlantic coast than in the semi-enclosed seas.

Figure 8b identifies the most influential wave component when comparing different TWL modeling approaches. Using the 100-yr TWL as an indicator, 29% of CTPs are most sensitive to dataset selection, while 0.1% are most influenced by the foreshore slope approximation and 70.1% are most influenced by the wave setup method. Dataset selection is the dominant factor along the Atlantic coast. In contrast, the wave setup formulation becomes more important in the semi-enclosed Baltic and Mediterranean Seas. On the other hand, the marginal seas, such as the Baltic or the Mediterranean, exhibit lower TWLs making the accuracy of wave setup estimation crucial. Even the smallest errors in this component can lead to significant discrepancies in the TWL predictions, potentially affecting the accuracy of coastal impacts assessments. Additionally, the low influence the foreshore slope approximation may be due to the conservative approach adopted for capping and normalizing the foreshore slopes (Melet et al., 2020).

Figure 8c presents the influence each wave contribution component tested has on the resulting 100-yr TWL FA, averaged per European basin. Across the entire study area, the most influential element is the wave setup formulation with values ranging from 49.3% along the NE Atlantic basin to 84.0% in the Central Baltic Sea. When comparing the dataset and the slope approximation elements, the Baltic Sea presents more influence from the slope approach, with the exception of the Kattegat Bay which could be described as a transition region rather than being part of the Baltic Sea. Moreover, Central Mediterranean also presents higher influence of the slope approximation than the dataset. The difference, however, is only 0.2%. The remaining basins across the study area, show the slope approach as the least influential element. When comparing the influences of the dataset and the wave setup formulation, most of the study area present a negative correlation. Along the Atlantic coast and the Baltic Sea, as the influence of the dataset increases, the influence of the wave setup formulation decreases.



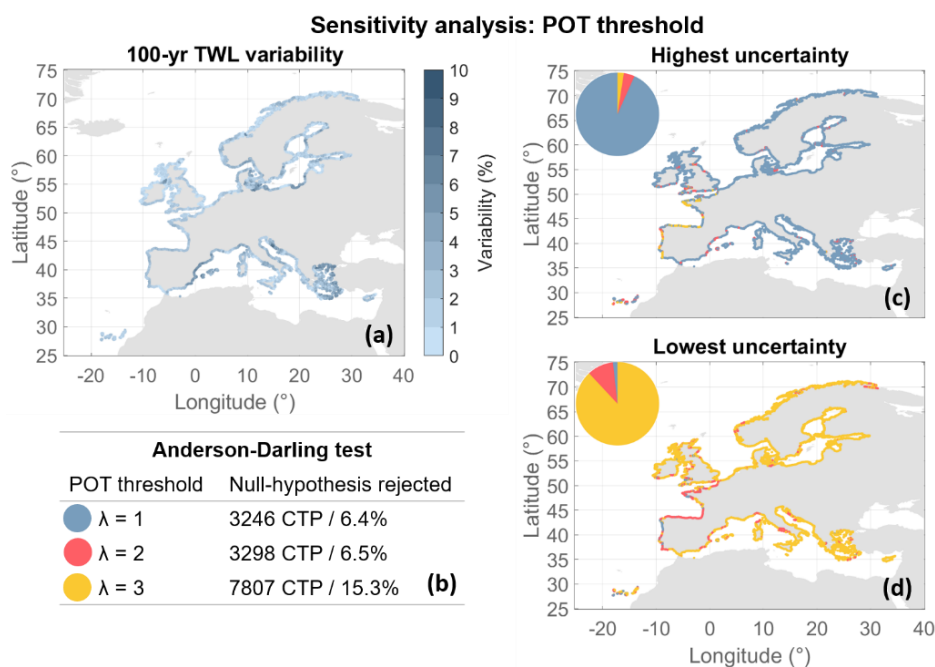
**Figure 8: Sensitivity analysis of the wave setup components: data resolution, foreshore slope approximation, and wave setup formulation. (a) Differences in 100-yr TWL when applying offshore instead of nearshore wave conditions. This plot isolates the influence of wave data resolution by comparing TWL approach D against A. (b) Dominant component explaining the 100-yr TWL variability. (c) Influence of wave setup components on the 100-yr flooded area variability averaged per EU basin.**

### 3.7 Sensitivity analysis: POT threshold

Figure 9 presents the sensitivity analysis regarding the POT threshold, comparing EVA results when adopting a variable threshold corresponding to an average of 1 event per year ( $\lambda=1$ ), 2 events ( $\lambda=2$ ), or 3 events ( $\lambda=3$ ). Although the corresponding TWL threshold varies up to 30%, the resulting 100-yr TWL varies up to 20% (Fig. 9a) and the corresponding 95<sup>th</sup> confidence interval up to 4%. The highest TWL threshold variabilities occur in the Baltic Sea, whereas the highest 100-yr TWL variabilities occur in the Mediterranean. Meanwhile, the Atlantic coast is the least sensitive to the POT threshold selection. Figure 9b displays results of an Anderson-Darling test when adopting the different POT thresholds. A rejected null-hypothesis indicates that the sample does not fit an exponential curve. Out of the 51,010 CTPs, only 121 CTPs indicate that none of the options lead to robust samples. The results shown for  $\lambda=1$  and  $\lambda=2$  are similar, with approximately 6.5% of the CTPs not

adjusting well to an exponential fit. Meanwhile,  $\lambda=3$  leads to more than double of the number of CTPs rejecting the null hypothesis (15.3%). Figure 9c–d indicate which threshold tested lead to the highest and lowest uncertainty as represented by the 95<sup>th</sup> confidence interval of the 100-yr TWL EVA results. Although,  $\lambda=1$  indicates an overall good fit of the samples in the Anderson-Darling test, this is the threshold leading to the highest uncertainty of return levels. A possible explanation is that  $\lambda=1$  leads to excessively small sample sizes, leading to instability in the EVA. Meanwhile, although adopting  $\lambda=3$  yields the lowest uncertainty in most CTPs, the confidence interval difference between  $\lambda=2$  and  $\lambda=3$  remains below 1% for the majority of CTPs, indicating minimal impact for 73.3% of the CTPs.

Ultimately, there is no single threshold which solves the different challenges in all 51,010 CTPs. However, given the poorer results from the Anderson-Darling test with  $\lambda=3$ , the higher uncertainties of the estimated return levels provided by  $\lambda=1$ , and the overall low 100-yr TWL variability across the different methods,  $\lambda=2$  represents a reasonable balance between accuracy, robustness, and sample size to be used in EVA.



**Figure 9: Sensitivity analysis of the POT threshold selected. Variable thresholds corresponding to an average of 1 extreme event per year ( $\lambda=1$ ), 2 events ( $\lambda=2$ ), and 3 events ( $\lambda=3$ ) were tested. Spatial variability 100-yr TWL variability (a). Anderson-Darling test results when adopting  $\lambda=1$ ,  $\lambda=2$ , and  $\lambda=3$  (b). A rejected null hypothesis indicates that the sample does not fit the exponential fit. Uncertainty of results indicating which  $\lambda$  leads to the highest uncertainties (c) and the lowest uncertainties (d) of return levels, based on the 95<sup>th</sup> confidence level.**

### 3.8 Uncertainty in extreme values of TWL

Table 3 presents the overall characterization of the study area regarding TWL extremes, aggregated per EU basin, including the overall uncertainties of 100-yr TWL and its corresponding FA. When looking at the entire study area, 100-yr TWL has an

uncertainty of 24.8%. As more detail is added to the analysis, the 100-yr TWL uncertainty decreases to 17.0% along the Atlantic coast and 15.8% in the Baltic Sea, while it increases to 39.9% in the Mediterranean Sea. In most cases, 100-yr TWL uncertainty increases with increasing relative contribution of wave setup. Notably, the European basins with the highest relative contributions of wave setup, Central Mediterranean and Ionian Sea, are also the ones with the highest uncertainties of the 100-yr TWL. Meanwhile, the basins with the lowest relative contributions of wave setup (English Channel and Kattegat Bay) are also the ones with the lowest 100-yr TWL uncertainty.

**Table 3: Characterization of the study area based on TWL extremes and aggregated per EU basin, region, and the entire study area. EVA results are summarized in TWL threshold corresponding to a  $\lambda=2$  and 100-yr TWL (mean  $\pm$  standard deviation), following the proposed methodology. Relative contributions refer to average per basin and considering all peaks identified by POT. Uncertainty quantification considers all sensitivity analysis performed (i.e., wave contribution, wave setup elements, TWL resolution, POT threshold, and POT minimum interval).**

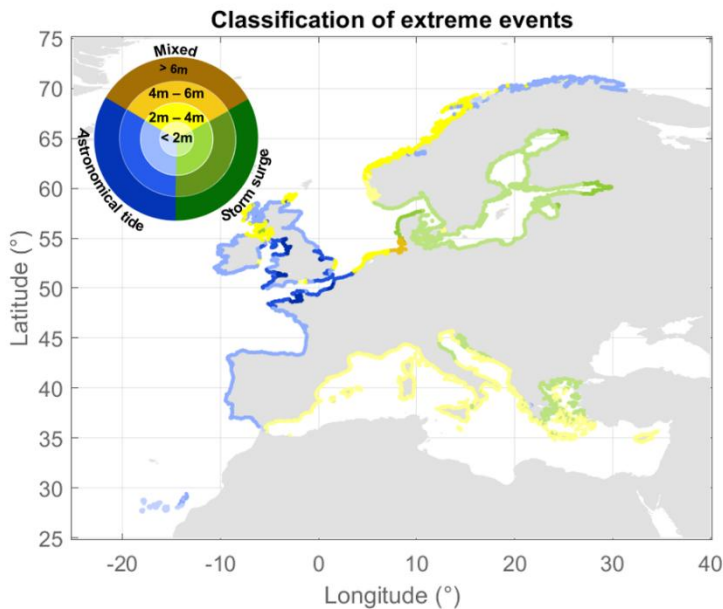
EU basin	EVA		Relative contribution			Uncertainties	
	Threshold (m)	100-yr TWL (m)	Tide (%)	Storm surge (%)	Wave setup (%)	100-yr TWL (%)	100-yr FA (%)
Norway Sea	1.7 $\pm$ 0.3	2.4 $\pm$ 0.4	55.7	31.8	12.5	20.2	3.9
NE Atlantic	3.1 $\pm$ 1.1	3.9 $\pm$ 1.2	71.1	20.8	8.2	15.4	20.1
North Sea	2.5 $\pm$ 1.0	3.5 $\pm$ 1.1	50.1	42.4	7.5	12.1	6.0
English Channel	4.4 $\pm$ 1.5	5.1 $\pm$ 1.6	88.0	7.7	4.3	9.4	4.7
Iberia and Biscay	2.5 $\pm$ 0.5	3.0 $\pm$ 0.6	84.1	4.8	11.1	19.1	4.2
Macaronesia	1.6 $\pm$ 0.1	1.8 $\pm$ 0.2	81.3	1.2	17.5	24.0	10.7
<b>Atlantic coast</b>	<b>2.4 <math>\pm</math> 1.1</b>	<b>3.2 <math>\pm</math> 1.2</b>	<b>65.0</b>	<b>24.8</b>	<b>10.2</b>	<b>17.0</b>	<b>3.2</b>
Kattegat Bay	0.9 $\pm$ 0.1	1.7 $\pm$ 0.2	7.3	87.1	5.6	9.4	10.7
Central Baltic	0.7 $\pm$ 0.1	1.4 $\pm$ 0.3	0.6	88.1	11.3	20.9	3.2
Gulfs	0.8 $\pm$ 0.1	1.7 $\pm$ 0.3	0.3	91.7	8.0	15.6	21.9
<b>Baltic Sea</b>	<b>0.8 <math>\pm</math> 0.1</b>	<b>1.6 <math>\pm</math> 0.3</b>	<b>2.0</b>	<b>89.5</b>	<b>8.5</b>	<b>15.8</b>	<b>5.0</b>
Central Mediterranean	0.5 $\pm$ 0.1	0.7 $\pm$ 0.1	22.6	55.6	21.8	47.7	10.4
Ionian Sea	0.4 $\pm$ 0.1	0.7 $\pm$ 0.1	14.5	64.1	21.4	42.8	4.4
Adriatic Sea	0.6 $\pm$ 0.1	1.1 $\pm$ 0.2	24.4	66.1	9.5	22.2	1.6
<b>Mediterranean Sea</b>	<b>0.5 <math>\pm</math> 0.1</b>	<b>0.8 <math>\pm</math> 0.2</b>	<b>19.5</b>	<b>61.6</b>	<b>18.9</b>	<b>39.9</b>	<b>8.0</b>
<b>EUROPE</b>	<b>1.4 <math>\pm</math> 1.2</b>	<b>2.0 <math>\pm</math> 1.3</b>	<b>35.3</b>	<b>51.8</b>	<b>13.0</b>	<b>24.8</b>	<b>3.8</b>

Regarding the 100-yr FA, the study area presents an uncertainty of 3.8% which decreases to 3.2% along the Atlantic coast and increases to 5.0% and 8.0% in the Baltic and Mediterranean Seas, respectively. However, this does not mean that the European scale map is more reliable but perhaps that it is masking some key responses coastal impacts might have when a different wave contribution is adopted, for example. Overall, uncertainties in 100-yr TWL and FA do not necessarily correlate. The results

show that large uncertainties in 100-yr TWL, when occurring in areas where TWL is typically low, may lead to minimal variability in FA. This is because the variations in 100-yr TWL remain small compared to changes in terrain elevation, limiting their impact on flood extent.

### 3.9 Classification of extreme TWL events

Figure 10 presents the European classification of extreme TWL events, considering the 100-yr TWL and the dominance of TWL components regarding their relative contributions. The results show that there are no regions where the wave setup dominates, although it could be an important source of extreme events, as previously shown. The highest 100-yr TWLs are found in areas where the astronomical tide dominates, mainly along the Atlantic coast. In the Baltic Sea, storm surge dominates all CTPs analyzed and the most sheltered areas are where the 100-yr TWL are the highest. In the Mediterranean Sea, the most exposed regions are where extreme events are mostly mixed, probably indicative of the wave setup contribution which might not be enough to dominate but still relevant in decreasing the influence of the other TWL components.



495 **Figure 10: European classification of extreme events considering the range of 100-yr TWL and the dominance of TWL components regarding their relative contributions.**

## 4 Discussion

The large-scale TWL calculation is complex and presents several challenges. We have presented a methodology for estimating large-scale TWL extremes and applied it to the European coast at the highest resolution to date. For the first time, this was done using nearshore wave conditions, particularly important for semi-enclosed European seas, where offshore data can lead to TWL overestimation. The methodology developed addresses the key challenges introduced regarding the inclusion of the

wave contribution in the TWL, the TWL reconstruction, and the EVA method selected to characterize extreme events. As a consequence, it also introduces limitations inherited from some of the approaches adopted.

Concerning the wave contribution, even though the use of static wave setup as a component of TWL was validated, we  
505 acknowledge that this component may be slightly overestimated in areas such as the Baltic and Mediterranean Seas. Notably,  
this may also influence the relative contributions of the TWL components as well. Several factors influence the accuracy of  
wave contribution estimates. First, we use a semi-empirical setup formulation developed for open-coasts and beaches. It should  
be noted that although both our results and the literature indicate semi-empirical formulations as more appropriate and  
representative of current best practice, they also present limitations (Dodet et al., 2019). For example, the formulation adopted  
510 here was not designed for use in all types of wave regimes or beach profiles (Plant & Stockdon, 2015). Second, we rely on  
modeled foreshore slopes with capping and normalization, which may not fully capture the range and distributions of observed  
beach slopes in different coastal environments. This simplification can lead to the underestimation of the wave contribution in  
steep slopes or overestimation in gentler ones (Melet et al., 2020). Third, the adoption of a constant sediment grain size in the  
application of the Sunamura (1984) foreshore slope approximation is a necessary generalization at this scale and cannot fully  
515 represent the entire study area. This decision might overestimate the resulting foreshore slopes in areas such as wetlands and  
mudflats which have finer sediment grain sizes. However, although beach types and sediment materials are diverse across the  
European continent, the value adopted here follows the global study by Rueda et al. (2017) and agrees with values observed  
in smaller scale studies across the study area (Anthony & Héquette, 2007; Duo et al., 2020; Egon et al., 2025; Horn & Walton,  
2007).

520 With respect to the TWL reconstruction, the linear approach adopted inevitably simplifies some coastal processes, particularly  
in areas such as bays and estuaries, for example. Although hydrodynamic modeling would be the ideal method to solve local  
processes, it is not yet feasible at this scale due to data and computational requirements. Nevertheless, although our TWL  
reconstruction does not explicitly resolve non-linear interactions, these are partly accounted for in the underlying databases.  
For instance, the storm surge database includes non-linear effects with the astronomical tide and the nearshore wave dataset  
525 considers sea level changes due to tide. Moreover, the TWL validation as a composite variable is inherently challenging,  
particularly at the large scale. While SWL can be validated with tide gauges, wave conditions can be validated with buoys.  
However, the increase in coastal water level as a result of the wave contribution is not captured by wave buoys. An alternative  
is to use instruments such as pressure sensors or ADCPs that capture local observations. Even these, however, are limited by  
the need for corrections, regular maintenance, and their localized nature. The lack of observational data and instruments  
530 allowing for the validation of the TWL has led previous studies to validate TWL indirectly, by validating each component  
individually. This strategy was adopted in China (P. Liu et al., 2025), New Zealand (Dalinghaus et al., 2025), and globally  
(Vousdoukas, Mentaschi, et al., 2018), for example. Therefore, given the successful validation presented for the TPXO,  
ROMS, offshore wave, and nearshore wave databases as well as the foreshore slope, we consider the TWL hindcast to be  
robustly validated for the purposes of this study. Nevertheless, a TWL validation was presented by comparing historical storms  
535 identified by previous studies. Even though most of the observed storms were referenced to tide gauges, which do not capture

540 wave contribution, our analysis showed that the extreme events analyzed were well represented by the TWL hindcast. Given the lack of data to validate TWL, the use of tide gauge information to validate TWL is common practice in the literature (Kirezci et al., 2020; Le Gal et al., 2023; Treu et al., 2024; Wing et al., 2024). Similar to the results found here, when comparing validation results with and without wave contribution, Kirezci et al. (2020) detected more accurate estimates when including wave setup. The authors debate whether this is an underestimation of the SWL or the influence of waves given that these results are even more pronounced under extreme than average conditions. Even though the reasons behind these results remain inconclusive, it is clear that when considering the wave contribution, extreme events are better captured by TWL reconstructions.

545 Regarding the EVA methodology, its adoption in tide-dominated regions require careful evaluation. Tides are deterministic, periodic, and autocorrelated, which contradicts the EVA assumption of a sample of independent, abnormally extreme events. However, by applying EVA to the TWL composed of both deterministic and stochastic variables, random exceedances emerge due to the interaction of residuals with tides, introducing variability in the extremes. Therefore, fitting an extreme value distribution to the aggregated TWL, as performed in this study, remains a valid and informative approach and allows the tail behavior to be characterized even in regions dominated by astronomical tide.

550 One key outcome is the strong influence that methodological choices concerning the wave contribution to TWL reconstruction have on the results. The proportion of the European floodplain which is inundated under various TWL approaches varies by up to 10%, depending on wave dataset selection, foreshore slope approach, and wave setup formulation, with the largest differences arising from the choice of wave setup formulation. Using the Guza & Thornton (1981) method yields 4 – 7% more flooded area than the Stockdon et al. (2006) formulation, regardless of the foreshore slope assumption, whereas using a constant foreshore slope yields 1% more flooded area than incorporating variable slopes. Similarly, nearshore TWL produces 3% more flooded area than offshore TWL. The regional differences in the importance of the three components of wave setup can be explained by the geometry and exposure of the different regions. On the one hand, the steeper bathymetry of the Atlantic coast enhances wave shoaling processes, leading to higher waves as they approach the shore. In contrast, the shallower waters and gentler slopes in the Baltic and Mediterranean Seas tend to result in greater wave energy dissipation, as waves encounter the seabed earlier and lose energy more quickly. Although not accounted for in this study, non-linear processes involved in wave transformation, such as refraction and wave-breaking mechanisms, could also help explain such differences as wave transformation processes affect TWL mainly in areas with steep beach slopes and complex offshore bathymetry (Serafin et al., 2019). On the other hand, the Atlantic coast has a larger fetch, allowing for the development of higher waves, and is more exposed to coastal winds, which amplify wave energy. Conversely, the Mediterranean Sea has a smaller fetch, producing more localized and less energetic waves. These factors also result in greater discrepancies in wave conditions within the semi-enclosed seas compared to the Atlantic coast, where offshore and nearshore wave conditions are more similar. While these effects merit further analysis at a higher resolution and on a smaller scale, the results suggest that relying on offshore wave conditions, common in the literature, may lead to an underestimation of the actual flood extent.

565

Besides the adoption of nearshore wave conditions in the estimation of the wave setup, the 1 km TWL hindcast has an  
570 unprecedented spatial resolution for this study area. When working with nearshore (downscaled) wave conditions, higher  
resolution is preferable to fully exploit the quality of the available information. Nearshore wave conditions capture local-scale  
variability, and the higher the resolution, the more faithfully the methodology represents these processes, making the best  
possible use of the data. A sensitivity analysis of the TWL resolution showed that the reduction in spatial resolution had  
minimal effect on the 100-yr TWL, although the flood extent varied up to 14% in some areas (Supplementary Table S3,  
575 sensitivity analysis on TWL resolution). The basins most sensitive to TWL resolution are located in the Atlantic region likely  
because of the increased need for accurate data when modeling a wider range of wave conditions given the highly energetic  
and variable wave climate in this region (Lobeto et al., 2024).

A second outcome is the application of an EVA methodology that is appropriate for large-scale studies. The POT approach  
focuses on high-magnitude events, unlike annual maxima (AM) methods, which assume that a single extreme event occurs per  
580 year. Our findings show that threshold selection greatly influences sample size, particularly when applying a constant  
percentile threshold. Comparing our thresholds to previous studies, we found that 77.1% of the study area exhibits more than  
12 events per year when the threshold of P97 used by Le Gal et al. (2023) is adopted. Meanwhile, if the thresholds of P98 and  
P98.5 used by Kirezci et al. (2020) and Paprotny et al. (2016) are adopted instead, this value decreases to 61.9% and 49.8%,  
respectively. These percentages indicate that large portions of the coast would not be adequately represented using these  
585 thresholds, leading to inconsistent analysis.

A third key insight is the importance of understanding the sources of extreme TWL events. Dominant TWL components  
provide insight into storm behaviors and potential impacts. The spatial patterns we observed have also been reported  
previously. The more exposed, tide-dominated Atlantic coast and the more sheltered and storm surge-dominated Baltic and  
Mediterranean Seas were identified by previous studies, even when neglecting waves (Merrifield et al., 2013). However, our  
590 results differ from others in the literature primarily because of our wave setup characterization. For example, compared to  
Vitousek et al. (2017) our storm surge contributions were higher for the European coast. This is perhaps due to the wave setup  
formulation adopted by the authors, which is specific to dissipative beaches and tends to overestimate wave setup. Additionally,  
their 111 km spatial resolution excluded marginal seas, where we observed the highest storm surge contributions. Meanwhile,  
we identified similar patterns of storm surge contributions in the Baltic Sea and of wave setup dominance in the Mediterranean  
595 Sea, when compared with Melet et al. (2018). Yet, the authors found lower tidal contributions than what we encountered for  
the Atlantic coast, likely due to the inclusion of the swash component, which gives more importance to the wave contributions  
and decreases the tidal influence. In our study, we excluded swash by using static wave setup, as swash operates on a scale of  
seconds to minutes, whereas flood events typically last hours to days (Hinkel et al., 2021; Parker et al., 2023).

Finally, we highlight the importance of considering TWL as a combination of its three components. For example, as we move  
600 towards more extreme return levels, the relative contribution of storm surge increases, exposing the coast to prolonged high  
TWL, which can also heighten wave setup processes (Su et al., 2024). However, we acknowledge that wave setup alone cannot  
drive coastal flooding. On the one hand, wave setup represents an increase in mean sea level of only a few centimeters to a

couple of meters (Idier et al., 2019). On the other hand, the width of the coast affected by this increase in mean sea level is only tens to a couple hundred meters wide (Dodet et al., 2019). The volume of water being propagated towards the coast potentially leading to coastal flooding is not large when compared to tides and storm surges, which increase mean sea level over several kilometers of coastal extent (Woodworth et al., 2019). Additionally, although our study shows that astronomical tide modulates extreme TWL in many regions, it should be pointed out that the main drivers of TWL associated with coastal flooding are the unexpected extreme sea levels due to waves or storm surges as a result of storm conditions. This is because in physical terms, the tide is an expected oscillation to which coastal communities are well adjusted to. However, our analysis shows that without the inclusion of astronomical tide, coastal flooding would probably not occur in many regions of the study area. The results show that even in parts of the Baltic and Mediterranean Seas the tide reaches a contribution of more than 20%, being particularly relevant under average conditions (see Supplementary Fig. S16). Therefore, astronomical tide becomes crucial during the most extreme conditions, even in microtidal areas. Lastly, storm surges, often the primary source of extreme TWL, tend to sustain elevated water levels for extended periods. Therefore, neglecting any one of these components may lead to an underestimation of TWL, particularly if a storm coincides with a spring high tide, thereby increasing coastal risk.

## 5 Conclusions

Currently, there remains a strong need to provide coastal flooding maps at large geographic scales. TWL is the primary input for coastal flood modeling and must be characterized appropriately for the scale of analysis. Without a proper assessment of TWL, estimating extreme conditions becomes more uncertain (Rohmer et al., 2021; Toimil et al., 2020). To address this need, the methodology presented incorporates spatial variability typical of large-scale coastal studies. This was achieved by using spatially and temporally variable foreshore slopes during TWL reconstruction and by applying spatially variable thresholds in the POT method. An analysis of relative contributions helped to interpret extreme TWL behavior along the European coastline. The characterization of three macro-regions (Atlantic coast, Baltic Sea, and Mediterranean Sea) supported the understanding of the degrees of uncertainty observed across different regions at distinct steps of the methodology. Extreme TWL events can cause severe coastal impacts as water overflows inland, reaching communities, infrastructure, assets, and buildings (Vousdoukas, Bouziotas, et al., 2018). Proper reconstruction of TWL and its extreme events is the first step for accurate flood hazard assessment. Large-scale evaluations help us identify hotspots for more detailed risk assessments. Two key limitations of large-scale studies remain, which also point to priorities for future work. First, the computational demands of working with large datasets are high. However, improvements in data availability and computational efficiency have enabled this study to deliver high-resolution, high-quality TWL extreme estimates across Europe. Second, simplifications and assumptions are often required to handle diverse coastal environments. Here, representing TWL as a linear sum of its components overlooks possible additional inputs such as river discharges and non-linear interactions between TWL components, which are important in regions with wide continental shelves and enclosed lagoons (Bertin et al., 2012; Lorenz et al., 2023), such as the Baltic and Mediterranean Seas. According to Arns et al. (2020), not considering non-linear interactions

635 between tide and storm surge, for example, can lead to a 30% increase in estimated extreme water levels, a 16% increase in coastal flooding costs, and an 8% increase in exposed people globally. A strategy to address this issue at the large-scale is to run a hydrodynamic model with both tidal and meteorological forcings combined (Haigh et al., 2014b). This approach, however, does not consider the contribution of waves.

640 Finally, we point out that although this study presents the highest-resolution estimates of extreme TWL for the entire European coastline to date, one could find a considerable variety of beach profiles and types, from sandy shores to rocky formations, within the 1 km distance adopted. Thus, while the TWL resolution adopted here is unprecedented at this scale, it remains insufficient for local-scale applications, where higher-resolution data are needed to support detailed planning.

### **Data availability**

645 The data generated in this study have been deposited in the Zenodo database under accession code 10.5281/zenodo.15111961.

### **Author contributions**

Conceptualization: AT, IJL. Methodology: AT, IJL, MM, HL. Formal analysis: CC. Validation: CC, AT, MM, HL. Writing – original draft: CC. Writing – review and editing: AT, IJL, MM, HL. Project administration: AT, MM. Funding acquisition: AT, IJL.

### **650 Acknowledgements**

This study was funded by the grant COASTALfutures (PID2021-126506OB-100); the Government of Cantabria through the FENIX Project; and the European Union's Horizon 2020 CoCliCo Project (grant agreement No 101003598). CC acknowledges the financial support of the Concepción Arenal Fellowship 2021 of the Universidad de Cantabria (UC-21-19). AT was economically supported by the Spanish Government through the Ramon y Cajal Programme (RYC2021-030873-I).

### **655 References**

- Almar, R., Ranasinghe, R., Bergsma, E. W. J., Diaz, H., Melet, A., Papa, F., Vousdoukas, M., Athanasiou, P., Dada, O., Almeida, L. P., & Kestenare, E. (2021). A global analysis of extreme coastal water levels with implications for potential coastal overtopping. *Nature Communications*, 12(1), 1–9. <https://doi.org/10.1038/s41467-021-24008-9>
- 660 Anthony, E. J., & Héquette, A. (2007). The grain-size characterisation of coastal sand from the Somme estuary to Belgium: Sediment sorting processes and mixing in a tide- and storm-dominated setting. *Sedimentary Geology*, 202(3), 369–382. <https://doi.org/10.1016/j.sedgeo.2007.03.022>
- Arns, A., Wahl, T., Haigh, I. D., Jensen, J., & Pattiaratchi, C. (2013). Estimating extreme water level probabilities: A

- comparison of the direct methods and recommendations for best practise. *Coastal Engineering*, 81, 51–66. <https://doi.org/10.1016/j.coastaleng.2013.07.003>
- 665 Arns, A., Wahl, T., Wolff, C., Vafeidis, A. T., Haigh, I. D., Woodworth, P., Niehüser, S., & Jensen, J. (2020). Non-linear interaction modulates global extreme sea levels, coastal flood exposure, and impacts. *Nature Communications*, 11(1), 1–9. <https://doi.org/10.1038/s41467-020-15752-5>
- Barboza, F. R., & Defeo, O. (2015). Global diversity patterns in sandy beach macrofauna: A biogeographic analysis. *Scientific Reports*, 5, 1–9. <https://doi.org/10.1038/srep14515>
- 670 Barnard, P. L., Erikson, L. H., Foxgrover, A. C., Hart, J. A. F., Limber, P., O’Neill, A. C., van Ormondt, M., Vitousek, S., Wood, N., Hayden, M. K., & Jones, J. M. (2019). Dynamic flood modeling essential to assess the coastal impacts of climate change. *Scientific Reports*, 9(1), 1–13. <https://doi.org/10.1038/s41598-019-40742-z>
- Bertin, X., Bruneau, N., Breilh, J. F., Fortunato, A. B., & Karpytchev, M. (2012). Importance of wave age and resonance in storm surges: The case Xynthia, Bay of Biscay. *Ocean Modelling*, 42, 16–30. <https://doi.org/10.1016/j.ocemod.2011.11.001>
- 675 Bevacqua, E., Vousdoukas, M. I., Zappa, G., Hodges, K., Shepherd, T. G., Maraun, D., Mentaschi, L., & Feyen, L. (2020). More meteorological events that drive compound coastal flooding are projected under climate change. *Communications Earth and Environment*, 1(1). <https://doi.org/10.1038/s43247-020-00044-z>
- 680 Bezak, N., Brilly, M., & Šraj, M. (2014). Comparaison entre les méthodes de dépassement de seuil et du maximum annuel pour les analyses de fréquence des crues. *Hydrological Sciences Journal*, 59(5), 959–977. <https://doi.org/10.1080/02626667.2013.831174>
- Cabrera, P., Montes, J., Duo, E., Brunetta, R., & Ciavola, P. (2024). The Role of Different Total Water Level Definitions in Coastal Flood Modelling on a Low-Elevation Dune System. *Journal of Marine Science and Engineering*, 12(6), 1003. <https://doi.org/10.3390/jmse12061003>
- 685 Calafat, F. M., & Marcos, M. (2020). Probabilistic reanalysis of storm surge extremes in Europe. *Proceedings of the National Academy of Sciences of the United States of America*, 117(4), 1877–1883. <https://doi.org/10.1073/pnas.1913049117>
- Camus, P., Mendez, F. J., Medina, R., & Cofiño, A. S. (2011). Analysis of clustering and selection algorithms for the study of multivariate wave climate. *Coastal Engineering*, 58(6), 453–462. <https://doi.org/10.1016/j.coastaleng.2011.02.003>
- 690 Camus, P., Mendez, F. J., Medina, R., Tomas, A., & Izaguirre, C. (2013). High resolution downscaled ocean waves (DOW) reanalysis in coastal areas. *Coastal Engineering*, 72, 56–68. <https://doi.org/10.1016/j.coastaleng.2012.09.002>
- Coles, S. (2001). *An Introduction to Statistical Modeling of Extreme Values*. Springer series in statistics: London (UK).
- Collings, T. P., Quinn, N. D., Haigh, I. D., Green, J., Probyn, I., Wilkinson, H., Muis, S., Sweet, W. V., & Bates, P. D. (2024). Global application of a regional frequency analysis to extreme sea levels. *Natural Hazards and Earth System Sciences*, 24(7), 2403–2423. <https://doi.org/10.5194/nhess-24-2403-2024>
- 695 Copernicus. (2019). *DEM - Global and European Digital Elevation Model*.
- Dalinghaus, C., Coco, G., & Higuera, P. (2025). Assessing ToTal Water Level Uncertainties Using Global Sensitivity Analysis. *Journal of Geophysical Research: Oceans*, 130, 460–466. <https://doi.org/10.1029/2025JC023011>
- Dodet, G., Melet, A., Ardhuin, F., Bertin, X., Idier, D., & Almar, R. (2019). The Contribution of Wind-Generated Waves to Coastal Sea-Level Changes. In *Surveys in Geophysics* (Vol. 40, Issue 6). Springer Netherlands. <https://doi.org/10.1007/s10712-019-09557-5>
- 700 Duo, E., Sanuy, M., Jiménez, J. A., & Ciavola, P. (2020). How good are symmetric triangular synthetic storms to represent real events for coastal hazard modelling. *Coastal Engineering*, 159(May). <https://doi.org/10.1016/j.coastaleng.2020.103728>
- Egbert, G. D., & Erofeeva, S. Y. (2002). Efficient inverse modeling of barotropic ocean tides. *Journal of Atmospheric and*

- 705 *Oceanic Technology*, 19(2), 183–204. [https://doi.org/10.1175/1520-0426\(2002\)019<0183:EIMOBO>2.0.CO;2](https://doi.org/10.1175/1520-0426(2002)019<0183:EIMOBO>2.0.CO;2)
- Egon, A., Farrell, E., Iglesias, G., & Nash, S. (2025). Measurement and Modelling of Beach Response to Storm Waves: A Case Study of Brandon Bay, Ireland. *Coasts*, 5(3), 1–24. <https://doi.org/10.3390/coasts5030032>
- Gomes da Silva, P., Coco, G., Garnier, R., & Klein, A. H. F. (2020). On the prediction of runup, setup and swash on beaches. *Earth-Science Reviews*, 204(March), 103148. <https://doi.org/10.1016/j.earscirev.2020.103148>
- 710 Guza, R. T., & Thornton, E. B. (1981). Wave set-up on a natural beach. *Journal of Geophysical Research*, 86(C5), 4133–4137. <https://doi.org/10.1029/JC086iC05p04133>
- Haigh, I. D., Wijeratne, E. M. S., MacPherson, L. R., Pattiaratchi, C. B., Mason, M. S., Crompton, R. P., & George, S. (2014). Estimating present day extreme water level exceedance probabilities around the coastline of Australia: Tides, extra-tropical storm surges and mean sea level. *Climate Dynamics*, 42(1–2), 121–138. <https://doi.org/10.1007/s00382-012-1652-1>
- 715 Harley, M. (2017). Coastal Storm Definition. In *Coastal Storms: Processes and Impacts* (Issue April 2017). <https://doi.org/10.1002/9781118937099.ch1>
- 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., ...
- 720 Thépaut, J. N. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146(730), 1999–2049. <https://doi.org/10.1002/qj.3803>
- Hinkel, J., Feyen, L., Hemer, M., Le Cozannet, G., Lincke, D., Marcos, M., Mentaschi, L., Merkens, J. L., de Moel, H., Muis, S., Nicholls, R. J., Vafeidis, A. T., van de Wal, R. S. W., Voudoukas, M. I., Wahl, T., Ward, P. J., & Wolff, C. (2021). Uncertainty and Bias in Global to Regional Scale Assessments of Current and Future Coastal Flood Risk. *Earth's Future*,
- 725 9(7), 1–28. <https://doi.org/10.1029/2020EF001882>
- Horn, D. P., & Walton, S. M. (2007). Spatial and temporal variations of sediment size on a mixed sand and gravel beach. *Sedimentary Geology*, 202(3), 509–528. <https://doi.org/10.1016/j.sedgeo.2007.03.023>
- 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. *Surveys in Geophysics*, 40(6), 1603–1630. <https://doi.org/10.1007/s10712-019-09549-5>
- 730 IHCantabria, Institute of Hydraulics of Cantabria. (2024). *CoCliCo. D3.4: Pan-European storm surge and wave hindcast and projections*. Accessible at: <https://coclicoservices.eu/workpackages/worckpackage-3/#toggle-id-2>
- 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. *Scientific Reports*, 10(1), 1–12. <https://doi.org/10.1038/s41598-020-67736-6>
- 735 Kumar, R., Lemos, G., Semedo, A., & Li, J. G. (2025). A global high-resolution CMIP6 ensemble of wave climate simulations and projections using a coastal multigrid: Configuration and performance evaluation. *Ocean Modelling*, 197(March), 102566. <https://doi.org/10.1016/j.ocemod.2025.102566>
- Le Gal, M., Fernández-Montblanc, T., Duo, E., Montes Perez, J., Cabrita, P., Souto Ceccon, P., Gastal, V., Ciavola, P., & Armaroli, C. (2023). A new European coastal flood database for low-medium intensity events. *Natural Hazards and Earth System Sciences*, 23(11), 3585–3602. <https://doi.org/10.5194/nhess-23-3585-2023>
- 740 Liu, H., Yang, F., & Wang, H. (2023). Research on Threshold Selection Method in Wave Extreme Value Analysis. *Water (Switzerland)*, 15(20). <https://doi.org/10.3390/w15203648>
- Liu, P., Din, A. H., & Hamden, M. H. (2025). Reconstructing hourly coastal total sea levels and assessing current and future extreme sea levels threats to the Coast of China. *Scientific Reports*, 1–16. <https://doi.org/10.1038/s41598-025-22493-2>
- 745 Lobeto, H., Semedo, A., Lemos, G., Dastgheib, A., Menendez, M., Ranasinghe, R., & Bidlot, J.-R. (2024). Global coastal

wave storminess. *Scientific Reports*, 14(1). <https://doi.org/10.1038/s41598-024-51420-0>

- 750 Lorenz, M., Arns, A., & Gräwe, U. (2023). How Sea Level Rise May Hit You Through the Backdoor: Changing Extreme Water Levels in Shallow Coastal Lagoons. *Geophysical Research Letters*, 50(21), 1–11. <https://doi.org/10.1029/2023GL105512>
- MacPherson, L. R., Arns, A., Dangendorf, S., Vafeidis, A. T., & Jensen, J. (2019). A Stochastic Extreme Sea Level Model for the German Baltic Sea Coast. *Journal of Geophysical Research: Oceans*, 124(3), 2054–2071. <https://doi.org/10.1029/2018JC014718>
- 755 Melet, A., Almar, R., Hemer, M., Le Cozannet, G., Meyssignac, B., & Ruggiero, P. (2020). Contribution of Wave Setup to Projected Coastal Sea Level Changes. *Journal of Geophysical Research: Oceans*, 125(8). <https://doi.org/10.1029/2020JC016078>
- Melet, A., Meyssignac, B., Almar, R., & Le Cozannet, G. (2018). Under-estimated wave contribution to coastal sea-level rise. *Nature Climate Change*, 8(3), 234–239. <https://doi.org/10.1038/s41558-018-0088-y>
- 760 Merrifield, M. A., Genz, A. S., Kontoes, C. P., & Marra, J. J. (2013). Annual maximum water levels from tide gauges: Contributing factors and geographic patterns. *Journal of Geophysical Research: Oceans*, 118(5), 2535–2546. <https://doi.org/10.1002/jgrc.20173>
- Muis, S., Verlaan, M., Winsemius, H., Aerts, J., & Ward, P. (2015). The first global-scale hindcast of extreme sea levels. *E-Proceedings of the 36th IAHR World Congress*, 1, 6177–6182.
- 765 Muis, S., Verlaan, M., Winsemius, H. C., Aerts, J. C. J. H., & Ward, P. J. (2016). A global reanalysis of storm surges and extreme sea levels. *Nature Communications*, 7(May). <https://doi.org/10.1038/ncomms11969>
- Neumann, B., Vafeidis, A. T., Zimmermann, J., & Nicholls, R. J. (2015). Future Coastal Population Growth and Exposure to Sea-Level Rise and Coastal Flooding - A Global Assessment. *PLoS ONE*, 10(3), 1–34.
- Paprotny, D., Morales-Nápoles, O., & Nikulin, G. (2016). Extreme sea levels under present and future climate: A pan-European database. *E3S Web of Conferences*, 7. <https://doi.org/10.1051/e3sconf/20160702001>
- 770 Paprotny, D., Morales-Nápoles, O., Vousdoukas, M. I., Jonkman, S. N., & Nikulin, G. (2018). Accuracy of pan-European coastal flood mapping. *Journal of Flood Risk Management*, 12(2), 1–16. <https://doi.org/10.1111/jfr3.12459>
- Parker, K., Erikson, L., Thomas, J., Nederhoff, K., Barnard, P., & Muis, S. (2023). Relative contributions of water-level components to extreme water levels along the US Southeast Atlantic Coast from a regional-scale water-level hindcast. *Natural Hazards*, 117(3), 2219–2248. <https://doi.org/10.1007/s11069-023-05939-6>
- 775 Plant, N. G., & Stockdon, H. F. (2015). How well can wave runup be predicted? Comment on Laudier et al. (2011) and Stockdon et al. (2006). *Coastal Engineering*, 102, 44–48. <https://doi.org/10.1016/j.coastaleng.2015.05.001>
- Pugh, D. (1987). *Tides, surges and mean sea level* (A handbook). J. Wiley and Sons.
- Pugh, D., & Woodworth, P. (2014). *Sea-Level Science: Understanding Tides, Surges, Tsunamis and Mean Sea-Level Changes* (2nd ed.). Cambridge University Press.
- 780 Rashid, M. M., Wahl, T., & Chambers, D. P. (2021). Extreme sea level variability dominates coastal flood risk changes at decadal time scales. *Environmental Research Letters*, 16(2). <https://doi.org/10.1088/1748-9326/abd4aa>
- Rasmussen, D. J., Kulp, S., Kopp, R. E., Oppenheimer, M., & Strauss, B. H. (2022). Popular extreme sea level metrics can better communicate impacts. *Climatic Change*, 170(3–4). <https://doi.org/10.1007/s10584-021-03288-6>
- 785 Rohmer, J., Lincke, D., Hinkel, J., Le Cozannet, G., Lambert, E., & Vafeidis, A. T. (2021). Unravelling the importance of uncertainties in global-scale coastal flood risk assessments under sea level rise. *Water (Switzerland)*, 13(6), 1–18. <https://doi.org/10.3390/w13060774>
- Rueda, A., Vitousek, S., Camus, P., Tomas, 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. *Scientific Reports*, 7, 5038.
- 790 Serafin, K. A., Ruggiero, P., Barnard, P. L., & Stockdon, H. F. (2019). The influence of shelf bathymetry and beach topography on extreme total water levels: Linking large-scale changes of the wave climate to local coastal hazards. *Coastal Engineering*, 150(January), 1–17. <https://doi.org/10.1016/j.coastaleng.2019.03.012>
- 795 Shchepetkin, A. F., & McWilliams, J. C. (2005). The regional oceanic modeling system (ROMS): A split-explicit, free-surface, topography-following-coordinate oceanic model. *Ocean Modelling*, 9(4), 347–404. <https://doi.org/10.1016/j.ocemod.2004.08.002>
- Stockdon, H. F., Holman, R. A., Howd, P. A., & Sallenger, A. H. (2006). *Empirical parameterization of setup, swash, and runup*. 53, 573–588. <https://doi.org/10.1016/j.coastaleng.2005.12.005>
- Su, J., Murawski, J., Nielsen, J. W., & Madsen, K. S. (2024). Coinciding storm surge and wave setup: A regional assessment of sea level rise impact. *Ocean Engineering*, 305(January). <https://doi.org/10.1016/j.oceaneng.2024.117885>
- 800 Sunamura, T. (1984). Quantitative Predictions of Beach-Face Slopes. *Bulletin of the Geological Society of America*, 95(2), 242–245. [https://doi.org/10.1130/0016-7606\(1984\)95<242:qpobs>2.0.co;2](https://doi.org/10.1130/0016-7606(1984)95<242:qpobs>2.0.co;2)
- Toimil, A., Losada, I. J., Nicholls, R. J., Dalrymple, R. A., & Stive, M. J. F. (2020). Addressing the challenges of climate change risks and adaptation in coastal areas: A review. *Coastal Engineering*, 156(July 2019). <https://doi.org/10.1016/j.coastaleng.2019.103611>
- 805 Tolman, H. L. (2009). User manual and system documentation of WAVEWATCH-III/ITM version 3.14. *Technical Note*, 3.14, 220.
- Treu, S., Muis, S., Dangendorf, S., Wahl, T., Oelmann, J., Heinicke, S., Frieler, K., & Mengel, M. (2024). Reconstruction of hourly coastal water levels and counterfactuals without sea level rise for impact attribution. *Earth System Science Data*, 16(2), 1121–1136. <https://doi.org/10.5194/essd-16-1121-2024>
- 810 USACE, U. A. C. of E. (1984). Shore Protection Manual. In *US Government Printing Office* (4th ed., Vol. 1).
- Vafeidis, A. T., Nicholls, R. J., McFadden, L., Tol, R. S. J., Hinkel, J., Spencer, T., Grashoff, P. S., Boot, G., & Klein, R. J. T. (2008). A new global coastal database for impact and vulnerability analysis to sea-level rise. *Journal of Coastal Research*, 24(4), 917–924. <https://doi.org/10.2112/06-0725.1>
- 815 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. *Scientific Reports*, 7(1), 1–9. <https://doi.org/10.1038/s41598-017-01362-7>
- Vousdoukas, M. I., Bouziotas, D., Giardino, A., Bouwer, L. M., Mentaschi, L., Voukouvalas, E., & Feyen, L. (2018). Understanding epistemic uncertainty in large-scale coastal flood risk assessment for present and future climates. *Natural Hazards and Earth System Sciences*, 18(8), 2127–2142. <https://doi.org/10.5194/nhess-18-2127-2018>
- 820 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. *Nature Communications*, 9(1), 1–12. <https://doi.org/10.1038/s41467-018-04692-w>
- Vousdoukas, M. I., Voukouvalas, E., Annunziato, A., Giardino, A., & Feyen, L. (2016). Projections of extreme storm surge levels along Europe. *Climate Dynamics*, 47(9–10), 3171–3190. <https://doi.org/10.1007/s00382-016-3019-5>
- 825 Vousdoukas, M. I., Voukouvalas, E., Mentaschi, L., Dottori, F., Giardino, A., Bouziotas, D., Bianchi, A., Salamon, P., & Feyen, L. (2016). Developments in large-scale coastal flood hazard mapping. *Natural Hazards and Earth System Sciences*, 16(8), 1841–1853. <https://doi.org/10.5194/nhess-16-1841-2016>
- Wing, O. E. J., Bates, P. D., Quinn, N. D., Savage, J. T. S., Uhe, P. F., Cooper, A., Collings, T. P., Addor, N., Lord, N. S., Hatchard, S., Hoch, J. M., Bates, J., Probyn, I., Himsworth, S., Rodríguez González, J., Brine, M. P., Wilkinson, H.,

830 Sampson, C. C., Smith, A. M., ... Haigh, I. D. (2024). A 30 m Global Flood Inundation Model for Any Climate Scenario. *Water Resources Research*, 60(8). <https://doi.org/10.1029/2023WR036460>

Woodworth, P. L., Melet, A., Marcos, M., Ray, R. D., Wöppelmann, G., Sasaki, Y. N., Cirano, M., Hibbert, A., Huthnance, J. M., Monserrat, S., & Merrifield, M. A. (2019). Forcing Factors Affecting Sea Level Changes at the Coast. In *Surveys in Geophysics* (Vol. 40, Issue 6). Springer Netherlands. <https://doi.org/10.1007/s10712-019-09531-1>

835