Reply to the Reviewer

Re: Manuscript ID Preprint egusphere-2025-2998

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

Camila Cotrim, Alexandra Toimil, Iñigo Losada, Melisa Menéndez, and Hector Lobeto

Natural Hazards and Earth System Sciences – NHESS

Response to Reviewer

We would like to thank the Reviewer for their careful revision of our manuscript and for the thoughtful and constructive comments provided. We appreciate the time and effort dedicated to the review process, and we believe the suggestions have significantly improved the quality and clarity of the study. Below, we provide a detailed, point-by-point response to each comment, outlining the corresponding revisions made in the manuscript.

Review Report

The manuscript presents results from developing and analyzing a high resolution hindcast of total water levels for Europe. The authors address the important topic of including wave contributions into TWL calculations and subsequent flood impact/risk analysis. They use different approaches to derive wave setup estimates and quantify the effect of the methodological choice on the resulting flood extent. They also test the effect of using different thresholds in the extreme value analysis that is applied to derive the 100-year TWL levels, which are then used for the static flood mapping. The manuscript is very well written and I commend the authors for the comprehensive analysis. I have few general comments and some minor ones listed below which I believe should be addressed to further improve the quality of the paper.

General comments:

1. The authors emphasize that their new hindcast has an unprecedented spatial resolution of 1km. This is great! However, it is never shown how going from a coarser resolution (for which other hindcasts already exist) to a finer resolution affects the flood hazard assessment, does it make any difference whether I use 1km or 10km or 25km? Personally, I would find this a more interesting sensitivity analysis than, for example, the sampling in the POT approach (which has been done in previous studies for different locations).

A sensitivity analysis on the TWL resolution has been included in the Supplementary Material.

"Sensitivity analysis: TWL resolution

The adoption of large-scale (offshore) wave conditions has been discussed throughout the study, while the effects of the TWL spatial resolution are examined in this sensitivity analysis. To this end, two alternative resolutions have been tested at spatial resolutions of 10 km and 25 km. In contrast to the 51,010 CTPs resulting from the 1 km resolution, these led to 4,039 and 1,324 CTPs, respectively. To check whether this loss in TWL resolution and detail would affect the resulting 100-yr TWL and the corresponding flooded area (FA), a sensitivity analysis was performed. Results from all three cases were compared based on the methodology adopted in this study: nearshore wave conditions, Sunamura foreshore slope, Stockdon static wave setup, POT with lambda 2, 72h minimum interval, and an exponential fit.

Table S3 displays the 100-yr TWL variability per European basins and regions, based on the corresponding average and standard deviation values. For the entire study, the effect of the TWL resolution increases the more detail we add to the analysis. For example, for the entire study area, 100-yr TWL is 2.0 ± 1.3 m and 2.1 ± 1.4 m under 1 and 25 km spatial resolution, respectively. Meanwhile, in the Central Baltic Sea basin the 100-yr TWL changes from 1.4 ± 0.3 m with 1 km resolution to 1.7 ± 0.2 m with 25 km resolution. Overall, the higher the TWL resolution the lower are the return values, which indicate a possible overestimation of extreme TWL when lower resolution TWL is used instead. These results possibly represent a greater level of detail captured by the 1 km spatial resolution data as opposed to the one provided by 10 and 25 km resolutions.

For the European study area, the flooded areas vary by 0.5% and 1% when lowering the TWL resolution to 10 and 25 km, respectively. As with other analyses in this study, these values represent a balance of what occurs across different areas within the European floodplain. For example, in the Kattegat Bay the flooded area increases 0.5% with a lower TWL resolution, whereas in the NE Atlantic basin it decreases 14% with a 25 km resolution compared to the 1 km resolution adopted. In most basins, lower TWL resolution results in smaller flood extent. The most sensitive basins to the TWL resolution are located in the Atlantic region, which also showed to be the most sensitive region to the wave data selection in the sensitivity analysis of the wave setup components."

Table S3 Average ± standard deviation of 100-yr TWL (m) and the corresponding flooded area relative to the area of the corresponding floodplain (%) per European basin, region, and study area. Results are shown for three scenarios of TWL spatial resolution: 1 km, 10 km, and 25 km.

EU basin / region	100-yr TWL (m)			100-yr relative FA (%)		
	1 km	10 km	25 km	1 km	10 km	25 km
Norway Sea	2.4 ± 0.4	2.5 ± 0.4	2.5 ± 0.4	35.0	35.0	35.0
NE Atlantic	3.9 ± 1.2	3.9 ± 1.1	4.0 ± 1.2	28.8	26.9	24.8
North Sea	3.5 ± 1.1	3.7 ± 1.1	3.6 ± 1.1	57.0	56.8	56.9
English Channel	5.1 ± 1.6	5.0 ± 1.7	5.0 ± 1.7	47.6	47.5	47.6
Iberia and Biscay	3.0 ± 0.6	3.0 ± 0.6	3.0 ± 0.6	34.2	33.9	33.5
Macaronesia	1.8 ± 0.2	1.9 ± 0.2	2.0 ± 0.2	12.3	12.0	11.3
Atlantic coast	**3.2 ± 1.2**	**3.3 ± 1.2**	**3.4 ± 1.3**	**49.4**	**49.0**	**48.8**
Kattegat Bay	1.7 ± 0.2	1.8 ± 0.2	1.9 ± 0.2	19.8	19.7	19.9
Central Baltic	1.4 ± 0.3	1.7 ± 0.3	1.7 ± 0.3	22.8	22.7	22.7
Gulfs	1.7 ± 0.3	1.9 ± 0.3	1.9 ± 0.3	17.3	17.3	17.3
Baltic Sea	**1.6 ± 0.3**	**1.8 ± 0.3**	**1.8 ± 0.3**	**20.7**	**20.6**	**20.6**
Central Mediterranean	0.7 ± 0.1	0.8 ± 0.1	0.8 ± 0.1	16.2	16.3	16.2
Ionian Sea	1.1 ± 0.2	1.1 ± 0.2	1.1 ± 0.3	11.1	11.0	11.0
Adriatic Sea	0.7 ± 0.1	0.8 ± 0.1	0.8 ± 0.2	37.1	37.0	37.0
Mediterranean Sea	**0.8 ± 0.2**	**0.9 ± 0.2**	**0.9 ± 0.2**	**24.4**	**24.3**	**24.3**
EUROPE	**2.0 ± 1.3**	**2.1 ± 1.4**	**2.1 ± 1.4**	**36.8**	**36.6**	**36.5**

The TWL resolution was also addressed in the Discussion section. Lines 569 – 577: ”Besides the adoption of nearshore wave conditions in the estimation of the wave setup, the 1 km TWL hindcast has an 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, sensitivity analysis on TWL resolution). The basins more 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).”

2. The authors go into a lot of detail about different EVA methods, specifically the choice of distribution and threshold selection, but pass over other relevant aspects quickly such as the declustering window to derive independent events; they use a constant value following other studies (but no references are provided); that's also an assumption (e.g., https://doi.org/10.1016/j.wace.2024.100701) with potential impact (as shown in the Arns et al., 2023 study).

A sensitivity analysis on the declustering time (minimum interval) adopted in POT has been included in the Supplementary Material:

"Sensitivity analysis: POT interval

The interval used when declustering events to guarantee their independence can be spatially constant or variable, depending on local climatic characteristics. For example, for waves in the Atlantic Ocean, previous evidence suggests that a minimum interval of at least 3 days (or 72h) is required, even though the Poisson assumption is well satisfied with intervals up to 6 days (Méndez et al., 2006). Meanwhile, a storm surge study along the German Bight showed that the influence of the declustering time on the return level outcomes is minimal when adopting POT, compared to AM (Arns et al., 2013). Ultimately, in a POT analysis, the independence criterion should reflect the geophysical origin and duration of the extreme events under study. On the one hand, when defining a spatially constant value, the interval is well-established with 72h in storm surge studies (Caspers et al., 2025; Dullaart et al., 2023; MacPherson et al., 2019; Pupić Vurilj et al., 2025; Vousdoukas, Voukouvalas, Annunziato, et al., 2016) although there are applications ranging from 1.5 days (Arns et al., 2013) to 6 days (Martín et al., 2024). In wave storm studies, a 48h interval has been commonly adopted (Lobeto et al., 2024) as well as a 12 – 24h calm period (Martzikos et al., 2023; Martzikos et al., 2021). On the other hand, spatially variable intervals have not been as explored, especially in large-scale studies. Similar to the adoption of a spatially variable POT threshold, the minimum interval would likely affect the estimation of return values as well as coastal flooding projections. However, the development of a methodology to define such heterogeneous interval values is out of the scope of the present study.

Following the decision of a spatially constant interval, a sensitivity analysis was performed to attest for the robustness of the adoption of the 72h instead of 48h and 96h. Figure S8 presents the EVA results obtained considering the different possibilities of declustering time. The 100-yr TWL variability resulting from the different POT minimum intervals tested is low compared to the other sensitivity analyses performed in this study (Figure S8a). The 95th confidence intervals of the 100-yr TWL indicate that 96h is the most uncertain option (Figure S8c), while 48h presents the lowest uncertainty (Figure S8d). However, the Anderson-Darling test analysis, indicates that the 72h interval leads to the highest proportion of CTPs adequately adjusting to the exponential fit (Figure S8b). Additionally, out of the 3,298 CTPs in which the 72h interval does not result in robust distributions, 2,098 CTPs rejected the null-hypothesis with all three POT minimum intervals tested.

The results indicate that, among the different steps of the methodology tested, the independence time between events has the least influence on the estimated return periods. Yet, this does not imply that the POT interval is irrelevant. Our analysis shows that it is still relevant, though less so than other methodological choices."

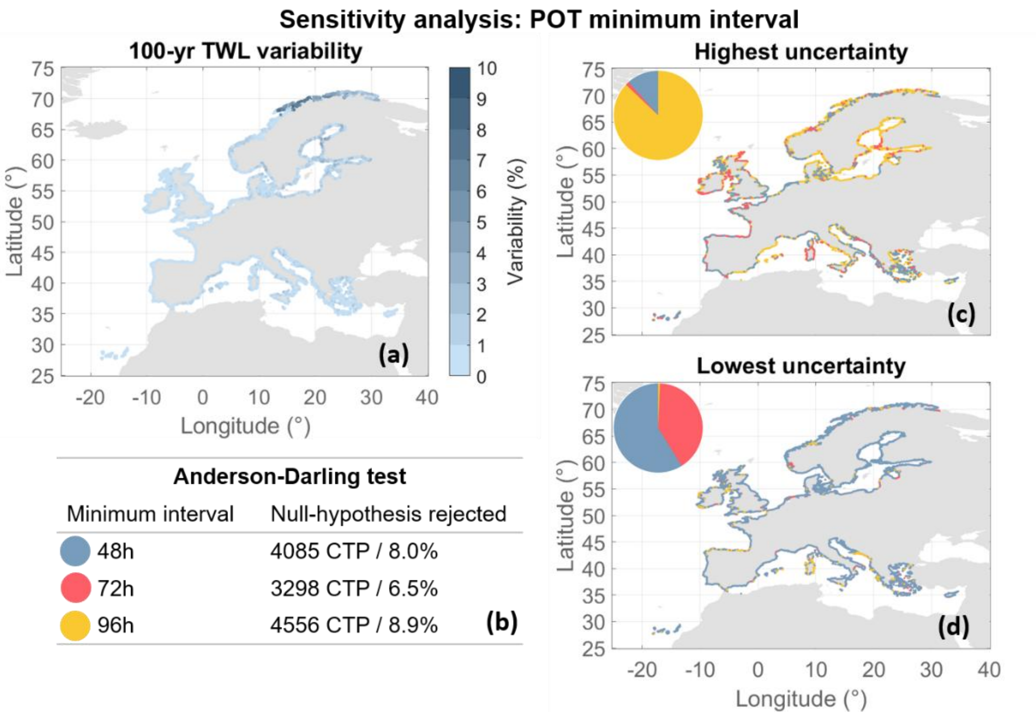

Figure S8 Sensitivity analysis of the POT interval selected (48h vs. 72h vs. 96h). Spatial variability 100-yr TWL variability (a). Anderson-Darling test results when adopting the different intervals (b). A rejected null hypothesis indicates that the sample does not fit the exponential fit. Uncertainty of results indicating which interval leads to the highest uncertainties (c) and the lowest uncertainties (d) of return levels, based on the 95th confidence level.

3. Overall, the paper is very long, especially introduction and discussion (but then there are no conclusions). The intro, for example, includes a lot of textbook style information about extreme value analysis, while the discussion repeats some info that is already touched on in the results section. I don't mind reading a longer paper, but the large amount of (sometimes tangential) information distracts a bit from what I think is the biggest novelty, namely the effect of using different approaches to account for wave contributions. I think with some

additional effort it can become more concise while highlighting the key novelties and related conclusions.

A detailed revision of the text has been performed.

Introduction section. The introduction was revised to replace what the reviewer referred to as "textbook style information" with a discussion of recent advances in extreme value modeling. Lines 71 – 99: "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 98th percentile (P98) of the TWL time series was used globally (Kirezci et al., 2020) while the 97th 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)."

Results section. Some parts from sections 3.4 and 3.6 were transferred and properly adjusted into the Discussion.

Discussion section. Parts of the Discussion have been rewritten to incorporate the text removed from the Results sections mentioned above as well as to make the manuscript more concise, to highlight the improvement in TWL resolution, and to discuss the main limitations of the study. Lines 498 – 615: "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 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 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 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).

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 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 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 (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 identified by previous studies. Even though most of the observed storms were referenced to tide gauges, which do not capture 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.

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.

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.

Besides the adoption of nearshore wave conditions in the estimation of the wave setup, the 1 km TWL hindcast has an 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, 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 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 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 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 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 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."

A new Conclusions section has been added to the manuscript to place more emphasis on the key novelties and limitations of the study. Lines 616 – 642: "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 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.

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."

Specific comments:

31 river discharge effects are not included, something worth mentioning somewhere, maybe when discussion limitations. Additional effects have been included. Lines 631 – 634: "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."

75-105 This part includes a lot of basic information that can be found in text books. At the same time it doesn't touch on more recent developments in extreme value modelling (for example, the approach proposed by Calafat and Marcos: https://www.pnas.org/doi/10.1073/pnas.1913049117, or approaches where distributions are fitted to the stochastic components of TWL after removing deterministic tides). The introduction was revised to replace what the reviewer referred to as "textbook style information" with a discussion of recent advances in extreme value modeling. Lines 71 – 78: "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."

157-159 something is wrong with this sentence. Sentence modified.

166 “global median beach slopes reported locally” sounds very strange and I am not sure what it is supposed to mean. Clarified.

205-210 I got very confused here because the same numbering style is used for the list and the equations; I suggest switching to i, ii, iii or a, b, c for the list and keep (1) (2) style for the equations. Numbering style of the list has been modified to i, ii, and iii.

215 Refer to table 1 when mentioning approach A. I was wondering what it is and only realized on the next page. Table 2 referred to.

235 “two extreme events per year”. Text modified.

242 “reconstructed…hindcast”; I would drop one. Text modified.

277 “Elbe Estuary”. Text modified.

310-315 (and other places) When the tides are the main driver, does it make sense to fit an extreme value distribution because its assumptions are not met? I know it’s a commonly used approach, but that doesn’t make it necessarily right.

The reviewer raises a valid concern and it has been included as a limitation of the study. Lines 544 – 549: “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.”

348 delete “the” Deleted.

Figure 7 (and related text) How does it look when no wave contributions are included (which is still often the approach used for large-scale assessments)

We have included SWL to the sensitivity analysis on the wave contribution component. Lines 380 – 398: “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.

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."

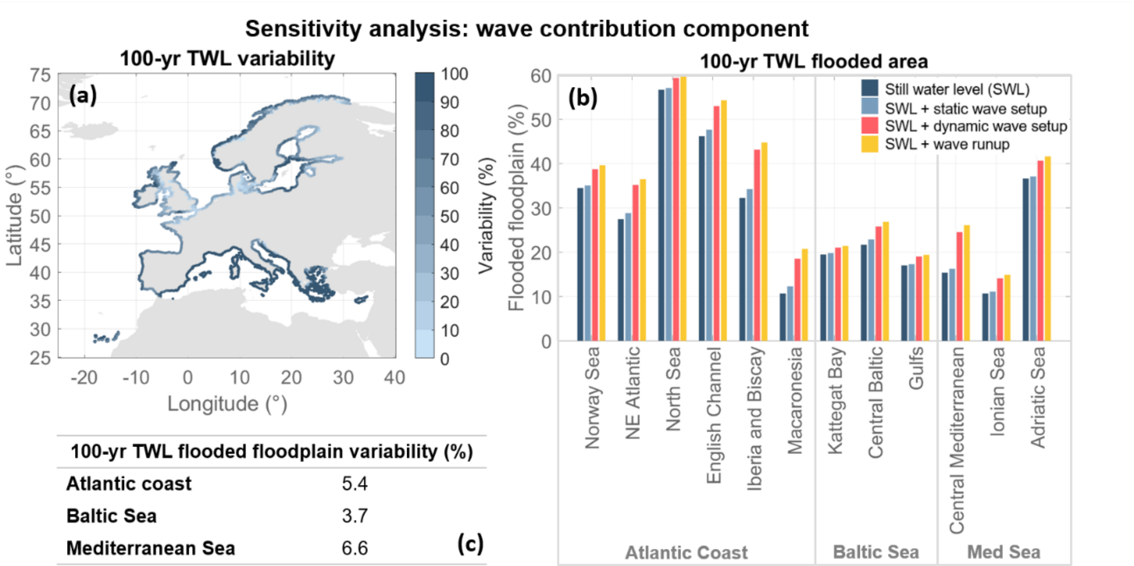

Figure 7 Sensitivity analysis of the wave contribution considered in the TWL reconstruction: still water level (SWL) vs SWL combined with static wave setup vs SWL with dynamic wave setup vs SWL with wave runup. (a) 100-yr TWL variability found in the results. (b) Results of maximum flooded area proportional 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.

392-403 This and other parts of the results section read more like Discussion. Section moved to the discussion.

492 delete "an". Deleted.

506 "applying a specially variable" (or "thresholds"). Text modified.

Supp. Fig. 8 "significancy" should be "significance" in the title. Figure title and caption modified.

References

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

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

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

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

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

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

Caspers, J. J., Kindermann, P. E., Rongen, G. W. F., & Geerse, C. P. M. (2025). Storm surge hydrographs and characteristics along the Dutch coast. An analysis from simulation data. Coastal Engineering, 201(May). https://doi.org/10.1016/j.coastaleng.2025.104776

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

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

Dullaart, J. C. M., Muis, S., De Moel, H., Ward, P. J., Eilander, D., & Aerts, J. C. J. H. (2023). Enabling dynamic modelling of coastal flooding by defining storm tide hydrographs. Natural Hazards and Earth System

Sciences, 23(5), 1847–1862. https://doi.org/10.5194/nhess-23-1847-2023

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

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

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

Harley, M. (2017). Coastal Storm Definition. In Coastal Storms: Processes and Impacts (Issue April 2017). https://doi.org/10.1002/9781118937099.ch1

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., Vousdoukas, 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, 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

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

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

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

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

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

Martín, A., Wahl, T., Enriquez, A. R., & Jane, R. (2024). Storm surge time series de-clustering using correlation analysis. Weather and Climate Extremes, 45(April). https://doi.org/10.1016/j.wace.2024.100701

Martzikos, N., Malliouri, D., & Tsoukala, V. (2023). Shape investigation and probabilistic representation of coastal storms. Applications to Mykonos and Barcelona. Applied Ocean Research, 135(April). https://doi.org/10.1016/j.apor.2023.103563

Martzikos, N. T., Prinos, P. E., Memos, C. D., & Tsoukala, V. K. (2021). Statistical analysis of Mediterranean coastal storms. Oceanologia, 63(1), 133–148. https://doi.org/10.1016/j.oceano.2020.11.001

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

Méndez, F. J., Menéndez, M., Luceño, A., & Losada, I. J. (2006). Estimation of the long-term variability of extreme significant wave height using a time-dependent Peak Over Threshold (POT) model. Journal of Geophysical Research: Oceans, 111(7), 1–13. https://doi.org/10.1029/2005JC003344

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

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

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

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

Pupić Vurilj, M., Antolínez, J. A. Á., Muis, S., & Morales Napoles, O. (2025). Storm surge hydrographs from historical observations of sea level along the Dutch North Sea coast. Natural Hazards, Ipcc 2021. https://doi.org/10.1007/s11069-025-07351-8

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.

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

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

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

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

Treu, S., Muis, S., Dangendorf, S., Wahl, T., Oelsmann, 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

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

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

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., 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