Reply to the Reviewers

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

We would like to thank the Reviewers 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 list of main relevant changes made to the manuscript and detailed, point-by-point response to each comment, outlining the corresponding revisions made in the manuscript

Relevant changes made to the manuscript

Revised structure of the manuscript:

- The results section has been shortened by moving the sensitivity analysis on POT threshold to the Supplementary Material.

Inclusion of new analyses in the Supplementary Material:

- Besides the change mentioned above, a new sensitivity analysis has been included in the Supplementary Material concerning the selection of the marine dynamics databases nodes used in each CTP by comparing a linear interpolation with the selected method of nearest neighbor.

Response to Reviewer 1

Review Report

Starting from the original manuscript, the authors have improved the paper by enhancing the description of the methods, sensitivity tests, and model validation. I really appreciate the effort. However, although the paper aims to provide high-resolution and high-quality TWL estimates across Europe, in my opinion it does not convincingly demonstrate that this goal has been achieved.

- Regarding high-resolution TWL: the added value of the 1 km dataset is not demonstrated, since the results for the 100-year TWL and the corresponding flooded areas are not significantly different from those obtained with the 10 and 25 km resolution datasets (page 25 of the main article and Table S3). Even though the authors stress that the flood extent varied by up to 14% in some areas, there is no evidence that the 1 km results are more accurate than those obtained with the other two resolution datasets.

As highlighted in Line 104, the aim of the study is to provide consistent, Europe-wide, high-resolution TWL boundary conditions for large-scale coastal flood hazard assessments. In that context, the added value of the 1 km product lies in its improved representation of nearshore forcing and local coastal variability, which is expected to be particularly relevant in areas with strong spatial variability in wave conditions and coastal morphology.

We agree with the Reviewer that, when flood extent is used as the evaluation metric, the differences between TWL resolutions are not always large. However, this is not unexpected. The inland flood response is strongly conditioned by the resolution and vertical uncertainty of the DEM used in the inundation modeling. In our case, the 25 m DEM and its non-negligible vertical error can mask the sensitivity to modest differences in coastal boundary forcing, such that differences in flooded area become more apparent mainly where the forcing contrasts are larger. This explains why the 100-year TWL and associated flooded areas may appear similar in some regions despite differences in the coastal boundary conditions. At the same time, we note that the sensitivity analysis does show non-negligible differences in some basins, with flooded area varying by up to 14%, particularly in Atlantic regions. These areas are likely more sensitive because of the larger discrepancies between TWL magnitudes and terrain heights, as well as the greater importance of accurately resolving a wider range of wave conditions.

To address this concern, we have revised the Discussion to better frame the added value and limitations of the 1 km dataset, and to avoid implying that higher TWL resolution necessarily translates into a uniformly detectable improvement in inundation extent across all regions. The revised text now reads:

Lines 554 – 565: "We emphasize that the objective is not to improve flood extent estimates directly, but to provide physically consistent high-resolution boundary conditions required for next-generation coastal flood models. With 1 km resolution data, the physics of the processes are better captured because we better resolve wave-height gradients, coastline orientation effects, and exposure heterogeneity. This is particularly relevant in complex coastlines (e.g., embayed or headland-dominated), where coarse resolution fails to capture gradients that are resolved at 1 km. A sensitivity analysis of the TWL resolution showed that the reduction in spatial resolution affected the flood extent up to 14% in some areas (Supplementary Table S3, sensitivity analysis on TWL resolution). The limited differences in some flooded area comparisons are expected given the DEM resolution and its vertical error, which can mask sensitivity to small changes, revealing differences primarily for larger forcing contrasts. Notably, the basins most sensitive to TWL resolution are located in the Atlantic region likely because of the greater discrepancies between TWL magnitudes and terrain elevations in addition to 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)."

- Regarding high-quality TWL: the validation of both the slope and the TWL reconstruction is qualitative (as mentioned in section 2.4 and Table 1) and does not provide a robust assessment of the quality of the main results of this study. The slope estimated using Sunamura approach has been qualitatively compared against those obtained with a traditional method based on high-resolution data or field observations. However, no details are provided about the wave conditions used in the Sunamura equation, no statistical validation is presented, and no analysis is offered regarding how the uncertainty in the slope propagates to the 100-year TWL and the corresponding flooded areas. Moreover, it is not clear whether the morphological characteristics of the coast have been taken into account in the slope and TWL estimations. The TWL validation is not robust because it is based on comparisons with tide gauges, which do not capture the wave contribution. Indeed, the results of the TWL validation (Fig. 3) show that the SWL and SWL+static wave setup perform better than the SWL+dynamic wave setup and SWL+wave runup cases. The SWL underestimation could be due to the well-known underestimation of the ERA5 atmospheric forcing (as also mentioned in the Supplementary Material).

To conclude, without a robust assessment of the TWL reconstruction, the main findings of this study - including return levels, relative contributions, and the classification of extreme events - cannot be considered reliable.

We agree that direct validation of the TWL at the European scale is inherently limited by the availability of observations. Tide gauges do not capture the wave contribution, and spatially extensive observations of wave setup, runup, or shoreline TWL are not available across Europe in a form that would allow a systematic quantitative validation of the reconstructed TWL field. We agree that this is a key limitation, and we have revised the manuscript to better reflect the level of confidence that can be assigned to the results.

That said, we note that this limitation does not preclude the interpretation of the results within the intended large-scope of the study. At this scale, TWL reconstruction is typically evaluated indirectly through the quantitative validation of its main components. This validation approach is consistent with previous large-scale TWL studies (e.g., Dalinghaus et al., 2025; Liu et al., 2025; Vousdoukas et al., 2018a). Accordingly, the results should be interpreted as large-scale, physically consistent estimates rather than site-specific validated TWL values. In our case, we quantitatively validate the astronomical tide and storm-surge components against tide-gauge observations, and the nearshore wave conditions used to estimate setup and runup against wave observations. The Sunamura-based slope is then derived from this validated wave forcing. We also include comparisons of the estimated slope against local estimates in selected hotspots and assess, through sensitivity analyses, the influence of slope uncertainty on the 100-year TWL and the resulting flooded areas. Sensitivity analyses indicate that uncertainty in slope propagates to TWL but remains secondary compared to surge and wave contributions at continental scale.

We agree, however, that the manuscript did not explain these aspects with sufficient clarity. In particular, we have now clarified:

(i) the wave conditions used in the Sunamura equation (Lines 510 – 511: “the Sunamura-based slope is derived from validated wave forcing”);

(ii) that local morphology and engineering structures cannot be comprehensively resolved at this continental scale (Lines 511 – 512: “it cannot fully resolve local morphological complexity or the influence of coastal structures at this scale”);

(iii) that the slope validation is intended as a consistency check against local estimates rather than a full site-specific calibration (Lines 193 – 194: “The slope validation is intended solely as a consistency check against local estimates rather than a full site-specific calibration.”); and

(iv) that the TWL comparison with tide gauges should be interpreted as an indirect consistency check of storm-event representation, not as a full validation of the wave-resolving TWL signal (Lines 522 – 524: “it is clear that when considering the wave contribution, extreme events are better captured by TWL reconstructions, with the TWL validation serving as a consistency check to assess the representation of TWL during storm events”).

Regarding the point that SWL or SWL plus static setup sometimes compare better with tide gauges than TWL formulations including dynamic wave contributions, we agree that this reflects an important limitation of the validation dataset rather than evidence that the wave contribution degrades the TWL reconstruction. Because tide gauges do not measure wave runup and only incompletely represent setup effects, better agreement with gauge records does not necessarily imply a better reconstruction of shoreline-relevant TWL. In this sense, the comparison shown in Fig. 3 is not intended to demonstrate that tide-gauge observations fully validate the wave-resolving TWL, but rather to assess whether the reconstructed water-level series remain consistent with observed storm-related sea-level variability. Moreover, as the Reviewer notes, part of the SWL underestimation may also be related to the known limitations of ERA5 atmospheric forcing.

To address this concern, we have revised the manuscript to better distinguish between component validation, consistency checks, and remaining uncertainties, and to avoid overstating the degree of direct TWL validation that is possible at this scale.

Lines 498 – 514: “Moreover, direct validation of reconstructed TWL as a composite variable at the European scale is inherently challenging as it is limited by observational constraints. 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 and tide gauges do not capture wave runup and only partially represent setup effects, while spatially extensive observations of shoreline TWL are not available across Europe. An alternative is to use instruments such as pressure sensors or ADCPs that capture local observations. However, even these are limited by the need for corrections, regular maintenance, and their localized nature. The lack of observational data and instruments enabling direct validation of TWL has led previous studies to validate TWL indirectly, by validating each component individually (e.g., Dalinghaus et al., 2025; Liu et al., 2025; Vousdoukas, et al., 2018a). The evaluation presented here also relies primarily on the quantitative validation of the individual TWL components, including tides, storm surge and nearshore wave conditions, together with consistency checks of the reconstructed TWL during storm events. 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. It should be noted that although the Sunamura-based slope is derived from validated wave forcing and compared against local estimates in selected hotspots, it cannot fully resolve local morphological complexity or the influence of coastal structures at this scale. Sensitivity analyses further indicate that, across most of Europe, uncertainty in slope propagates to TWL though it remains secondary compared to surge and wave contributions at continental-scale.”

Response to Reviewer 2

Review Report

The authors have adequately addressed my previous concerns. I appreciate the added sensitivity analysis and validation steps. I have no further objections to move the manuscript forward in its current form.

We thank the reviewer for their positive feedback and for the constructive comments provided throughout the review process. We are pleased that the additional analyses have addressed the concerns and strengthened the manuscript.

Response to Reviewer 3

Review Report

This manuscript presents a high-resolution hindcast of Total Water Level (TWL) for Europe, addressing an important and often overlooked aspect of coastal hazard assessment, the contribution of wave-driven processes to TWL. The authors describe multiple approaches for deriving wave setup and evaluate their influence on flood extent, representing a meaningful advancement over methods that either omit wave contributions or approximate them coarsely.

In addition, they are very through in assessing the impacts of different components and how they are created. e.g. variable POT thresholds to capture, 1, 2 or 3 events per year. They also clearly show the uncertainty around the different components.

The methodology for constructing the individual TWL components appears well conceived, the manuscript is clearly written, and the effort invested is evident throughout. The authors also clearly highlight the limitations in their approach, e.g. variability of coastlines and how their approach could be improved in future works. I recommend acceptance subject to minor revisions, as detailed below.

Comments

1. Linear summation of TWL components

While the individual component datasets appear robustly constructed and are validated in the hindcast context, the linear summation approach used to combine them warrants further discussion. The authors acknowledge this but then do not really consider any other approaches for constructing TWL in such contexts, including joint probability analysis of tide and surge, or the use of derived metrics such as skew surge, which better captures the statistical dependence between tidal phase and surge magnitude. A reasoning why they did not use these approaches would be beneficial.

The limitations of constructing TWL through linear summation are now discussed in greater detail. Lines 624 – 633: “At large scale, several approaches have been proposed to better represent the nonlinear interactions and statistical dependence between tidal and storm-surge components. For instance, this has been addressed through hydrodynamic modeling (Haigh et al., 2014b) and joint probability statistical techniques (Palmer et al., 2024). Additionally, previous studies propose the use of skew surge, which allows for a more reliable representation of the dependence between tidal phase and storm surge magnitude (Marcos and Woodworth, 2017; Williams et al., 2016). Although physically and statistically more robust, these approaches generally require more site-specific information and are computationally less tractable and less transferable across regions. For this reason, linear summation was retained here as a spatially consistent and computationally efficient approach. Importantly, the focus of this study is consistency across Europe rather than site-specific accuracy. Further work is nevertheless needed to incorporate wave contributions within nonlinear TWL frameworks at this scale."

2. Length of the results section

The results section is lengthy and could benefit from condensation. The authors may wish to consider whether even more material could be moved to supplementary information without diminishing the paper.

We have moved the sensitivity analysis on POT threshold to the supplementary material. The description within the methodology was also moved. Figure 1 has been updated to agree with the analyses being presented in the manuscript.

3. Downscaling/Interpolation of components

The authors state that the nearest grid point was assigned for each CTP for some variables and for others hourly time series were "obtained". A clearer explanation of how the datasets were downscaled/interpolated or why they used nearest neighbour is needed.

Clarified. Lines 148 – 152: "At each CTP, hourly time series were obtained for the astronomical tide, storm surge, offshore wave conditions, and nearshore wave conditions. While nearshore wave conditions were downscaled directly at each CTP, the offshore wave hindcast, storm surge hindcast, and TPXO databases do not share the same spatial resolution. To address this mismatch, data from the nearest grid point was assigned to each CTP for those variables. A sensitivity analysis confirmed the robustness of the node selection method from the marine dynamics databases (see Supplementary Figure S6)."

"Sensitivity analysis: selection of nodes within marine dynamics databases

A sensitivity analysis was performed to assess how the data selection with nearest grid point as opposed to interpolation influence the P95 results of astronomical tide, storm surge, and TWL in a selection of CTPs. The nearest grid point was selected based on the closest node of each database to the CTP at hand.

The interpolation method was based on a linear interpolation with nearest-neighbor extrapolation. Figure S6 presents the results for three locations, one in each of the three European regions identified in the study. The largest differences in astronomical tide and TWL are found in the Baltic Sea, while the largest differences in storm surge are found along the Atlantic coast. However, the statistical metrics analyzed indicate that such differences are too small to play a relevant role in the estimation of extreme values as R^2 range from 0.88 to 1, RMSE from 0 to 0.002 m, and bias from -0.001 to 0.001. Therefore, the different results obtained with both approaches do not justify the use of interpolation methods in this case, particularly at this scale, when there are 51,010 points in which the method is required.

Therefore, when selecting the astronomical tide and storm surge data, the nearest grid point approach was adopted instead of interpolation for its computational efficiency after ensuring that such a decision has no influence on the results. We highlight that the resolution of the different databases is higher than the resolution at which such variables usually vary. The astronomical tide TPXO spatial resolution is around 3.5 km and the storm surge spatial resolution is 10 – 15 km. Meanwhile, astronomical tide and storm surge processes present a spatial variability of approximately hundreds to thousands of kilometers (Woodworth et al., 2019). Therefore, the resolution of the databases used provide a sufficiently dense sampling to capture their regional gradients.”

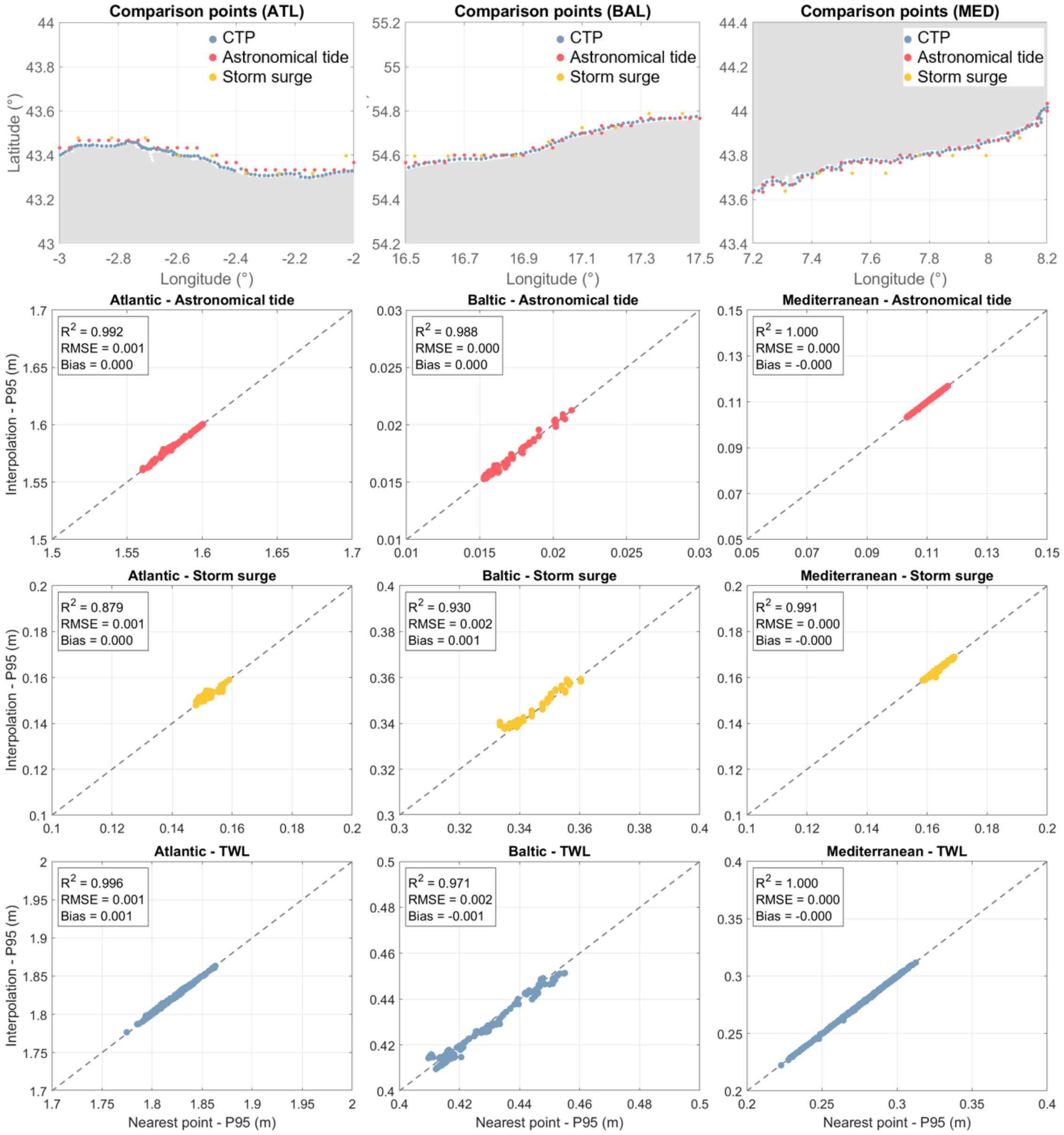

Figure S6 Sensitivity analysis of selection of corresponding marine database information based on P95. Examples are shown for smaller areas located in each of the three European regions identified in the study: Atlantic coast (left column), Baltic Sea (middle column), and Mediterranean Sea (right column).

4. Issue with Figure

Figure 2: it's hard to see the different time series for each location across all the parameters particularly Tide and TWL.

Figure 2 has been updated. A different yaxis is now being used for the points located in microtidal environments as the magnitudes of tide and TWL are much lower than the remaining points, which is the reason behind the difficulty in visualization observed by the reviewer.

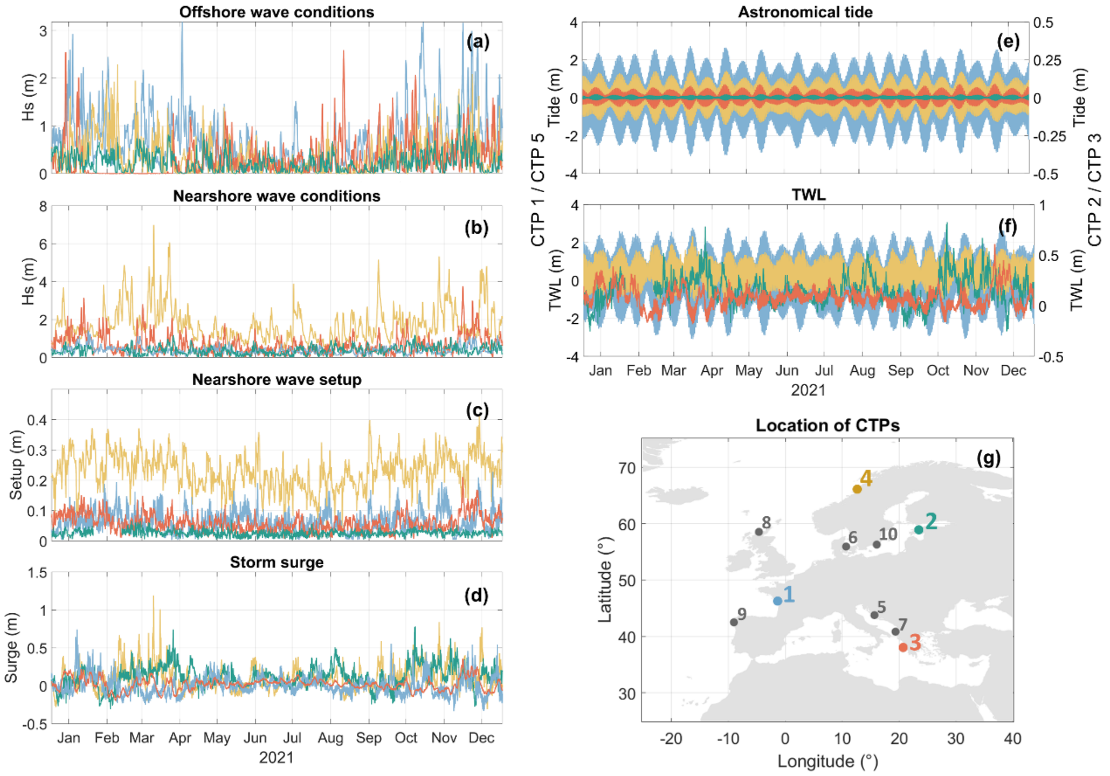

Figure 2: Hindcast time series referred to 2021 for the following variables: (a) offshore significant wave height, (b) nearshore significant wave height, (c) nearshore wave setup, (d) storm surge, (e) astronomical tide, and (f) TWL. (g) Location of the CTPs used as examples. Time series are represented in the same colors as the corresponding CTP. Gray CTPs represent the remaining six test points selected with K-means. Note. Left axis in (e) and (f) refer to CTPs 1 and 4, right axis to CTPs 2 and 3.

References

Dalinghaus, C., Coco, G., and Higuera, P.: Assessing ToTal Water Level Uncertainties Using Global Sensitivity Analysis, J. Geophys. Res. Ocean., 130, 460–466, https://doi.org/10.1029/2025JC023011, 2025.

Haigh, I. D., Wijeratne, E. M. S., MacPherson, L. R., Pattiaratchi, C. B., Mason, M. S., Crompton, R. P., and George, S.: Estimating present day extreme water level exceedance probabilities around the coastline of Australia: Tides, extra-tropical storm surges and mean sea level, Clim. Dyn., 42, 121–138, https://doi.org/10.1007/s00382-012-1652-1, 2014.

Liu, P., Din, A. H., and Hamden, M. H.: Reconstructing hourly coastal total sea levels and assessing current and future extreme sea levels threats to the Coast of China, Sci. Rep., 1–16, https://doi.org/10.1038/s41598-025-22493-2, 2025.

Lobeto, H., Semedo, A., Lemos, G., Dastgheib, A., Menendez, M., Ranasinghe, R., and Bidlot, J.-R.: Global coastal wave storminess, Sci. Rep., 14, https://doi.org/10.1038/s41598-024-51420-0, 2024.

Marcos, M. and Woodworth, P. L.: Journal of Geophysical Research : Oceans of the North Atlantic and the Gulf of Mexico, J. Geophys. Res. Ocean., 7031–7048, https://doi.org/10.1002/2017JC013065.Received, 2017.

Palmer, K., Watson, C. S., Power, H. E., and Hunter, J. R.: Quantifying the Mean Sea Level, Tide, and Surge Contributions to Changing Coastal High Water Levels, J. Geophys. Res. Ocean., 129, https://doi.org/10.1029/2023JC020737, 2024.

Vousdoukas, M. I., Mentaschi, L., Voukouvalas, E., Verlaan, M., Jevrejeva, S., Jackson, L. P., and Feyen, L.: Global probabilistic projections of extreme sea levels show intensification of coastal flood hazard, Nat. Commun., 9, 1–12, https://doi.org/10.1038/s41467-018-04692-w, 2018a.

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

Woodworth, P. L., Melet, A., Marcos, M., Ray, R. D., Wöppelmann, G., Sasaki, Y. N., Cirano, M., Hibbert, A., Huthnance, J. M., Monserrat, S., and Merrifield, M. A.: Forcing Factors Affecting Sea Level Changes at the Coast, Springer Netherlands, 1351–1397 pp., https://doi.org/10.1007/s10712-019-09531-1, 2019.