Global application of a regional frequency analysis on extreme sea levels

Remarks to the Author

The authors present an application of regional frequency analysis to estimate extreme sea level exceedance probabilities along the global coastline from tide gauge data and GTSM-ERA5 hindcast. The paper is well-written and well-structured and addresses the relevant issue of estimating extreme water levels at a global level. The regional frequency analysis is a useful approach when dealing with areas where data from local tide gauges is absent or insufficient, that was previously applied to different regions and with this paper to the entire world. The specific objectives of the study are outlined in the Introduction and properly referred to in the manuscript. The methodology is organized in a fluent way although some passages need to be better explained and justified. The validation with Cyclone Yasi is very specific, there is no explanation of how it was chosen nor is there any discussion of the possibility that it was a specific case. The Discussion requires some revision as the limitations of the methodology were not fully explored. Assumptions and simplifications must be done in a global assessment, especially when dealing with global coarse databases. However, the implications and possible errors introduced in the result need to be addressed and discussed.

Abstract

Line 45: The authors refer here to their methodology as innovative, however, there is no mention in the discussion of how the methodology from Sweet et al., 2022 was modified and possibly improved.

1. Introduction

Line 113-115: while this is true, it is not a limitation that the RFA can overcome. The data used within the paper have a time constraint as well (1979-2018 for the GTSM-ERA5) meaning this is a limit that remains.

2. Data

Line 211: what is the spatial resolution of the HYBRID-CNES-CLS18-CMEMS2020 MDT dataset?

3. Methods

Line 243: can the authors explicitly describe what areas are “unsuitable” for the RFA? Although these areas are explained later in the methods, I suggest the authors give some hints already to better understand Figure 1 as well. Another suggestion would be to include the wording “unsuitable” and “suitable” in Figure 3 for coherence with Figure 1, although the colours are already a good indication.

Line 256: were the authors applying a quality check to detect spikes, wrong values, large data gaps, and vertical datum issues?

Line 262: “*the FES2014 tidal elevations performed better than those of GTSM”:* how was the comparison between GTSM-ERA5 and FES2014 tidal elevations performed? Were they evaluated against tide gauge data or was it proved in previous studies?

Line 267: Estimating the wave set up as 20% of Hs is indeed very common and convenient and probably the most appropriate choice given the global scale of this application. However, there are limitations to this method and considerations to be made. There needs to be a mention here or in the discussion of the fact that this is a rule of thumb that does not capture the influence on wave setup of the specific coastal geometry, bathymetry and local waves and wind conditions.

Line 271: The statement suggesting that wave setup is lacking in sheltered areas, such as bays and estuaries, needs clarification. While it might be necessary to make a simplification and consider no wave setup in sheltered areas there are limitations in doing so by only looking at the topography of the coastline. Wave setup in bays and estuaries may indeed be lower compared to exposed coastal regions but is strongly influenced by the specific orientation, shape, and depth characteristics.

Line 282: if a classification of sheltered/exposed coastline was already done why the authors why are the tide gauges "assumed to be located in sheltered regions" and a check was not performed?

Line 292: Why is there an upper limit to the amount of water level records considered in a 400km radius area ("maximum of 10")? Was that the maximum number observed? There is no explanation or sensitivity analysis of how such thresholds were defined.

Line 306: if 836 tide gauges were used in the application of the RFA why in "3.1 Data processing" do the authors state that "A total of 2,223 tide gauges with a mean record length of 21.4 years were used in the RFA."?

Line 308: In Fig. 2A the "example grid cell" colour doesn't match the actual colour due to overlapped layers. Would suggest changing the colour in the legend to match that of the cell.

Line 321: There is no mention of the reason why in some areas the density of record locations from GTSM-ERA5 is too low (which I suppose means a minimum of 3 points?). Especially in Europe where the resolution is even higher (1.25km, as visible in Figure 1 from Muis et al., 2020) why are there green regions (single-site GTSM-ERA5 analysis) in Figure 3?

Line 330: Why 19 years? Is it the minimum number of years available from the tide gauges? Is this vertical datum adjustment applied to GTSM-ERA5 data (which only extends until 2018) or only to tide gauges?

Line 338: empty line.

Line 340: The implications of choosing a unique value of 4-days for the declustering around the global coastline must be addressed, as High et al., 2016 analysed events in the coastline around the UK. The effectiveness of the 4-day window may vary in regions with different climates, meteorological patterns, and coastal configurations and it is important to acknowledge such limitations (see the chapter from Harley 2017 in Coastal storm definition, "*Coastal storms: processes and impacts*", for the meteorological independence

criterion variability). Was a sensitivity analysis performed to understand how the declustering method affects the resulting return period curves? Moreover, what do the authors mean by “moving window of the storm”? How is a storm defined?

Line 341: “This window is selected…”

Line 342: No reference for estimating the index flood as the 98th percentile of the declustered daily highest water levels?

Line 365: By applying a fixed threshold without local considerations isn’t there the risk to exclude exceptionally extreme events from the analysis? Was a sensitivity analysis performed on this value? Once again, extreme water level events can exhibit considerable variability based on the local characteristics of the region, including bathymetry, coastal morphology, and storm dynamics. Applying a fixed threshold globally may not capture this local variability effectively. In fact, both the Gulf of Mexico and Japan are subjected to extreme weather events and storm surges.

Line 380: “The index u is then estimated…”. First, the u index is estimated, and then the LEWLs can be estimated as well. The order of the sentences must be inverted.

Line 398: Would substitute “Q99 tidal elevations” with 99th percentile for coherence with line 343 “98th percentile”.

Line 394: It isn’t clear nor mentioned why a bias correction is needed.

Line 420-423: I would move this part of the methods in the “2. Data” section as it describes how COAST-RP was derived.

4. Results

Line 487: the example of cyclone Yasi wasn’t under sampled in the historical record. Two gauges out of 10 recorded it because those were at the impacted locations. I would remove line 485 to 487 from here.

Line 503: I do not agree. The Cardwell tide gauge is in a unique location compared to the rest of the gauges in the area, as it is located at the back of a semi-enclosed bay and at the northern mouth of the Hinchinbrook Channel. The specific shape and morphology around the Cardwell tide gauge can contribute to the funneling and amplification of the storm surge on the left side of the storm track (where the onshore winds push water toward the coast). Moreover, not “any local effects due to surge” can be accounted for using the index flood but only those variations that are within the scale of resolution of the analysis. That would correspond to 1 km when considering the interpolated points, but 1.5/2.5 km if considering instead the original resolution of GTSM-ERA5 (which is using bathymetric data from EMODnet and GEBCO therefore even lower resolutions). It is important to clarify the spatial scale of the application of extreme water level values derived from global models, as overconfidence in the results can lead to errors at the local level.

Line 519: I would adjust for coherence the plot in Fig. 5D and 5E so that the return periods on the x-axis extend from 1 to 1000 (see Fig. 4C and 4D).

Line 520: To demonstrate the increase in robustness from the RFA would have been interesting to include a mention of the confidence intervals, how do they compare with those from the individual tide gauges?

Line 538: The results shown in Fig. 5E show an increase in the return period curves and therefore a "spread" of the extreme event due to the regionalization process. However, the Cardwell curve significantly decreased and the information relative to the 3 m extreme event was lost. The authors should not only reinforce the strength of the RFA in propagating the risk in the entire region considered but should discuss a) the fact that at Cardwell and Clump Point, there is an underestimation (or "smoothing out", as the authors wrote in Line 559) of the extreme water levels from the return period curves estimated with the RFA approach and b) the possible negative implications of increasing the return levels at the other gauge locations. While RFA is a valuable tool, it is crucial to consider its limitations and potential challenges, especially when interpreting results for coastal management and safety assessments. An overestimation of extreme water levels can have several implications, potentially resulting in unnecessary costs for risk mitigation measures and infrastructure design that often needs to be optimized.

Line 555: After looking at the example in Fig. 5, I would say that the increase in the RFA is not only related to the under sampling of rare surge events. During cyclone Yasi there were 10 tide gauges in the area actively recording and only two of those were impacted by the rare surge event. Therefore, the fact that in Fig. 6A there are positive differences is associated with the RFA regionalizing extreme events. The wording "under sampling" gives an idea of a lack of in situ measurements or short records, but the GTSM-ERA5 data are also available for 40 years only and with 1.5/2.5 km resolution. I would write a disclaimer about the fact that extreme events are under sampled because of their nature, they are rare and might have never occurred at all at a specific location apart from not being recorded because of the scarcity of in situ tide gauges.

Line 557: Why were the authors not assessing the impact of the limiter on the shape parameter? It is a relevant difference whether the decreases observed in the return levels are associated with a threshold that could be modified or rather by the regionalization method itself.

Line 573: Here and along the paper I would use "quantification of the increase" or quantification of the incorporation of tropical cyclones as their associated risk is in fact better captured. Assuming that an increased extreme water level exceedance probability by the RFA is an improvement is a step further that needs to be discussed. The authors assume from the Introduction onwards that increasing the ESL exceedance probabilities through the RFA approach is an improvement before discussing the implications of this assumption. Just because a region could possibly be affected by e.g. tropical cyclones it doesn't mean that the solution to account for that is to have an increased return period curve of extreme water levels.

Line 583: Why is there an oversampling of extreme events and how was that checked?

Line 615: For consistency use under sampled not “under-sampled”.

Line 620: In Fig. 8C I mostly see an increase (in some cases by 50%) in ESL exceedance probability from the RFA when compared to the single gauges, with the distribution being overall positively biased (overestimation of extreme water levels from the RFA).

Line 651: The water level at the gauge is not mischaracterized, that is an in-situ measurement. What can be mischaracterized is the overall extreme event in case there aren’t tide gauges in the location where the maximum water levels are reached.

Line 652: Change “undersampled” to “under sampled”.

Line 654: I think would be more correct to state that the issues were improved, not overcome.

Line 664: While it is true that a single site analysis of tide gauge data can underpredict the regional risk of ESL generated by tropical cyclones, overpredicting it has also its drawbacks. Please discuss the implications of disseminating the hazards over a larger area.

Line 668: “high degree of accuracy”

Line 677: In addition, Hudson Bay showed an unexplained increase in extreme water levels that does not reflect any cyclone activity.

Line 681: While RFA is designed to handle short and incomplete tide gauge records, the reliability of the analysis is still contingent on having a sufficient number of extreme events for robust parameter estimation. In regions with very limited data, uncertainty in estimates may be high. However, while it is true that GTSM-ERA5 has a temporal limit, saying that “*As such, the RFA has little basis upon which to draw data from when characterising rare extreme events*.” seems in contrast with what was previously said:
“672 [….] Once again, the RFA provides a solution to this problem. As demonstrated in
673 Fig. 7, the distribution of increases to local return levels made by the RFA broadly follows
674 the same patterns globally as the differences between COAST-RP and GTSM-ERA5. This
675 highlights the ability of the RFA to characterise tropical cyclone hazard which is typically
676 underrepresented as a result of short records.”
Please consider rewriting this section to avoid contradictions when exploring the limits and advantages of the RFA.

Line 683: The term “tropical cyclone” was used 41 times in the manuscript, consider adopting the commonly used “TC” abbreviation.

Line 698: The downscaling process assumes that the extreme event observed at one location can be used to estimate the risk for nearby areas. This may be less reliable in regions with diverse coastal characteristics and complex topography. At specific coastal areas a much finer resolution analysis than the RFA is needed before applying inundation models. It should

be made clear that it is advisable to supplement RFA with local knowledge, consider uncertainties in the results, and conduct sensitivity analyses to understand the robustness of the findings, especially in the context of coastal management and safety assessments.

Line 721: This is why can be useful to have an additional plot (e.g. in the Appendix) that could complement Fig. 4A and 4B with confidence intervals or with some sort of indication of the amount of data that was used for the RFA.

Line 734: There is a wave buoy dataset distributed by CMEMS that integrates observations from around the world: the “INSITU_GLO_WAVE_REP_OBSERVATIONS_013_045” at https://doi.org/10.17882/70345).