

Authors Comments 2

RC2: 'Comment on sp-2025-8', Anonymous Referee #2, 22 Dec 2025

Major comments:

This is my first review of 'Metocean study for the planning of the FSRU terminal: a case study at the Estonian coast of the Gulf of Finland' by Raudsepp et al., in which the authors present extreme value statistics for metocean variables, based on model data and observations in the Baltic Sea. The study is motivated by the vulnerability of marine infrastructures – especially the Paldiski FSRU terminal. As the first reviewer pointed out, as it is now, the manuscript finds itself somewhere in between a report and a manuscript. The scope of the 'state of the planet' is such that a more report like submission may do, but below I make some suggestions for improvements – mainly focusing on reducing the uncertainties in the Baltic wide analysis (see comment 9) which I think should be the main focus of the manuscript (see comment 1). I consider this to be a major revision that requires some dedicated work, but at the same time I think the revisions can be quite straightforward to implement.

1) I would suggest flipping the manuscript upside down. I think the readers would find it much more interesting to learn about Baltic Sea wide extreme values, rather than this one very specific location off the coast of Estonia. I would therefore change the title to be something like 'Extreme winds, sea level, waves, and currents in the Baltic Sea' and start with section 3.5 and use the FSRU site as a validation case study → if possible, I would also consider adding more station observations (e.g. from the Copernicus in-situ TAC) for validation.

We thank the reviewer for the constructive restructuring suggestion and fully agree with the principle behind it: a Baltic-wide synthesis can be valuable if the underlying uncertainties are sufficiently constrained and the validation is extensive. In our original submission we indeed included a Baltic-wide extension (former Section 3.5) and highlighted it in the abstract and conclusions.

However, after considering both reviewers' feedback—especially concerns about novelty, dataset limitations, and uncertainty—we decided that retaining the Baltic-wide mapping would require a substantially larger effort (multi-dataset ensemble, broad spatial validation, and regional bias characterization) to reach a defensible scientific standard. Because this would fundamentally change the scope and length of the manuscript, we instead chose a major revision in the opposite direction: we removed the Baltic-wide analysis and substantially strengthened the Paldiski case study to deliver a clear, rigorous, and reproducible contribution.

Importantly, we still implement the spirit of the reviewer's suggestion by:

- expanding validation using additional observational data sources where possible (including broader in situ records / Copernicus in-situ products when available) to strengthen the robustness of the selected datasets and the transferability of the workflow, and

- explicitly reframing the manuscript away from a local “engineering report” toward a scientific evaluation of data suitability and uncertainty for nearshore extremes in the Gulf of Finland.

We believe this revised focus is more consistent with what can be robustly supported by the available datasets within a single paper, while still producing outputs of high practical value and methodological relevance.

2) In relation to the previous comment, framing the study around offshore infrastructure makes sense, but falls a bit short if it is not taken to the engineering level (computing actual expected loads to structures, which would allow structural engineers to directly use these results). I think such work is out of scope for the present study, which is why I suggest recasting the study along the lines of comment 1 – also in the abstract/introduction.

Authors’ responses:

We agree. The present study provides environmental boundary conditions (return levels and uncertainty) that can be used as inputs to engineering analyses, but it does not compute structural loads or responses. In the revision we will ensure the manuscript does not imply that it delivers full engineering design loads, and we will clarify the intended use of the results as metocean design criteria inputs.

3) The authors do not provide much quantitative details on the basis by which the higher resolution regional wind products were dropped out (apart from the general statements about the biased extreme values) – was this decision based on the Paldinski site alone? Although it is a bit of work, I would ask the authors to use CERRA data in addition to the ERA5. CERRA5 has in many occasions been shown to outperform ERA5 and I believe it should allow reducing the uncertainties in the Baltic wide analysis.

Authors’ responses:

We agree that the current manuscript does not provide sufficient quantitative justification for excluding COSMO-REA6, BaltAn65+, and NORA3. In the revision we will provide objective metrics (e.g., bias, RMSE, correlation, and tail-focused metrics such as the 95th/99th percentile and annual maxima) for the candidate wind products against available observations. We also will include CERRA winds in site specific analysis.

4) Overall I think it is a bit misleading to list the different datasets in table 1 and then not use them. However, linked to the comment 9, I would encourage the authors to reconsider using those datasets (and CERRA) in order to narrow down the uncertainties in the extreme values.

We agree. In the revision we will make Table 1 consistent with actual usage. Where feasible, we will also use multiple datasets as an ensemble to quantify dataset-driven uncertainty, especially for winds.

5) L148 – which model output was used to fill the tide gauge data-gaps?

Authors' responses:

Thank you for pointing this out. In the revision we will specify explicitly that the gaps in the Paldiski North Port tide gauge record (2010–2016) were filled using the CMEMS Baltic Sea Physics Reanalysis sea surface height time series (product ref. no. 1) at the nearest grid cell, after aligning the model series to the gauge datum and mean level over the overlapping period.

6) L164-166. The authors mention the rather large (15% to SWH and 5% to periods!) corrections to wave properties, but there is no figure or reference to back this up. These corrections are quite large so perhaps the authors could add a supplementary figure with a bit more careful description on how these numbers were decided. These percentages seem to be quite round so the reader is left wondering if they are really coming out of a fit, or rather based on expert judgment?

Authors' responses:

We agree. The current manuscript states fixed corrections for significant wave height and wave periods based on a short buoy deployment, but does not provide the supporting comparison plots or a clear description of how the correction factors were computed. In the revision we will add a supplementary figure comparing modelled and observed wave parameters (scatter and Q–Q plots) and clearly describe the calculation of the correction factors.

7) L176 – to my understanding the Peaks Over Threshold method requires some way of ensuring that the same event is not considered multiple times. How did the authors do this?

Authors' responses:

We agree that POT requires an explicit independence treatment (declustering) and that this must be described. In the revision we will add a clear declustering procedure for POT/PUT analyses (e.g., run-length declustering with a minimum temporal separation window), and we will retain only the cluster maxima/minima as independent extreme events.

8) section 2.6 – did the authors apply the site specific corrections for winds and waves to the full regional dataset? I am assuming not, but it would be good to explicitly state that.

Authors' responses:

We thank the reviewer for raising this point. Site-specific correction factors were not applied to any basin-wide/regional dataset. In the original manuscript, the wind and wave corrections were introduced only for the Paldiski case-study location (e.g., wind scaling to better match local maxima and wave-height/period adjustments based on the short-term local buoy comparison). These corrections were not intended to be transferable spatially.

In the revised manuscript, we have removed the Baltic Sea-wide analysis and the associated “Regional extreme value analysis” section (former Section 2.6). Consequently, there is no longer any regional EVA to which such site-specific corrections could be applied. To avoid ambiguity, we will add an explicit statement in the local wind and wave methodology sections clarifying that the correction factors are site-specific (Paldiski only) and are not propagated beyond the case study.

9) The wind speed return values seem low and the spread is very large – almost to the point that is becomes useless. The error has also a very peculiar round maximums, possible around points where ERA5

has assimilated (or not?) some actual observations. I think the authors should make their utmost to reduce that uncertainty range and provide explanation for the peculiar shapes. For example, it seems very unlikely to me that moving a few kilometers east to west in the region north of Gotland island, one would have close to the same return period (22 m/s, which, btw seems very low) but a difference in the confidence interval that ranges from 1 m/s to well over 10 m/s (colormap is saturated).

I had a look at some other publications, for example this one

<https://www.sciencedirect.com/science/article/pii/S2212094721000530> who show similar maps of return values. Their colormap is not the easiest, but it looks to me like many models produce much higher values over the Baltic for 30-year return values than what the authors show for 50-year return values. Especially in coastal areas, I wonder how much of this is due to the ERA5 resolution.

I would ask the authors to check if the other datasets produce similar uncertainty range to the return period values and/or if the peculiar shapes could be attributed to the assimilation methods used in ERA5 (would be an important result I think). Use of CERRA dataset might help with robustness. I also encourage the authors to explore other statistical methods to narrow down the confidence interval, although I don't have direct suggestions here.

Some of the other variables show similarly variable uncertainty fields, some of which might be understandable and linked to strong gradients, but for example the one for maximum SLA is rather difficult to grasp. Why is the relatively low return period in the central Baltic so uncertain (by well over a meter)? Again, I think it would be important that the authors review their statistical methods and aim to narrow the uncertainty range and/or explain where the uncertainty comes from.

Authors' responses:

We thank the reviewer for this detailed and very helpful assessment. We agree that the Baltic Sea-wide return-value and uncertainty maps (in particular for wind speed and maximum sea level anomaly) in the original manuscript were not sufficiently robust and were difficult to interpret, including the unusually large and spatially heterogeneous confidence intervals and the “patchy/round” patterns noted by the reviewer.

After considering this comment together with the other reviewers' concerns regarding regional applicability, uncertainty, and defensibility, we have therefore removed the Baltic Sea-wide extension of the extreme value analysis from the revised manuscript (former Section 2.6 “Regional extreme value analysis” and former Section 3.5 “Extension of the Extreme value analysis to the whole Baltic Sea”, including Fig. 4). Consequently, the revised manuscript no longer presents basin-wide wind/SLA return values nor their spatial uncertainty fields, and we do not attempt to interpret potential spatial artifacts related to resolution or assimilation at the Baltic-basin scale.

Importantly, we still fully acknowledge the underlying methodological concern raised by the reviewer (i.e., the need for robust extremes and meaningful uncertainty estimates). In the revised manuscript—now focused on the Paldiski case study—we strengthen the site-specific analysis by:

- clarifying the wind metric and averaging used (to avoid inappropriate comparisons with studies using gusts or different averaging periods),

- improving the uncertainty quantification and EVA documentation for the Paldiski site (including declustering/independence treatment for POT and sensitivity checks), and
- where feasible, adding an additional dataset comparison (e.g., CERRA vs ERA5) as a local robustness check, consistent with the reviewer's recommendation—without extrapolating such comparisons to basin-wide conclusions.

A comprehensive investigation of whether the Baltic-wide uncertainty patterns arise from ERA5 assimilation/resolution effects versus EVA/bootstrapping choices would require a dedicated basin-scale multi-dataset ensemble and extensive validation, and is therefore beyond the scope of the revised (Paldiski-focused) manuscript.

10) Data Availability statement. I would ask the authors to share the resulting return periods (i.e. data that can be used to reproduce figures 2-3 and especially 4), their uncertainties, along with codes to reproduce them from original data. This could be done via Zenodo repository. Sharing the data and the codes is a standard today, and allows the interested readers to understand what has been done exactly (especially since the manuscript has no equations), and since the resulting data can be very useful, and is not very large, I see no reason why the authors should not share it.

Authors' responses:

We will publish the data on return periods and scripts to reproduce the figures.