

Response to Reviewer

Manuscript: "Metocean study for the planning of the FSRU terminal: a case study at the Estonian coast of the Gulf of Finland" (sp-2025-8)

Dear Editor and Reviewer,

We thank the reviewer for their careful assessment of the manuscript and constructive suggestions. We have substantially revised the manuscript to clarify its scientific contribution, strengthen the Baltic Sea-wide analysis, document dataset selection and bias treatment, and improve transparency and reproducibility.

Below, the reviewer comments are reproduced in italics, followed by our point-by-point responses and descriptions of the corresponding changes made in the revised manuscript.

RC1: 'Comment on sp-2025-8', Anonymous Referee #3, 04 Sep 2026

General Assessment:

The manuscript investigates metocean extreme conditions at the site of the FSRU terminal near Pakrineeme/Paldiski and across the Baltic Sea using preprocessed reanalysis/hindcast products, site specific observations and modelling products in combination with standard extreme value analysis methods. The main objective is to demonstrate the usefulness of the data sets and preprocessing methods for the assessment of metocean conditions and to provide Baltic Sea wide assessment results for regional extreme value studies. The authors conclude that the available local data, preprocessing and assessment methods are suitable for site specific metocean studies the site of the FSRU terminal, but that regional assessment results can only be used to evaluate the regional variability of the results. Overall, the manuscript addresses several interesting questions that are relevant to metocean assessments.

The study has several strengths, particularly the strong data foundation, which combines state-of-the-art multi-year products, with local observations and high-resolution model results to provide a comprehensive data basis for the extreme value assessments. The study demonstrates the authors strong expertise in the numerical modelling and the hydrography of the Baltic Sea. However, several issues, outlined below, should be addressed before the manuscript can be considered for publication.

My main comments are given below

Major comments:

The manuscript investigated the met.ocean extreme conditions at the site of the FSRU terminal near Pakrineeme/Paldiski and in the entire Baltic Sea. The first part of the paper presents a comprehensive, but somewhat site-specific study, focusing on the site of FSRU terminal. This might be of less scientific interest for non-local readers, who are not familiar with the site. To generalize the results, one could maybe try to address the benefits of the applied data sets and data harmonization methods for similar studies at other sites in the Baltic Sea, but this is not easily done, as the Baltic Sea environment is very divers, as the authors are expressing as well. In the second part of the study, the local, site-specific assessment at the

FSRU terminal is complemented by a Baltic Sea-wide analysis of extreme conditions. The geographically larger scope of this study makes the paper interesting for a larger user community. However, there is not a clear research question formulated. The study is motivated by the relevance of the Baltic Sea wide met.ocean assessment for similar studies.

Response: We agree that the research question and the broader relevance of the site-specific analysis were not stated explicitly enough. We have therefore revised the end of the Introduction to formulate the study around two linked questions: first, how long-term multi-source metocean datasets can be used to characterize extreme conditions at a specific infrastructure site when local observations and high-resolution modelling are available; and second, how those site-specific results can be placed within the spatial variability of metocean extremes across the Baltic Sea. We also clarify that the transferable element of the study is the workflow for combining regional products with local evaluation and preprocessing, rather than the numerical correction factors derived at Paldiski, which are site-specific.

The study follows a logical progression from a local assessment at a specific site (Pakrineeme/Paldiski) to a broader, more general assessment covering the entire Baltic Sea. However, this progression could be communicated more clearly by establishing a stronger narrative that connects the two parts of the study. In its current form, the local and Baltic Sea-wide assessments appear somewhat separate rather than as an extension of the same analysis. This might be motivated by the difference in data preprocessing prior to the analysis. But the local, site-specific EVA for the FSRU terminal could provide information on the performance of the regional, Baltic Sea wide assessment.

Response: We agree that the relationship between the two scales should be clearer. We have revised Sections 2.6 and 3.5, the uncertainty discussion, and the Conclusions to establish a stronger narrative connecting the local and Baltic-wide assessments.

The study employs (1) numerical methods for wave and ocean circulation modelling and downscaling, (2) post-processing and harmonization techniques to improve the quality and consistency of the input data, and (3) established statistical methods to estimate return levels and their associated uncertainties. The applied data sets and models are state-of-the art and well suited for the assessment of metocean conditions. The limitations of the modelling approach are well communicated and analyzed. The use of the data postprocessing and harmonization methods is justified, but could be communicated more clearly within the paper. Currently, they are described in the supplementary material. The applied methods for extreme value analysis are standard, but their limitations with regard to analyzing extreme events with long return periods (100-years) could be communicated in the paper. The strength and weakness of the methods could be described in a bit more detail. It could also be mentioned why Block maxima method was used, while the POT method was discarded for the regional assessments, even though this might just have been done to limit the computational effort of this study.

Response: We agree and have clarified the methodological choices in the revised manuscript. We now describe the main strengths and limitations of the block-maxima and POT approaches, qualify the

extrapolation involved in the 100-year return levels, and explain why annual-maxima GEV fitting was used for the Baltic-wide analysis. The regional choice was made primarily to maintain a consistent procedure across all grid cells, whereas POT would require location- and variable-specific threshold selection and declustering criteria. We therefore emphasize the 50-year return levels for the regional comparison and retain the 100-year site-specific estimates as explicitly qualified longer-return-period extrapolations.

The used data sets are of best-available quality and appropriate for the study at hand. I only miss a detailed description in the form of a table, including spatial and temporal resolution and method of pre-processing. Maybe such a table could be added.

Response: We have added a new table (Table A2) summarizing the datasets used in the study, including their temporal and spatial resolution, period used, and study-specific preprocessing. The table also distinguishes where different preprocessing was applied for the site-specific and Baltic Sea-wide assessments.

The results are clearly presented, but some figures are rather crowded with too many panels. An example is figure 4, with 12 panels (6x2). More details are given in the extended assessment below.

Response: We appreciate the suggestion. We considered splitting Figure 4, but decided to retain it as a single multi-panel figure because the panels follow a consistent structure, with the 50-year return level shown alongside its corresponding uncertainty for each variable. Keeping these panels together facilitates direct comparison of the spatial patterns and uncertainties across the different metocean variables and supports the integrated Baltic Sea-wide interpretation. We have checked the figure layout and labeling to ensure that the individual panels remain readable at publication size.

Generally, the figures support the conclusions of the paper, but sometimes there is a mismatch. This might be because the figures were changed at a later stage without correcting the manuscript. Example is the distribution functions of SWAN and WAM (Copernicus Marine reanalysis) results, which show a consistently higher distribution for SWAN for waves above 0.4 m, whereas in the document states that this is only the case for moderate waves. More details are given in the extended assessment below.

Response: We agree. The previous text overinterpreted the difference between the two wave datasets and was not fully consistent with Figure 2. We have removed the statement that the difference between moderate and large waves results from the different record lengths and revised the discussion to describe only the distributional difference shown in the figure. We have also corrected the Figure 2 caption so that the actual time period used for each dataset is clear rather than implying that all datasets necessarily cover the same period.

.

The interpretation of the results is sound and is based on a thorough analysis of the assessment results, supported by a good understanding of the physical conditions in the Baltic Sea and the Gulf of Finland. The conclusions are supported by the presented results. The main findings are clearly identified and support the rationale of the paper.

Response: We thank the reviewer for kind words.

Minor comments

Line 33: “place local extreme” is not standard English.

Response: We replaced the sentence ending with: "The local estimates are compared with Baltic Sea-wide patterns of metocean extremes."

Chapter 1. Introduction

Line 41: “to make informed decisions”

Response: We agree and rewrote "to make informed decisions on subsequent engineering assessment and operational planning."

Line 43: “underestimating the importance of this step.”

Response: We replaced the wording with "Omitting the metocean assessment, or underestimating its importance, can compromise the environmental information available for subsequent engineering decisions."

Line 51-52: “is such a case” rather than “is a case in point”

Response: We corrected the text.

Line 54: “operational under all conditions” rather than “operational when most needed”, which could depend on the geopolitical situation.

Response: We agree that the geopolitical wording should be removed. However, "operational under all conditions" would imply a capability that the study does not assess. We rewrote instead: "Reliable metocean information is needed to assess structural safety and define permissible operating conditions."

Line 59-62: This sentence is rather a long and could be restructured for clarity.

Response: We divided the sentence: "The Gulf of Finland is a long, narrow sub-basin with complex coastal geometry. Its geometry and wind forcing can produce substantial non-tidal water-level variability, including storm-driven setup and the seiches (Medvedev and Kulikov, 2021; Barzandeh et al., 2025; Lorenz et al., 2025)."

Line 62: “Limited fetch” rather than “enclosed fetch”. Maybe “enclosed geography”

Response: We rewrote: "The gulf's geometry limits...".

After line 68, introduction of the Gulf of Finland: It would be good to describe the study site with a link to

Figure 1. I can't find where figure 1 is referenced in the document.

Response: We added a short site-description paragraph immediately after the Gulf of Finland context: "The Pakrineeme/Paldiski study site lies on the Pakri Peninsula, on Estonia's southern Gulf of Finland coast (Fig. 1), in a predominantly bedrock-controlled coastal setting. The FSRU berth has a representative water depth of approximately 16 m. Compared with the open Baltic Proper, the site is relatively sheltered from waves but remains susceptible to sea-level extremes and episodic wind-driven events." Figure 1 was already cited earlier in the Introduction.

Figure 1: Full references for the cited publications in the figures caption are not provided. Abbreviations like GIE and GEM; and even EMODnet should be explained. What does the red line in panel (b) represent.

Response: We have revised the caption of Figure 1 to provide the full references for the cited publications and to define the abbreviations GIE, GEM, and EMODnet. We have also clarified that the red rectangle in panel b) indicates the drone's field of view and the extent of the photograph shown in panel c).

Table1: Abbreviations like QUID and PUM should be explained. I think it would be appreciated if area and time coverage as well as spatial and temporal resolution of the data sets could be provided. This could be done in a separate table.

Response: QUID (Quality Information Document) and PUM (Product User Manual) are already expanded in the first entries of Table 1. We added the requested spatial and temporal information in a separate analysis-input table.

Table1, or rather a separate table: It would be good to add a table listing the data sets, spatial and temporal resolution and data pre-processing methods applied (see table1 in the review). This should be done for the site-specific assessment and the Baltic Sea wide assessment, as both assessments use not exactly the same pre-processing methods.

Response: We added an analysis-input table in the appendix.

Line 82-85: NEMO and GETM are introduced as individual ocean models. How are they employed in the assessment. It is not clear how they are configured. Is NEMO providing boundary conditions for GETM?

Response: We have added a clarification stating that "NEMO-EST05 and GETM are independent downscaling simulations that both use boundary conditions from the CMEMS reanalysis".

Line 96-99: I would distinguish engineering design studies from the assessments of metocean conditions. Engineering design studies are not in the focus of this paper.

Response: We agree. To avoid confusion, we deleted the text on Lines 96-99.

Line 104: I would rather write: "to ensure that the estimates are realistic." This is maybe just a small point, but reality is something that models can only approach.

Response: We used "to assess consistency with the available observations"

Line 105: Is it possible to describe how the observations are used to constrain the model predictions. Is this done by data assimilation?

Response: We removed the term "constrain" in the revised paper. We have used observations for comparison with models and for bias correction.

Line 108: "more reliable metocean characterization than" This is a comparative construction

Response: We replaced the claim with "provides a multi-source characterization and identifies differences among the available estimates."

Figure1(c): Is this a design study of the LNG terminal or the figure of the final construction?

Response: We have clarified in the figure caption that panel (c) is a drone photograph of the LNG terminal after its construction

Chapter 2. Data and methods:

The whole paper: Please change from CMEMS reanalysis of Copernicus marine reanalysis

Response: We have retained the term "CMEMS reanalysis" because CMEMS is defined at its first occurrence as the Copernicus Marine Environment Monitoring Service and the abbreviation is used consistently throughout the manuscript. This concise term also distinguishes the CMEMS marine products from the ECMWF ERA5 atmospheric reanalysis. The complete product names are provided in Table 1 and the corresponding references.

Would it make sense to provide a figure of the time series used for the assessment of the metocean conditions at the FSRU site.

Response: We appreciate this suggestion. However, plotting several decades of hourly records for all metocean variables would produce compressed time-series figures in which individual events and relevant variability would be difficult to distinguish. Since the study focuses on extreme-value behaviour, we instead present the distributions, distribution tails, and EVA diagnostics used in the assessment. To ensure transparency and reproducibility, we have added an explicit reference to Raudsepp et al. (2026b), where the complete site-specific time series and analysis scripts are available for download.

Line 157-158: Could you provide a few more details on the preprocessing of the wind fields for the sitespecific assessments in the paper? You could add the information in the supplement as an appendix to the paper. Otherwise, this part remains a bit unclear.

Response: The wind-field correction methodology and supporting figures are provided in Supplementary Section S1, "Wind correction." We have added a reference to this section in the manuscript.

Line 189-233: Here, the methods for extreme value analysis are introduced. I would recommend separating these paragraphs from the rest of the data-set description chapter and to introduce a separate chapter introducing the methods.

Response: We agree that the extreme-value methodology should be clearly separated from the dataset descriptions. As Section 2 already covers both data and methods, we reorganized the material into a dedicated Section 2.5, Extreme Value Analysis Framework, with separate subsections for the site-specific

analysis (Section 2.5.1) and its Baltic Sea-wide extension (Section 2.5.2). This provides a clearer distinction between the input datasets and the statistical methodology while preserving the overall structure of the Data and Methods section.

Line 190-200: Here the (a.) block maxima method and (b.) the peak over threshold method are introduced. It would be good to mention the strength and weakness of each method, to motivate the use of the methods for the analysis of metocean extremes.

Response: We added: "Block Maxima Method main strengths are straightforward implementation and the absence of a threshold-selection requirement, making it convenient for consistent comparisons across variables and locations. However, retaining only one value per year discards other potentially informative storm events, while including maxima from relatively calm years. By retaining several independent extreme events per year, POT can use the available record more efficiently and potentially improve the precision of return-level estimates. Its main limitations are sensitivity to threshold selection and the need to separate dependent observations associated with the same storm."

Line 196: Can you motivate your choice of assessing a 100-year return period. The data sets you are using is not covering that many years. Sea level: 1993-2009,...,2016-2021=28 years, Waves: 1993-2020=27years. Maybe you can mention that the expected uncertainties are large for that reason. But uncertainties are not provided for the site-specific assessments.

Response: We have added: "A 100-year return level is a fitted probability statement, not an observed event that requires a 100-year record by definition. Nevertheless, a few decades provide limited information about the tail, particularly when the shape parameter is uncertain. Therefore, we mainly focus on the 50-year estimates for local-regional comparison, and retain 100-year estimates only as explicitly qualified longer-return-period extrapolations."

Line 202: I think the term EVA is used here for the first time, without introduction.

Response: "Site-specific Extreme Value Analysis (EVA)" was defined in the title of subsection 2.5.1 title.

Line 205-207: Here you motivate your choice for a Baltic Sea wide assessment of the metocean conditions. Could you add a few more words on how the site-specific and the Baltic-Sea wide assessments complement each other. How a reader should use the results of your assessments. Could they use the Baltic Sea results in a relative way, to compare different location with each other and to identify optimal areas characterized by frequent calm conditions, to support micro-or macro-siting?

Response: We added: "The regional estimates support preliminary comparison of rare-event exposure, while the local analysis addresses site-specific representativeness and preprocessing. However, return-level maps alone do not quantify the frequency or duration of calm conditions, operational weather windows or optimal siting. Those questions require additional time-series, directional, persistence and engineering analyses."

Line 208: Here you mention "correction factor" which is rather unspecific. I would suggest to include the material in the supplement as an appendix and to refer to it here.

Response: We have deleted the Lines 208-213 to avoid confusion. We did not do any bias correction or used any correction factors in regional scale analysis.

Line 208-213: I would suggest restructuring the paragraph, as the current order of the arguments makes it a bit difficult to follow. This paragraph describes the regional assessments, but starts by explaining the approach used for the site-specific assessments, only to state later that this approach was not applied to the regional assessments.

Response: We have deleted the Lines 208-213 to avoid confusion. We did not do any bias correction or used any correction factors in regional scale analysis.

This is a rather unusual sequence, to start with the data processing before the data has been introduced and to describe the method in the last paragraph of the section.

Response: We have changed the structure of the section.

Line 227-233: Here you introduce the extreme value and uncertainty calculation method. I would introduce it earlier. I almost missed it here.

Response: We have changed the structure of the section

Line 227-233: Could you mention why you did not use POT method here. Was there a reason for this?

Response: We added: “Annual-maxima GEV fitting provides a uniform framework for regional comparison. It avoids the location- and variable-specific threshold selection and storm-declustering diagnostics required by POT, simplifying regional implementation. This is a practical methodological choice—not evidence that POT is unsuitable—with the trade-off that additional extreme events within each year are discarded.”

Line 231-233: There are a few unfinished sentences here.

Response: We finished the sentences.

Chapter 3. Results and Discussion:

Line 236: It could be mentioned that the modelled winds represent the “mean wind speed”, representing a 10 minutes average. Wind gusts are not included in the assessment. So, the results of the assessment can be used to characterize the sustained wind loading (Line 249), not the peak loading.

Response: We added that wind gusts were excluded.

Line 238: Here you mention “correction factor” again, which is rather unspecific. I would suggest to include the material in the supplement as an appendix and to refer to it here.

Response: The wind-field correction methodology and supporting figures are provided in Supplementary Section S1, “Wind correction.” We have added a reference to this section in the manuscript.

Line 272: Not all ocean model data sets are hindcasts. Copernicus Marine BaLMFC reanalysis is not a hindcast. I assume NEMO-EST05 and GETM are providing hindcast.

Response: We agree. The Copernicus Marine product is reanalysis will be called a reanalysis, while NEMO-EST05 and GETM results are hindcasts. We have corrected the text.

Line 278-282: I would assume that storm Gudrun is covered by the modelled period. However, the highest sea surface height event I can identify is a near-50-year event of just below 1.6 m, including in the Copernicus Marine BAL MFC product (Fig. 3c). So, should the modelled 100-year event be higher than extreme sea level for storm Gudrun? Is Gudrun not a 100-year event at the FSRU site?

Response: We agree that our original classification of Gudrun as an approximately 100-year event at the FSRU site, based on its exceptional severity across the Baltic Sea, was not justified. A return period describes the exceedance probability of a particular variable at a particular location, not the storm as a whole. We have removed this unsupported interpretation.

Line 292: For visual consistency, it would be preferable to arrange Sections 3.3 and 3.4 in the same order as the corresponding panels in Figure 2.

Response: We have changed the order of subplots in Fig. 2.

Line 298: See my comment before to the use of “correction factors”.

Response: Details of wave correction are provided in Supplement "Wave correcton"

Line 301-306: I can't see that SWAN and WAM results have a qualitatively different behavior for moderate versus large waves. Fig2 (g,h) shows SWAN to be consistently higher than CMEMS WAM wave reanalysis, except for waves below 0.4m. But the figure caption states that all the time series cover the same period 2012-2021. So, what is written here is maybe not correct, as it is mentioned that SWAN and WAM time series have different length.

Response: We agree. We have removed the unsupported moderate-versus-large-wave distinction.

Chapter 3.4: How big is the difference of the daily averaging of the ocean currents? Hourly currents should be stronger and provide a much more suitable design foundation. Why is the Baltic Sea wide regional assessment using hourly currents (Line 376-377), while the site-specific assessment is using daily averages. Could you motivate your choices.

We agree that this part needs further clarifications. In fact, the NEMO-EST05 current fields used for the site-specific analysis are available at 1-hour resolution. They were averaged to daily means only for the direct comparison with the CMEMS current product in Fig. 2, because the latter is available as daily means. The site-specific extreme-value analysis itself uses the native hourly NEMO-EST05 current series and therefore retains the sub-daily current maxima.

We have clarified this distinction in Sections 2.4 and 3.4. We have also clarified that the Baltic-wide current assessment does not use hourly currents; it is based on the original daily-mean CMEMS surface-current product. Consequently, the regional current return levels are intended as spatially consistent screening estimates and may not represent sub-daily peak currents relevant to detailed local engineering assessment.

Figure 3 and discussions. The uncertainty in the assessment is relatively large because the available datasets are rather short, providing only limited information to constrain the tail of the extreme-value distribution. As I understand it GEV calculates the shape parameter from the data and is therefore better able to provide a realistic distribution, especially for short data sets.

Response: We agree that the GEV distribution provides greater flexibility than the Gumbel distribution because its shape parameter is estimated from the data, allowing different forms of tail behaviour to be represented. However, for relatively short annual-maxima records, the shape parameter can itself be weakly constrained and sensitive to individual extreme years. Thus, the additional flexibility of the GEV does not necessarily imply lower uncertainty for long-return-period extrapolation. In Fig. 3 we therefore compare GEV and Gumbel fits rather than assuming that the GEV is automatically preferable, and we retain the distribution supported by the fit and physical plausibility. We have clarified this point in the Methods and further emphasize that the 100-year estimates represent relatively uncertain extrapolations beyond the available record, whereas the 50-year estimates are the main basis for local–regional comparison.

Chapter 3.5 and 3.6. the two chapters provide a description of the regional assessment results and a description of the uncertainties of the applied method. The structure of the chapters could be improved. The wind and wave assessment should be done separately. The assessment of the statistical uncertainties could be linked more consistently to the assessment of the 50-year return values. The statistical uncertainties of the ocean currents are currently not assessed. And there could be a separate chapter providing a synthesis of the site-specific assessments and Baltic Sea wide assessments. This is only done for minimum sea level (Line 395-397).

Thank you for this suggestion. We agree that the distinction between the regional return-level results, their statistical uncertainty, and the broader methodological limitations was not sufficiently clear in the previous structure. We have therefore reorganized the regional assessment so that wind, waves, surface currents, and sea level are discussed separately. For each variable, the 50-year return levels and the corresponding statistical uncertainty are now discussed together.

We have also included the statistical uncertainty of the regional surface-current estimates, which is shown in Fig. 4h and discussed together with the current return levels. In addition, we have introduced a separate section synthesizing the site-specific Paldiski assessment and the Baltic Sea-wide results. This comparison now covers wind, waves, currents, and sea level rather than only minimum sea level, and highlights the differences arising from local exposure as well as spatial and temporal resolution. The former "Uncertainties and limitations" section has consequently been moved (and renamed) after this synthesis and is reserved primarily for methodological limitations that are not represented by the statistical confidence intervals.

Line 352-364: Some sentence could be improved. The text looks a bit hastily written. A few words of introduction could be added. The paragraph jumps right into the analysis of the results.

Response: We have restructured the paragraph.

Line 355: what is a “full interval width”? Is this not the “uncertainty range”?

Response: Corrected

Line 360: “Our analysis shows”

Response: Corrected

Line 366: So far gusts have not been analyzed. This could be mentioned earlier in the paper.

Response: We are mentioning it at the beginning of the Results and Discussion section.

Line 370-371: “limited number of annual maxima”. I think you mean the limited number of years that are covered by the data set.

Response: Corrected

Line 378-386: Has the sea level data been preprocessed prior to the analysis? Was the long-term mean of the model state removed? I assume not, but it would be good to clarify. Usually, modelled long-term means show an easterly increase of sea level.

Response: We have clarified it by adding: "We should note that long-term mean sea level was not removed from the analysed data."

Line 387: “The uncertainties ...”

Response: Corrected

Line 389: I assume you use daily mean reanalysis because that is what the Copernicus Marine reanalysis provides. This is not so clear from reading the text. I suggested to provide a table with spatial and temporal resolution of the data sets. This could help to clarify this.

Response: We have provided the Table in Appendix.

Line 395: “Meanwhile”, maybe better “Correspondingly”.

Response: Corrected

Line 398: throughout the paper “LNG” and “FSRU” are used rather interchangeably. Is it possible to use only only one abbreviation?

Response: We have retained both abbreviations because they refer to different concepts. "FSRU" denotes the floating storage and regasification unit itself, whereas "LNG terminal" refers to the broader terminal area and associated infrastructure, including the quay, mooring facilities, transfer systems and onshore gas connection. We have reviewed the manuscript and standardized the terminology to ensure that each abbreviation is used consistently according to this distinction.

Line 401-406: This paragraph mentions the use of metocean data for defining engineering specifications, while earlier on (Line 97-99) you mentioned that combined structural and metocean assessments are required for this.

Response: We have corrected the first sentence of the paragraph: "Metocean criteria derived from local and Baltic-wide EVA provide environmental inputs for subsequent, separate engineering analyses for the engineering specifications for mooring systems, platform freeboards, breakwater heights, and operational limits."

Chapter 3.6: I would recommend to change the title and the use of the work "uncertainty" in this chapter to reflect the different terminology and meaning. Here we discuss "physical" uncertainties and shortcomings of the input data, rather than "statistical" uncertainties as a result of the analysis (Fig 4).

Response: We agree and have renamed this section to "Model and data limitations." We have also revised the terminology within the section to distinguish limitations associated with model formulation, unresolved physical processes, and input data from the statistical uncertainty of the fitted return levels shown in Fig. 4.

Line 408: Some introductory words would be nice, to provide a motivation for the assessment. You could mention the three types of uncertainty that you assess: (1.) model coupling impacts, (2.) morpho-dynamic changes and (3.) resolution of the atmospheric forcing.

Response: We agree and have added an introductory paragraph explaining the motivation and structure of the section.

Line 419: "h" has not been introduced as the water depth. You can maybe mention that the wave breaking limit is $H_s/h=0.78$. This however is only a local criterion. Sheltering coastlines and shallow water areas at upstream locations are not considered.

Response: We have defined h as local water depth in the text.

Line 424: Can you provide a reference for the impacts of wave-ocean circulation model coupling?

Response: We have strengthened the referencing of this discussion. Kanarik et al. (2021) is now cited directly after the statement describing the effect of modelled currents on Baltic Sea wave conditions. We have also added Haapaniemi et al. (2026), which demonstrates that explicit wave forcing can modify ocean-side stress, stokes-coriolis forcing, turbulence, and upper-ocean mixing. Together, these references support the discussion of both directions of wave–ocean interaction and the structural uncertainty associated with using uncoupled wave and circulation products.

Line 425: You mention wave model bias correction here, but I guess you mean the 15% scaling that was applied to account for the overprediction of extreme waves. This however is not a constant, mean bias which you remove. It reduces the extremes stronger than the medium-large waves.

Response: We agree. The previous term "wave-model bias correction" was imprecise because the applied adjustment was multiplicative rather than a constant additive bias removal. We have revised the terminology throughout the relevant discussion and now refer to it as a site-specific multiplicative adjustment. We also clarify in the Methods that significant wave heights were reduced by 15% and wave

periods by 5% proportionally, so that the absolute magnitude of the adjustment increases with the magnitude of the wave parameter.

Line 454 and 458-459: Could you make a reference to Fig.4.

Response: We have provided the references.

Line 466: I assume you mean the “regional-scale return levels of the uncorrected data set”. “uncorrected return levels” could be a bit ambiguous, meaning that you corrected the return levels itself earlier on.

Response: Corrected

Line 468-472: Could you go into bit more detail and explain what you mean with sensitivity analysis? Did you apply the correction factors derived from the site-specific assessment to the regional results for the Baltic Sea? But these correction factors are site-specific. They may vary between locations. I would provide these values as an example, not as a definite uncertainty estimate.

Response: We agree. The previous wording could be interpreted as if the correction factors derived at Paldiski had been applied to the Baltic-wide results. This was not the case. We have revised the text to clarify that the regional EVA uses the original uncorrected gridded datasets. The Paldiski adjustments of approximately 5–10% for wind speed, 15% for significant wave height, and 5% for wave period are now presented only as examples of the magnitude of model–observation differences at one coastal site, not as regional uncertainty estimates. We also clarify that model bias may vary spatially with exposure, bathymetry, coastal geometry, and observation-to-grid representativeness, and therefore requires site-specific evaluation.

Line 488-489: Do you mean that the regional EVA can be used during the siting process to identify suitable locations, or that it can be used for contingency planning by identifying safe locations to which the entire terminal could be towed in the event of an extreme storm?

Response: We agree that the previous wording was ambiguous. Our intention was not to suggest that the regional EVA could identify locations to which an FSRU or terminal could be towed during an extreme event. We have revised the text to clarify that the Baltic-wide EVA is intended as a regional hazard-screening tool for comparing the relative exposure of existing and prospective infrastructure areas. It is not intended for operational refuge planning or direct site selection, which would require additional analyses of persistence, directionality, navigation constraints, accessibility and engineering requirements.

Figure 4: I think 12 panels in one figure is a bit much. This means that each panel will be rather small. I would recommend to split this figure into several figures.

Response: We would like to keep the figure as it is. In the published version by clicking to the figure it appears separately.

Conclusions:

Line 500-501: The scientific contribution mentioned here includes 2 points, (a.) the harmonization of longterm input data sets, which has been carried out for site-specific assessments and could have been

described in more detail in this paper (instead of in the supplements) and (b.) the cross-validation of these data sets, which has only been carried by comparing observed and modelled distribution functions. I would call the validation study a statistical or distribution-based validation study, to distinguish them for a full validation study.

Response: We agree and have revised the terminology to avoid implying that the study constitutes a full validation exercise. Comparisons with observations and independent model products are now described as distribution-based evaluation rather than cross-validation. We have also added a concise description in the Data and Methods section of the main harmonization steps used for the site-specific assessment, including alignment to a common sea-level datum, temporal matching for direct comparisons, and observation-based wind and wave adjustments. These preprocessing steps are also summarized in Table A2. We further clarify that the Baltic-wide assessment retains the original gridded datasets to preserve spatial consistency. The Conclusion has been revised accordingly.

We have also added the copyright statements to the Figure 1 caption, as requested by the editor. If preferred, these statements could instead be placed in a dedicated Copyright notice, Acknowledgements, or Disclaimer section at the end of the manuscript.

We have also carefully revised product numbering.