

Review of “Metoccean study for the planning of the FSRU terminal: a case study at the Estonian coast of the Gulf of Finland”

General Assessment:

The manuscript investigates metoccean 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 metoccean 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 metoccean studies the site of the FSRU terminal, but that reginal 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 metoccean 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 Pakrineene/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.

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.

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 metoccean 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.

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.

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.

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.

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.

Minor comments

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

Chapter 1. Introduction

Line 41: “to make informed decisions”

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

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

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

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

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

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.

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.

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.

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.

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?

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.

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.

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

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

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

Chapter 2. Data and methods:

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

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

Line 157-158: Could you provide a few more details on the preprocessing of the wind fields for the site-specific 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.

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.

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.

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.

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

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?

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.

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.

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.

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

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

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

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.

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.

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.

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?

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.

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

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.

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.

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.

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).

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.

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

Line 360: “Our analysis shows”

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

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

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.

Line 387: “The uncertainties ...”

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.

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

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

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.

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).

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.

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.

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

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.

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

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.

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.

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?

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.

Conclusions:

Line 500-501: The scientific contribution mentioned here includes 2 points, (a.) the harmonization of long-term 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.