Earth Syst. Sci. Data Discuss., referee comment RC2
https://doi.org/10.5194/essd-2021-45-RC2, 2021

Comment on essd-2021-45

Anonymous Referee #2

Referee comment on "An integrated marine data collection for the German Bight – Part 2: Tides, salinity, and waves (1996–2015)" by Robert Hagen et al., Earth Syst. Sci. Data Discuss., https://doi.org/10.5194/essd-2021-45-RC2, 2021

The study provides results from a 20-year hydrodynamic and wave hindcast for the German Bight (North Sea), their validation, and integration into the EasyGSH products. Compared to earlier products the main novelty of these hindcasts are the use of time-dependent (annual) bathymetric data that can be expected to significantly influence the results in the shallow waters of the region, and the coupling of the hydrodynamic with one of the two wave models used. In principle, this makes the data set extremely interesting and valuable. The topic is perfectly suited for this journal and I in principle recommend publication.

However, from my perspective, the paper is still in a preliminary state and I think, there are a couple of major issues that need to be addressed before publication.

- Compared to earlier hindcasts this study uses time-dependent (annual) bathymetric data, which represents one of the main novelties. At the same time, statements are made that errors in earlier years are caused by uncertainties in earlier bathymetric data. This left me with two questions: First, how have these data been derived? Second, is there value-added compared to simulations with constant bathymetry? More specifically, I suggest that the authors

  (a) introduce some description of the bathymetric data use, for example, the extent to which the model domain was surveyed every year or how much time interpolation was done at individual grid points, or how good different surveys match spatially and whether or not interpolation was needed. Some Figures to illustrate the issues would be appreciated.

  (b) provide a comparison of results with a run with constant bathymetry to assess the value added by their approach. I am fully aware that the hindcast could not be fully repeated, but some strategy, e.g. to assess the sensitivity for average or extreme conditions would be appreciated.
- The domain of the hydrodynamic model basically covers the North Sea and the authors briefly mention that some calibration was made to account for surges at the open boundaries. Some more details are needed, for example, what data were used and how

was the calibration made? Can this set-up account for the effects of external surges? What are the impacts on extremes?

- The validation strategy appears a little limited. While it somehow depends on how the data will be used (the manuscript is not very specific here) I feel that extreme events may play a role. Most of the validation concentrates on average errors, some specific validation for extremes would be very beneficial.
- I had difficulties following the argumentation in section 2.1. From my perspective, this section requires major re-writing. The envisaged intention appears not to be in-line with the discussion. Figures 1 and 2 need better description and the general strategy needs to be explained. For example, the text discusses web and GIS-based applications that are not apparent in the scheme in Figure 1. I was more or less lost reading this section. Are the model outputs part of the data set or only the analyzed products? What were the criteria for selection of specific analyses (user request, other, ...), etc.
- The experimental design is not fully clear. The authors should discuss and justify the use of two different wave models. I understand that one was coupled to the hydrodynamic model but why then the other was needed? If there is a reason this should be elaborated.
- There is something wrong with the error metrics (equation #5 in the appendix). This is (a) not a typical definition of an index of agreement and (b) if a correlation is meant both sides of the equation do not match. Maybe this is just a typo, but please check potential consequences for analyses in the text. Also, there is no equation #4.
- The paper mentioned several times the involvement of stakeholders in the development of their product but their role or the benefit for the overall product is not addressed.
- The paper mentions several times that it is the second part of a two-part publication. I suggest that the authors at the beginning briefly introduce the overall concept and how work/results were split into different parts. This would make the strategy and content easier to assess. The parts where this statement is repeated could then be removed.

Minor comments

- In the title, what is the meaning of the CE in "2015 CE"?
- The abstract should be re-written. It should not contain references and should describe the work presented in this manuscript.
- Line 27: To me, the phrase "in a contested region" sounds strange in this context. I suggest replacing with, for example, "… region with conflicting or competing interests …"
- Line 34: What exactly is "short-term" with the respect to the previous sentence?
- Line 37: "For this reason …". I don't think that is the reason.
- Line 40: "Individual tides" Please be specific.
- Lines 41-45: At least the hindcast described in Weisse and Pluess is available at a higher resolution than mentioned here.
- Lines 47-51: There are different parameters mixed-up (e.g., resolution of tidal and bathymetric data sets are compared). Please revise.
- Figure 2: The abbreviations EEZ and AWZ are mixed up.
- Section 2.2: What is the vertical resolution of the hydrodynamic model? Please add. When currents are discussed, are depth-averaged currents used? Please be specific.
- Line 182: I assume that the meteorological forcing is not communicated between these two models. Please revise. What are the wave effects on hydrodynamics accounted for in the coupling?
- Line 223: Salinity is shown, not salt transport.
- Line 290: Not the model validity is shown, only the extent to which model parameters considered agree with observations.
- Section 4.1: While M2 is the dominant constituent, shallow water effects are probably

important in the model domain. A closer look at other components such as M4 would add value to the discussion.

- Line 320 “This led to worse error metrics …” I don’t understand the argument.
- Line 339: Shifts in tide-gauge locations should be documented.
- Discussion of Figures 10 and 11: This implies that most of the bias is due to errors in tidal low waters. This should be mentioned explicitly.
- Figure 13: The slope needs a unit. Discussion of R-squared needs to be clarified (see major comment #6)
- Table 2 and the discussion in the text: Is Tm02 or mean wave period discussed? Also, how useful are the error statistics for the peak period? Frequencies are discretized differently in both models and values can only take discrete numbers.
- Line 478 “relevant forcing parameters” Forcing of what?
- Line 484 “It should be mentioned … it is hoped …”. This is very vague. Is it possible to be more specific?
- Line 491: Which part or applications of sea level science are meant?
- Line 491 “Thus, we propose …” What exactly means proposed? Is there an ongoing attempt? If so, please specify.
- There are several typos or uncommon phrases. I suggest that the manuscript could benefit from some language editing.