

# Review report for “Incorporating historical information to improve extreme sea level estimates”

## 1 Summary

The article by MacPherson et al. presents a method to incorporate historical Extreme Water Level (ESL) information into classical Extreme Value Analysis (EVA) using a Bootstrapping approach. They demonstrated the method for Travemünde, Germany and indicated to a possible underestimation of the current design water level for the flood defense.

The analysis of this paper is generally well-written. However, I find that the discussion is very site-specific, and many potential avenues for application of such a method in other places are not touched.

I recommend acceptance of this article after clearing-up some of the details - particularly regarding the historical data - which is essentially the main attraction here. My comments are given below. LXX means the corresponding line number.

## 2 Major comments

### 2.1 On the study area

- A study area map is needed for the ease of the readers.
- I think Section 3.1 would be a easier read if the datums are first identified. For example, please consider indicating NHN in-terms of mean sea level (MSL) and the corresponding value of HW200, HW200+50cm etc in terms of MSL. Is NHN is equal to mean sea level?
- I think it is worth mentioning in the text that the tide condition, e.g., the fact that the low tidal range at Travemunde allows directly looking into the ESLs. I think, it would not be so straightforward without this micro-tidal setting. L203.

### 2.2 Historical ESL events

- L130-145 discusses the approach of incorporating the historical ESL, and gives a background. The “main issue” of incorporating such data is identified to be the fact that historical measurements are isolated data points, and not having the duration of the observation defined. While these are true, I think one other major element is not discussed here, or other places - is the consistency of the data itself (in terms of datum). When stretched backwards, even with systematic data - such as tide gauge - datum consistency can get quite tricky. As you have already discussed before, presence of a large event can have significant impact on the EVA, I believe this point needs further attention in the article. Essentially, I believe this might have been discussed in the chapter (?) by Jensen et al.
- Table 1. Please consider a indicating for each storm where they are sourced from.
- Which value of 1320 is taken? 3.10 or 3.20? As it is a large value, what do you expect in terms of uncertainty? More important question is related to point 1 - how comparable is this measurement in-terms of datum/height accuracy?
- L190: From a quick look at Jensen et al., for 1694 event there is a 11cm difference in ESL between Lubeck, and Travemunde. For the events, where Lubeck value is taken for the sake of extending the series, could it induce another set of bias? To put it differently, are the sea levels comparable between the two sites?
- I did not find the 1304 event in the table, which is shown in Figure 1. In Jensen et al. this event exists but no date nor height is reported.
- Is it necessary to consider the VLM corrections in the historical series? Particularly given that we are taking values from 700 years ago. As in L211, a detrending is done for the systematic data. It would definitely create a question regarding if you need to apply some corrections to the historical series too.

### 2.3 Comparison with maximum likelihood approach

- Among the 3 methods of DWA (2012), only method 2 seems to be compared to. If you do not consider the other methods to be tested, please remove them from the description to make it lean. Without multiple comparison, L15 needs a revision - consider changing from "... outperforms other commonly used approaches." to "... outperforms currently used approach for Travemunde."

### 3 Minor comments

- L18-19: please consider adding how much larger ESL estimates (in percentage?)
- L28: "will increase" → "projected to increase"
- L94: Could you please provide a more accessible literature on the inclusion of "Climate Surcharge", instead of these two German refs (reports, I presume)?
- L115: missing ref to Pickands-Balkema-de Hann theorem.
- L151: what does "int" mean?
- L296: what is the value of observed frequency?
- Please be consistent on the definition of the level (m above MSL) throughout the MS.
- L404: could this be cross-referenced to previous studies? (e.g., Catalog by Jensen et al.).
- L414: please consider replacing "significantly" → "substantially", as there is an expectation to see a significance test when the term significant is used in the context of probabilistic analysis.
- L416: ... 224 cm... please see point on the consistent identification of the datum.
- L424: are the floodable regions is under insurance-schemes? Could be a good addition their interest on such results.