Reply to Reviewer 2

We thank the reviewer for his/her detailed comments on our manuscript. In the following, the original reviewer comments are given in italics and our point-by-point responses to the reviewer's comments in roman font with planned changes to the text put into quotation marks.

The future "climate" of water levels is one of the core problems for low-lying areas. The manuscript addresses this problem by means of advanced statistical modeling of parameters of extreme value distributions for future water levels and a sort of ensemble projection of extreme water levels and their return periods.

The analysis is theoretically sound, relies on high-quality data sets, has been performed professionally and leads to an interesting set of results. The presentation is clear and well structured, uses correct English and brings enough details for understanding the material.

General comments:

I am thus generally happy to recommend the manuscript for publication.

Before sending to print, however, I recommend to expand the presentation a little bit to cover some aspects that may mislead inexperienced readers and to make a few adjustments that would make the interpretation more exact and the message clearer. The recommended changes and additions only address single wording features and interpretation aspects (most of which are technically acceptable as presented in the manuscript) and do not involve any large changes to the presentation.

A potential trap for some readers may be the interpretation of the limited set of arguments of the Weibull distribution. Even though the authors mention on lines 307–308 that the shown values [of the upper threshold for the argument of the reverse 3D Weibull distribution] should not be interpreted as actual limits for the sea level, I would recommend commenting the related aspects in more detail to make the situation clear. There are two aspects worth of mentioning.

Firstly, the limited region of validity of the 3-parameter Weibull distribution could be interpreted differently. On the one hand, there is a temptation to think that this distribution provides the final truth about some properties of the described processes. On the other hand, the existence of this kind of threshold is not really physical and could be interpreted as showing that the entire GEV approach loses its validity near and behind this threshold.

Secondly, the set of block maxima may contain elements of different water level "populations" of the Baltic Sea. The reason is the well-known property of the Baltic Sea:

its water volume may increase or decrease considerably for several weeks by water exchange through the Danish straits. The "population" of the background water level of the Baltic Sea roughly follows a Gaussian distribution whereas the local storm-driven surges roughly follow an exponential distribution [Soomere, T., Eelsalu, M., Kurkin, A., Rybin, A., 2015. Separation of the Baltic Sea water level into daily and multi-weekly components. Continental Shelf Research, 103, 23–32, doi: 10.1016/j.csr.2015.04.018]. It may thus easily happen that the block (annual) maxima do not necessarily come from the same distribution. In this case the GEV distribution is just a passable approximation of the distribution of the block maxima and nothing more. It may easily be that the large scatter of the threshold of question is a reflection of this feature.

In this sense it is better to remove the conjecture "From theoretical perspective, this suggests that there might be an upper limit that the sea level extremes can reach along the Finnish coast" on lines 369–370 from the manuscript and also to modify the sentence "This also suggests that the hierarchical models can be used to estimate theoretical upper limits of the extremes of short-term sea level variations along the Finnish coast" to make sure that the unexperienced readers are not mislead.

We had similar concerns about providing quantitative estimates for the theoretical upper limit in the manuscript. It indeed might give a false impression for the reader that this limit would be an actual physical upper boundary for annual sea level maxima in the study region, which it certainly is not. Thus, we have removed the comment about the theoretical upper limit from the Conclusions section and rephrased the similar sentence in the abstract to make it more neutral.

We will also comment on the alternative approach to using GEV distribution as a model for sea level extremes in the Discussion section as follows:

"We have used GEV distribution to model the overall annual sea level maxima (after removing long-term trends). However, different processes contribute to variations in sea level in the Baltic Sea and cause it to fluctuate at different temporal scales. Therefore, alternative modeling approaches could be considered to model separately the different sea level fluctuations and their contributions. For example,Soomere et al. (2015) have used an approach in which separate statistical models were fitted to weekly-scale and local storm-surge driven sea level fluctuations".

Specific comments:

The Abstract seems too long, e.g., the sentence on lines 3–5 could be removed without any loss to the message and the material on lines 11–13 could be made more compact and smooth.

After considering this comment, we slightly modified lines 11–13, but decided to keep lines 3–5 as they were. We felt that otherwise the abstract would have lost part of the motivation for the hierarchical modeling approach.

Line 21: probably "associated WITH" or similar.

Corrected.

Lines 23–24: even though the increase in the mean sea level has exceeded the global average during the past 50 years in the Baltic Sea in many locations, there are opposite examples, e.g., the sea level on the Latvian shores [Männikus, R., Soomere, T., Viška, M. 2020. Variations in the mean, seasonal and extreme water level on the Latvian coast, the eastern Baltic Sea, during 1961–2018. Estuarine Coastal and Shelf Science, 245, Art. No. 106827, https://doi.org/10.1016/j.ecss.2020.106827]. This feature very shortly reflected in (Weisse et al., 2021) and may easily be overlooked. Also, it seems to have local character.

We will include a remark about this in the same location:

"... with some local exceptions from this trend (e.g., Männikus et al., 2020)."

Line 33: it is recommended to insert a reference to the analysis of meteotsunamis in the study area even though such a reference appears later.

Done.

Lines 38–39: while piling up water in the ends of the Bay of Bothnia and Gulf of Finland for sure is one of the main reasons for very high water level in these locations, the role of piling and emptying the entire subbasin is probably minor there compared to harbor-type oscillations. See, for, example [Jonsson, B., Döös, K., Nycander, J., Lundberg, P. 2008. Standing waves in the Gulf of Finland and their relationship to the basin-wide Baltic seiches. Journal of Geophysical Research-Oceans, 113 (C3), C03004, doi: 10.1029/2006JC003862]. Still, this effect seems to be a decisive one in some other basins, such as the Gulf of Riga [Männikus, R., Soomere, T., Kudryavtseva, N. 2019. Identification of mechanisms that drive water level extremes from in situ measurements in the Gulf of Riga during 1961–2017. Continental Shelf Research, 182, 22–36, doi: 10.1016/j.csr.2019.05.014.].

We will slightly modify this sentence following the reviewer's comment:

“The largest sea level variations on the Finnish coastline take place in the ends of the Bay of Bothnia and Gulf of Finland due to the piling up effect and standing wave oscillations within the bay (Jönsson et al., 2008).”

Line 61: “However, they did not consider spatial dependencies explicitly in their analysis” is ambiguous and is better to be removed.

Done.

Line 80: consider replacing “extends” by “applies”.

Done.

Line 116: “which should reduce the correlation between the annual maximum values.” is of course correct but this operation most likely almost totally removes this correlation.

We will change “*which should reduce the correlation*” to “*which removes the correlation*” to better underline the removal of correlation between the annual maxima by this operation.

Line 139: What is the meaning of the plus sign at the end of square brackets?

The plus sign denotes that the function is defined only if the term within the square brackets is larger than zero.

Line 141: consider replacing “y has bounded upper tail” (that is mathematically nonsense for an argument) by perhaps a longer explanation that the GEV distribution function is only defined until a specific value of y which is often associated with the theoretical maximum or minimum value of the process under consideration.

We prefer to keep the explanation short but will replace “y has bounded upper tail” with “y has an upper limit”.

Line 173 and in several locations below: the simple use of “Common” (or similar) makes reading fairly complicated. Consider using “The COMM version/approach/ model” etc., e.g., as on line 246.

Thank you for this comment. We discussed changing the names of the methods, but decided to keep the original naming convention for both the Common and Separate model.

Line 199: Do you have a specific reason for using norm when evaluating the expression in square brackets?

We will change the norm to parentheses. There is not any particular reason for using a norm in this case.

Line 234: consider treating "elpdLOO" as a variable, e.g., $elpd_{LOO}$ unless you have reasons for using the text mode. Anyway, unify the use of P_{LOO} in Table 1 and as text on line 243.

We have loosely followed the notation style used in Vehtari et al. (2017) (upright for elpd and italics for p) but will unify the use of p_{loo}. As a minor modification, we will change the subscripts and both symbols to lowercase.

Line 248: probably "location and scale PARAMETERS" are meant.

Corrected.

Line 253: it is not recommended to start the sentence from a symbol or expression.

Corrected.

Line 267: consider expanding the expression "Weibull-type distribution" towards explanation that the GEV approach uses so-called reversed 3-parameter Weibull distribution (and not, e.g., the 2-parameter Weibull distribution that is common in the description of wind speed, wave heights, etc.). Just to make clear the scene for the reader.

We have changed "Weibull-type" to "3-parameter Weibull" in this sentence to avoid any confusion with the 2-parameter Weibull distribution.

Line 269–270: "they used ocean model output instead of observations in their analysis" is only partially true. They also used measured data from five locations and noted strong variations in the shape parameter depending on both the particular location and the method for evaluation of the parameters of the GEV distribution.

This is true and will modify this paragraph as follows:

“...A slightly contrasting result was obtained by Soomere et al. (2018), who estimated shape parameter values close to zero from an ocean model output along the Estonian coast. Furthermore, in their additional analysis based on tide gauge observations the shape parameter varied strongly depending on the location and method used to estimate the GEV distribution parameters...”

Line 276: “separate fits”: see comment to line 173.

Please see our reply above.

Section 6 Conclusions: it is recommended to remove the short names of scenarios from the text in order to make the section readable on its own.

Done.

References:

Coles 2001/2004 is missing from the list

Thank you for noticing this! The reference was lost due to an unknown reason, when we compiled the pdf from the latex template.