The paper presents the compound flood risk analysis across the Swedish coast in the presence of low record availability and the choice of copula. While the uncertainty due to the first can't be averted, the second can be improved by the appropriate choice of copula and its parameter. Often sentences are not clear and require attention in framing. In a few cases, the methodology adopted is not robust and needs a relook. Often, there are misleading interpretations that make the paper weak. The paper can be published after appropriate revisions. The reviews are summarized as below:

Thank you for your review and many helpful comments to improve our study. We agree with your point that the uncertainty due to low record availability can't be averted directly, however this paper aims to argue that using different data sources can be highly important to better estimate uncertainties linked to the available datasets (as often the length of observations is short but modelled data have inherently uncertainties and biases). We hope our responses given below marked in red as well as our changes in the manuscript have helped to address these and other issues raised, also with regards to the interpretations and conclusions made.

Updated references for section 2.1.1 have been made due to final adjustments to the now published methodology presented in Dubois et al. (2024).

1. In Abstract, line 12: "The compound flood risks…. Often estimated using statistical copulas". This line can be misleading since copulas are one of the methods for estimating joint probability between two random variables. There are other methods as well, for example, joint entropy, or bivariate distributions considering box-cox transformations of associated random variables. Please consider revising/discarding this sentence.

Thank you for your comment, we have adjusted the manuscript accordingly (lines 12-13).

2. Line 27: What about the coastal backwater effects that influence the occurrence of compound flooding?

Thank you for your comment, this is actually what we have tried referring to when pointing out the relation to the storm surges effect and therefore clarified this point in adding a sentence about it and a reference (lines 30-32).

3. Please use the SI unit for sea level measurement.

Thank you for your comment, we do not see the issue of using cm as the sea level measurement unit as it is done in many other studies because the sea level variations are rather low (maximum around a few meters).

4. Line 82: Please use the word 'copula' throughout and not the 'statistical copula'.

Thank you for your comment, we agree and have adjusted the manuscript accordingly.

5. For processing 13-year sea level observation, a re-analysis coupled observational analysis was performed. In cases of data scarcity, the peak-over-threshold (POT) approach is in use instead of annual maxima. On the other hand, coupling different data sources, as adopted in this study, often results in

underestimation due to scale mismatch issues and extremes often underestimated in gridded reanalysis runs. If you are purely interested in observational assessment, the POT approach may be more powerful considering on average 2-3 to events per year, as compared to mixing reanalysis runs with the local tide gauge records.

As mentioned also to reviewer 1, we decided to only use the annual extremes for the analysis, which is a common approach in the literature, also to maintain the same method across the different datasets used for the paper for a better comparison basis. Concerning the sea level univariate brief analysis (as this is not the core of the paper but is rather used to introduce each dataset and assess differences independently between each of them), we did not mix any reanalysis data with tide gauge observations data. The reconstructed time series data is only based on observations (see reference to Dubois et al. (2024) that now got accepted for final publication). The reanalysis data has well-recognized issues that makes it difficult to work with at this level of detail. However, when it comes to carrying out the sensitivity analysis, we do not see this as a major issue as the sea level data do not seem to have a strong impact on the copula results and this issue even strengthens our conclusion that, for this case, hydrological data influence the results the most. We adjusted the manuscript to clarify this point (lines. 414-416).

However, for the copula analysis, we also carried out a brief analysis on defining extreme events at sea level above the 95th percentile value and another test using the threshold value of the 99th percentile. This analysis did not seem to make any difference in our conclusions as also found in Ward et al., 2018; but a more extended sensitivity analysis could be, we think, highly relevant. However, we believe this is outside the scope of this study. We also refer to this point in the section 4. Limitations.

On page 5, line 110-125: how you have converted hourly records to daily? The tide gauge records in Sweden are available at a minute-scale temporal resolution.

We converted downloaded hourly sea level tide gauge data from the SMHI download webpage to daily time series using the maximum hourly data within the day. We then adjusted accordingly the manuscript lines 115-116 & 126-127.

1. On Fig.3: lower panel, clearly shows that the reanalysis-driven reconstructed sea level observations are largely underestimated, especially at larger return period values. Please show the sea-level observation measurement in meters (SI unit).

We agree with you that, in terms of median values, the model-based and reanalysis datasets seem to largely underestimate return levels. However, it is important to keep in mind that the observations, as well, are associated with large uncertainties as displayed by the background colours and the background colours of the other sets include also the median RLs of the observations datasets (for more information, refer to Dubois et al., 2024). To keep consistency through the study, we rather would like to keep the sea level unit as cm.

2. between lines #145-150: What are the different sources of uncertainty of these models? Please describe number of parameters involve for calibration, forcing

data requirements and their temporal resolution. The predictive skills of the hydrologic models in simulating daily river discharge are not discussed at all.

Thank you for your comment, indeed we did not discuss in detail the hydrological models as that information can be found within the references that can be found within the manuscript. Some clarifying sentences have however been added to the manuscript to be clearer about where the interested reader can find evaluation and more details about the hydrological modelling (section 2.1.2).

3. Fig. 4: Y axis label: use superscript for the discharge measurement. Further, the uncertainty estimates between E-Hype and S-Hype model can be quantitatively estimated by the ratio of upper bound to the lower bound across higher and lower return levels.

Thank you for your suggestion, we adjusted the y-axis label. Indeed, quantitatively estimating the uncertainties estimates between E-Hype and S-Hype can be really interesting, but we believe that this is outside the scope of this specific study focusing on the sensitivity analysis of compound flood events analysed within a copula approach.

4. Line 177-178: Is it maximum likelihood based estimates of GEV parameters? This might be problematic for estimation of shape-parameters of GEV. Often a Bayesian estimate is proposed.

We indeed used the maximum likelihood to estimate GEV parameters (for the univariate study) and we agree a Bayesian estimate might be better. However, as also mentioned to reviewer 2, the univariate analysis is, we think, important in order to add clarity to the paper but not as a core part of the study. Therefore, we think it was not necessary to follow this analysis path and rather maintain it at the current level of simplicity.

5. Lines 203-206 and elsewhere: sentences are erroneous, please consider revising. Both ‘OR’ and ‘AND’ approaches are suitable for modelling joint effect: while the former consider a time offset, the later considers co-occurrence.

Both approaches used here correspond to the ones in Serinaldi (2015) where the “OR” approach accounts for both for a time offset where only one of the variables is high enough to create a bivariate occurrence hazardous but also accounts for a co-occurrence where both drivers are high enough to make a bivariate occurrence hazardous.

6. On page 11: line 245: highlights a ‘discrepancy’.

As here, the dash lines assume independence between both drivers (annual river discharge and corresponding sea level) and the full yellow lines assume dependence between both drivers using the copula; then the fact that both lines are not superposed highlight that both variables are dependent otherwise they would be superposed. To clarify this point, the sentence has been revised (line 278).

7. Line 251: One do not assign any probability density function to each copula rather derives copula-based joint PDF.

Thank you for your comment, we agree with you and have adjusted the manuscript accordingly.

8. On page 13: line 280 onwards – this section and the subsequent ones are very confusing, rather much simpler and statistically robust methods should be adopted. The Gaussian copulas are not good while considering highly skewed data as here. The best method to select copulas are to apply the minimum AIC criteria with small sample corrections (in presence of limited data availability) followed by an appropriate goodness-of-fit measure, such as application of resample-based Cramer von Mises goodness-of-fit statistics.

This section is central to the paper and looks at the sensitivity of the choice of copula and aims to highlight the importance and challenges of choosing the best copula as you are suggesting and therefore, we would argue that including a large set of different copulas is important for proper context and discussion. Not because we argue that Gaussian fits are proper in this context. The method you proposed here, based on the AIC criteria, is indeed the one we use to select the best copula (see section 2.2 and corresponding references in this section) and we therefore agree with your comment. To clarify this point, precision has been added (line 300). While adopting different sampling strategies can be highly relevant, we decided to keep the same sampling strategy based on annual maxima across each dataset for consistency. We also conducted a brief analysis based on sampling values above the 95th and 99th percentiles but this did not seem to impact the analysis. However, a deeper sensitivity analysis on the sampling method could be really interesting and confirm this point but we think this is outside the scope of this study (see previous answer to comment 5).

9. Lines 335-340: Please explain in terms of hazards.

Thank you for your comment, we hopefully clarified this point (lines 358-359).

10. Line 339: Coincidence of independence line versus copula-derived dependence PDF does not necessarily stress the hypothesis that river discharge predominates over high sea levels. The other way around can also be possible.

Your comment seems to result from a misunderstanding of our paragraph as here, we refer to „corresponding dashed and full lines across the sets“ and we do not state that independence line versus copula-derived dependence PDF coincide, rather the opposite actually (see previous paragraph). We regret this misunderstanding and have modified this section in the manuscript for this not to be carried forward to other potential readers. As we can see on Fig. 7, for example the sets E-Hype / rec Halmstad and the set E-Hype / pred Halmstad, their independence lines of both sets are almost superposing as well as their copula-derived dependence PDF. And this, as it is also the case for the datasets obs Nissan / pred Halmstad and obs Nissan / rec Halmstad stresses that river discharge dominates over sea level inputs.

11. Line 346: What is the ‘most likely scenarios’ here?

Each time we refer to the ‚most likely scenario‘, we refer to the definition given in the methodology section 2.2 (this has been elaborated in the manuscript in lines 227-232) that we extended slightly to clarify those points. So here, we refer to the

scenario from the best copula fit (according to statistical criteria as AIC) with the highest density along the closed-form joint probability density function of the copula.

12. Line 390 and associated section: There are several uncertainties in return levels due to the incorrect and erroneous application of copulas. Please use an appropriate goodness-of-fit measure to select the best-fit distribution. Also, there is not enough evidence that the SL is least sensitive to compound flood hazards; – mere little shift in density contours does not justify this major finding.

As mentioned previously, we fully agree with this general approach and therefore did use the AIC methodology, as well as other statistical tests, to rank copulas (section 2.2) but we decided to keep all of them to highlight the importance of the choice of copula. Here, in this particular case, the study did not find significant dependency between sea level annual maxima and corresponding river discharge and we only found a significant dependency between annual river discharge and corresponding sea level (section 3, first paragraph & table A.1). Figure A.2 resume all of our results and Fig. 8 shows the sensitivity test. This sensitivity test on switching data sources is, we think, a strong enough evidence that, in our case, the sea level is least sensitive to compound flood hazards as, for one set of river discharge fixed, the results across the different associated sea level datasets do not change drastically (low NDV) compared to fixing a set of sea level data. For our studied site and area, we consider that our result and conclusion are robust. On other sites and regions, the results could however be different and we encourage that more studies may be needed and we have carefully checked the revised manuscript to hopefully reflect in a reasonable way that uncertainty and limitations exist, as is unavoidable in most studies. Please see also revisions related to other reviewer comments.

13. In section 4: first paragraph, what is the need of extreme sea level analysis using model-derived sea level observations? A purely observational assessment employing different sampling mechanisms can work too. In the second paragraph, the uncertainty resulting from the choice of copula can be constrained by adopting appropriate goodness-of-fit statistics for the selection of the best-fitting copula.

Thank you for your comment, good point; we have added clarification to this section on the limitations of our study. The need for extreme sea level analysis using model-derived sea level observations in this study was motivated by the short available time series at the station of interest (13 years) which, in the Extreme Value Theory would be associated with really high uncertainties resulting in difficulties to draw any reliable conclusion towards longer return periods. Also, the goal of this paper is to highlight the risk of using only one type of data sources which has inherent limitations as well as one potentially wrong copula as it has been seen in previous literature. Your point on adopting the appropriate framework to select the best-fitting criteria is indeed in agreement with our conclusion where we highlight the importance of the choice of copula.

Ward, P. J., Couasnon, A., Eilander, D., Haigh, I. D., Hendry, A., Muis, S., Veldkamp, T. I. E., Winsemius, H. C., and Wahl, T. (2018) Dependence between high sea-level and high river discharge increases flood hazard in global deltas and estuaries, Environmental Research Letters, 13(8), 084012. 10.1088/1748-9326/aad400.