Response to reviewer 1 comments:

Comment 1:

Specific comment 37

L 509, Why do you use the values maximising the dependence? Understanding when the dependence is maximised provides interesting information on the physical system, however, the dependence values that are relevant from a point of view of the impact is that between the variables at the same time. In fact, the storm tide and the river flow interact at the same time in the real world.

Response:

This is because in Method 3 the information on the temporal dynamics (i.e. relative timing) of storm surges and astronomical tides is discarded and only the peaks of flood drivers are considered via the use of a static tail water level, as discussed in section 2.4. This is one of the limitations of Approach 3 and thus, Method 3. The following statement has been added to improve clarity:
"This is because in this method the information on the temporal dynamics of storm surges and astronomical tides is discarded and only the only the peaks of flood drivers and their joint dependence are considered, as discussed in section 2.4."

In practice and in general, could using the lag that maximise the dependence lead to importantly overestimate the risk? If so (as I believe), such a potential overestimation should at least be stressed explicitly (in addition to noticing the related limitation) in the text.

Response:

This is a well-documented limitation of Approach 3, which is based on flood events. See (Zheng et al., 2015a; Zheng et al., 2015b; Zheng et al., 2013). As suggested by the reviewer, this limitation is mentioned at multiple locations in the manuscript, including section 2.4on Approach 3[1] and Section 5.4 on results comparison[2] and Section 6 on discussion[3].

---

[1] Line 300: "Despite these advantages, there are several simplifications involved in this approach when converting continuous meteorological data into a set of multivariate 'design events', which could lead to significant misspecification of flood probability if not taken into account. … During this process information on the temporal dynamics of storm surges and astronomical tides is discarded. … a significant difficulty arises when trying to align the timing of the storm surge and astronomical tide events with the timing of the flood-producing rainfall in the upstream catchments Santiago-Collazo, F. L., Bilskie, M. V., and Hagen, S. C.: A comprehensive review of compound inundation models in low-gradient coastal watersheds, Environmental Modelling & Software, 119, 166-181, 2019.. Indeed, this problem has not been resolved, with most current methods using a stochastic method to account for the temporal shape of surge peaks MacPherson, L. R., Arns, A., Dangendorf, S., Vafeidis, A. T., and Jensen, J.: A Stochastic Extreme Sea Level Model for the German Baltic Sea Coast, Journal of Geophysical Research: Oceans, 124, 2054-2071, 2019. or taking a simplified approach such as assuming 'static' lower boundary conditions rather than explicitly resolving the tidal dynamics Zheng, F., Leonard, M., and Westra, S.: Application of the design variable method to estimate coastal flood risk, Journal of Flood Risk Management, doi: 10.1111/jfr3.12180, 2015a. 2015a.. The extent to which this simplification leads to mis-specified flood risk (and whether this misspecification leads to an under- or over-estimation of probabilities) is not known."

[2] Line 629: "... However, in the joint probability zone (e.g. locations Sw10 and Sw12) where both flood drivers have a significant impact on resulting flood levels, the event-based Method 3 results in significantly higher flood levels for a given return period compared to Method 2. This is especially the case for location Sw12, where flood levels estimated using Method 3 are above the upper bound of the 95% confidence interval generated using Method 2 based on censored continuous simulation data. This over-estimation of flood levels for a given return period from Method 3 due to the use of a static tail water level and the associated assumption that the peaks of the two flood drivers with always concede can potentially lead to over-conservative estimation of flood risk and costly flood prevention infrastructure."

[3] Line 678: "... However, by translating continuous flood time series data into a set of 'flood events', the information on coincident timing between different flood drivers is often lost, and various simplifying assumptions often need to be made. For example, when implementing the design variable method (DVM), the tail water level is assumed to be static (i.e. no tidal dynamics) with a value that corresponds to the specified

Comment 2:

Specific comment 38

Fig 8, The 2D simulations receive as input time series of T and Q, therefore a question arises: which is the value of the time series that you consider as that to be reported on the x and y axes?
The plots, e.g., panel c, suggests that for a given 10 year return level of Q, when T becomes larger (from 0 to 1-year return period), H decreases. This is physically inconsistent. Such inconsistent behaviours seem to occur in the range of T AND Q below 1-year return levels. Do you have an explanation for that? If the explanation is convincing, one would then consider not showing values in this bivariate range (up to 1-year return level for both variables).

Response:

This variation is potentially caused by the interpolation method used. Additional discussion has been added in the revised manuscript to explain this.
"It can also be observed in Figure 8 that there are some variations in estimates of flood levels with very short return periods (e.g. return periods of 1 in 1 year or below), with the increase in one flood driver leading to decreased compound flood levels. Careful inspection of the results shows that this feature does not apply to any of the simulated data points, in the sense that simulation points with larger values of the boundary conditions always yield larger flood levels. Rather, the 'inflection' only occurs in a sparsely sampled region of the plot, and is thus suggestive of the limitations of using a log-linear interpolation scheme in this region. This therefore highlights the importance of carefully considering the sampling scheme as part of the analysis."

My suggestion to editor and authors is to add some hatching in the part of the plot that is not considered trustable such to highlight the issue to the reader. As all my comments, this is in the interest of the authors given that some readers may focus on the image at first and then on the text; hence, the image could look odd to the reader as unphysical. If that is computationally expensive, I think that the non-trustable area should at least be highlighted in words also in the caption, where the author would refer to the text for further explanations.

The paper that I originally suggested (Bevacqua et al. (2020)) is now published at https://www.nature.com/articles/s43247-020-00044-z I believe that the work is relevant for some relevant statements that the author make about the changes in the dependencies, i.e. "This is particularly the case if one is able to assume that the dependencies between variables are either not greatly affected by climate change or that changes in dependencies produce second- order effects on flood probability compared to changes in the marginal distributions."

This is the only work available in the literature where changes in both marginal and dependencies of the meteorological drivers of compound flooding was considered for Australia. Therefore it would serve as a basis some for the statements and I would suggest considering it.

Response:

Thank you for the suggestion. The following wording has been added in the Caption:

*"*Note: the "inflection" in the contour lines for very short return periods is due to the use of interpolation scheme noting the sparsity of samples in these regions*."*

In addition, the recommended reference (Bevacqua et al., 2020) has been added in the revised manuscript.

---

exceedance probability. This simplifies the probability estimation process by assuming that the peak of tail water will always intercept with the peak of fluvial flood at any given location within the model domain, but it ignores the dynamic interactions of the two flood drivers, including the possibility that the peak fluvial flood wave will not occur at precisely the same time as the peak tidal cycle. Consequently, this method will always lead to over-estimation of flood levels Zheng, F., Leonard, M., and Westra, S.: Application of the design variable method to estimate coastal flood risk, Journal of Flood Risk Management, doi: 10.1111/jfr3.12180, 2015a. 2015a., as have been observed from results for the case study system."

Comment 3:

Specific comment 24

L 390, It is not clear to me why you need to account for the low water level periods through the resampling approach, given that you will fit the GPD only to the extremes. I understand that is necessary to be aware of the time in between the peaks to estimate the return periods, but why simulating it?

Response:

One important reason that flood data during low water level periods are also 'simulated' using the resampling approach is because the actual threshold values that will be used to fit the GPD is not known a priori. The resampling approach will provide a reasonable transition of flood levels between 'flood periods' and 'low water level periods' compared to just using zero values and makes sure reasonable flood level estimates will be used for flood probability estimation.

Thanks for the explanation, please also explain this in the paper if this is not done already.

Response:

Thank you for the suggestion. It has been added in the manuscript[4].

Comment 4:

Specific comment 12

L 185, During a discussion among colleagues, it was hypothesised that this may be related to the fact that often there is interest in measuring either the sea level or the river discharge and therefore no stations are collocated at the interface between the two. What do you think about this? Discuss it if you think that this is relevant. I guess that this appears also discussed/hypothesised in Paprotny et al. ("Compound flood potential in Europe").

Response:

Thank you for this suggestion, the following comment has been added.

"The lack of gauges within estuaries are likely to be at least in part due to the fact that there has historically been greater interest in measuring either the sea level or the river discharge and therefore there is less interest to place stations at the interface between the two (Paprotny et al., 2018)."

Sorry for that, but I realize that the paper that I suggested to cite here (Paprotny et al.) was not accepted for publication, so I am not sure whether the journal allows for citing it. You could refer to Bevacqua et al., 2017 (already cited in the paper) who also discuss the same issue.

Response:

Thank you for the suggestion. The reference has been updated.

[4] Line 442: "Since water level information below the selected threshold for fitting a GPD is censored in the frequency analysis, a resampling approach is used to fill in water level information during the low water level periods, which also addresses the challenging of not knowing *a priori* the exact value of the boundary condition thresholds."

Comment 5:

Specific comment 22

L368, Authors tend to oppose GPD and GEV as alternative approaches. Do you expect any differences in terms of uncertainties? Also, you use the GPD to estimate return periods/level. Shouldn't you also provide an equation for that?

Response:

The difference in the estimation outcomes from GPD vs GEV is out of the scope of this paper. The equation for the GPD is included in section 4.1.

I simply meant to add an equation for the return period based on the GPD (given that the GPD equation is provided). This seems not in the paper. Authors and editor can judge whether the reader would benefit from such an equation or not.

Response:

Thank you for the clarification. Estimation return period using a frequency analysis is a very common approach and there are many existing R/Python Libraries include functions that can do it. Personally, I like to reduce the number of equations in a paper, as too many equations can be a distraction. As a result, an additional equation on this is not added.

References:

MacPherson, L. R., Arns, A., Dangendorf, S., Vafeidis, A. T., and Jensen, J.: A Stochastic Extreme Sea Level Model for the German Baltic Sea Coast, Journal of Geophysical Research: Oceans, 124, 2054-2071, 2019.

Santiago-Collazo, F. L., Bilskie, M. V., and Hagen, S. C.: A comprehensive review of compound inundation models in low-gradient coastal watersheds, Environmental Modelling & Software, 119, 166-181, 2019.

Zheng, F., Leonard, M., and Westra, S.: Application of the design variable method to estimate coastal flood risk, Journal of Flood Risk Management, doi: 10.1111/jfr3.12180, 2015a. 2015a.

Zheng, F., Leonard, M., and Westra, S.: Efficient joint probability analysis of flood risk, Journal of Hydroinformatics, 17, 584-597, 2015b.

Zheng, F., Westra, S., and Sisson, S. A.: Quantifying the dependence between extreme rainfall and storm surge in the coastal zone, Journal of Hydrology, 505, 172-187, 2013.

Response to reviewer 2 comments:

Comment 1:

L64-L66: Slightly unclear. “experiencing long term changes” Are these the long-term changes caused by the aforementioned long-term climate phenomena. If not, please provide more details of their sources.

Response:

The statement here intended to refer to the fact that “the joint probability of flood drivers” are changing over time, which has been reported in the two reference provided, i.e. Arns et al., 2020 and Bevacqua et al., 2019. To our knowledge these changes have not yet been formally attributed either to climate change or other processes, so it is not possible to provide a causal statement.

Comment 2:

L80: “Superposition on the astronomic tide”. The interaction of the surge and tide could be mentioned explicitly here.

Response:

Thank you for the suggestion. “i.e. the interaction of surge and tide,” has been added in the sentence.

Comment 3:

Figure 2 and following paragraphs text: The text describing the approach could be simplified (in parts) by referring to the pathways outlined in Figure. It seems strange that the pathways discussed in the Figure are not mentioned at all in the text.

Response:

The details of the pathways can be in the caption or the text or both. After careful consideration, we decided to keep the description of the pathways in the caption, as there will be readers who may look at the figure without reading the text in detail. The reason we did not repeat the details of the pathways in the text is to avoid duplication. In addition, readers’ understanding of the main text will not be compromised without these additional details in the text.

Comment 4:

L296: “two components mentioned above”. The two components being referred to are not immediately obvious and should be stated. I assume you are referring to the probability/statistical modeling and hydrodynamic modeling.

Response:

Thank you for this observation. The sentence has been changed to “the two components indicated above (i.e. flood surface and associated probability)”.

Comment 5:

Figure 3: Avon basin in the caption but Swan-Avon basin on the Map. Be consistent!

Response:

Thank you for this observation. It has been changed to “Swan-Avon basin”.

Comment 6:

L412-413: “as only values above certain high thresholds are full accounted for”. Is it not periods where there is at least one value about the high thresholds rather than only values above a certain threshold? Also “fully accounted for” is very vague please be more specific.

Response:

Thank you for this observation. It has been changed to “as only values above certain high thresholds need to be included as part of the joint probability calculation”.

Comment 7:

L423: I think “both” rather than “either” would be more accurate.

Response:

There are situations where floods can be caused by one of the two drivers. Therefore, it should be “either”.

Comment 8:

Figure 5: Please provide a more detailed explanation of how the ‘high water level periods’ are defined. For instance, is there always a gap between the end of the initial buffer period and the first exceedance of T during an event, why? How long is the gap? Does the post event buffer always begin once Q falls below its threshold?

Response:

The definition of “high water level period” is provided in the second paragraph of section 4.2 just above Figure 5:

“The combination of the flood periods and the time buffer periods is referred to as the high water level periods, ”

The reason a buffer period is used is because the threshold values for both flood drivers are not known a priori, as mentioned at the beginning of the second paragraph in section 4.2. By extending the simulation period of floods, we increase the likelihood that all water levels above the required threshold values (which will be estimated after all simulations are complete) are simulated.

Comment 9:

L607: Grammar. Remove “relatively”.

Response:

Thank you for the observation. It has been removed.

Comment 10:

L672: Dependence estimation does not include marginal distribution estimation. I would therefore change dependence estimation to “joint probability estimation” or similar.

Response:

“Dependence” is the terminology used in most relevant studies on the third approach, which is the focus of discussion at this location. Therefore, “dependence estimation” is used to be consistent with previous studies.

Comment 11:

L687: “2) the dependence between the two flood drivers is location specific” I am not sure how this is more of a disadvantage with this method compared with the other two approaches.

Response:

Out of all three approaches, only in the third approach the dependence is considered separately. Since the dependence is location specific and requires to be estimated using an appropriate statistical model, it is included here as a challenge for Approach 3 for the completeness of the statement. The statement has been revised as below to make it clearer: “… and 2) the dependence between the two flood drivers is location specific and needs to be estimated using an appropriate statistical model (Zheng et al., 2015a)”