Review of the revised paper: "Estimating the Probability of Compound Floods in Estuarine Regions Wu" by Wu et al.

I thank the author for considering my comments. Overall, I found the paper improved and reccomend it for pubblicaiton. However, I have some minor additional technical comments.

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 depndence lead to importantly overestimate the risk? If so (as I believe), such a potential overestimation should at least be stressed explicitely (in addition to noticing the related limitation) in the text.

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, Ithink that the non trustable area should at least be highlighted in words also in the caption, where the author would refer to te 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.

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.

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 publicaiton, 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.

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 equaition or not.

Best regards.