

The manuscript titled 'Multivariate statistical modelling of extreme coastal water levels and the effect of climate variability: a case study in the Netherlands' presents an interesting attempt to quantify the joint probability of coastal water levels and precipitation that ultimately create a compound flood hazard. They use 800 synthetic annual maxima events to define the marginals of copulas and create a trained impact function to relate predictands and predictors. The degree of uncertainty introduced by shorter records is also quantified as a commentary on the importance of data volume for such methodologies. While I do think the publication will ultimately be a quality contribution to the literature, it currently is vague on some methodology components that need further clarification. I recommend that the manuscript be returned to address the following thoughts:

- In general, there are a lot of references to figures in the supplemental information that feel as if they are written in the same manner that one would normally refer to an in-text figure. If showing these figures are crucial to communicating the results, then I feel they should be in the main paper. Otherwise I suggest rewriting the sections (i.e. 4.1.2, 4.1.3, 4.1.4, etc.) to explain the results in words without referencing a take-away point that a reader would need to see a figure to understand. You can then tell the reader that further information is available in the supplement.
- The introductory paragraph refers to the same author/lab groups efforts in 5 straight individual sentences. While subsequent paragraphs show the author's have a broad grasp on literature beyond this one lineage, I recommend broadening the background to highlight that the motivation for this work does not arise simply from one group's efforts. There are many other works that have identified and attempted to account for multivariate climate drivers of compounding events (e.g. Anderson et al. 2019, Serafin et al. 2014, Rueda et al. 2016).
- I'll admit I am confused by the tidal variability included in Figure 1. The text at Line 114 indicates that the tide cycle is added but doesn't give any specifics (I suggest adding these specifics to improve transparency). Figure 1 makes it look like all 800 events had the maximum occur at the same phase of the tide? Otherwise the bold tidal level would be a more flat line with a large envelope of variability around it? If the events do all occur at the same tidal phase then that would be a significant limitation of this work. Perhaps Figure 1 is only a single example taken from the 800 annual and the text caption for the Figure could be rewritten to prevent the interpretation that it is derived from all 800 scenarios.
- Are copulas fit to purely empirical distributions? At which point the underlying assumption is that the 800 events can accurately represent the tails of the distributions? If this is the assumption being made then I think it should be explicitly stated in the manuscript and acknowledged as a potential limitation for obtaining extremes.
- Although a paragraph at the beginning of Section 2 does describe the study site, I think an annotated figure of the coast, the physical point where all data is obtained, and the square area or arial footprint of the watershed catching the precipitation could aid the

manuscript. I was left wondering about the coastal configuration, proximity to open water, and proximity to human altered landscapes.

- I think the paragraph between lines 112-120 could grow to be multiple paragraphs that detail the methodology from van den Hurk et al. (2015), as this manuscript is heavily dependent on that work.
- Although the explanation of copulas is suitable for publication, I think that the author's dynamical interpretation of the final copulas could be useful. By that I mean, why does a Frank copula fit better and what does that tell us about the dynamics of the compound hazard?
- I think the usefulness of this case study to readers may be improved by including a commentary on what physical processes are or are not being wrapped up into the relatively broad predictors. Does the original modeling framework exhibit sea level anomalies at longer frequencies than just meteorological surges and tides (i.e. monthly or seasonal anomalies)? Does the location of this virtual tide gauge experience waves? Perhaps a paragraph at the end of the discussion could address limitations and how extensible the study is to other sites.