Reviewer 3:

I would like to compliment the authors for a well-written paper. The research gap this paper aims to fill is clear and relevant. Here are some minor comments to consider.

The authors would like to thank the Reviewer for their thoughtful comments. Below, we provide our responses to each comment and outline how we plan to adjust the manuscript following the NHESS review process.

- Consider adjusting the title to speak to a larger audience. Currently, the term 'mixed populations' does not speak for itself. Instead, it would be good to focus on the fact that you look at different storm types that can drive compound flooding.
    1. As a follow-up, the term 'populations' is never clearly defined.

Agree and the title will be changed as:

"A multivariate statistical framework for mixed storm types in compound flood analysis"

For the follow-up question: the sentence is revised from L54 to L55 as follows:

"However, most of these studies assume that all extreme events originate from a single population, which refers to a set of events or observations that share common characteristics and are generated by similar underlying processes"

- As RC1 and RC2 have highlighted there are some limitations and uncertainties in the methods that you have used. In the paper, I am missing a proper discussion of these limitations and how to overcome them in the discussion/conclusion.

Thank you for the comment. Two paragraphs are added to the end of the discussion, explaining the main limitations on L451 (of the original manuscript).

"One limitation of the proposed framework is the identification of compound events based on extreme flood drivers. In some locations, none of the flood drivers need to be extreme to cause compound flooding, as geographical exposure, and other factors (e.g., elevation, drainage, permeability) play dominant roles. Therefore, the focus here is on assessing the compound flood potential and how the joint probabilities of different flood drivers are linked to various storm types. To extend the proposed framework for fully characterizing the compound flood risk, the statistical approach can be combined with hydrodynamic numerical models (so called hybrid modeling, e.g., Moftakhari et al. 2019) to estimate flood inundation. However, analyzing only the most likely event (even though it may be the most plausible given the observations) does not capture the range of flood levels that could be generated by different combinations of flood drivers (NTR and RF) along an isoline. One way to address this limitation is to sample an ensemble of events (peak NTR-RF combinations) along the isoline and run them through flood models. Alternatively, a response-based approach can be employed, which involves simulating flood hazard for a large number of synthetic events from the multivariate statical model and then performing the statistical analysis on the response variable of interest (e.g., flood depth at a given location). The latter is computationally demanding, possibly necessitating the use of a surrogate model, however the return level estimates are likely to be more robust than when adopting an event-based approach (Jane et al., 2022). The simulated probabilistic flood depths and extents can then be incorporated with exposure and vulnerability data to perform a comprehensive flood risk assessment.

Additionally, despite the long data records used (Gloucester City; 1901 to 2021 and St. Petersburg; 1948 to 2021), the stratified TC samples contain a relatively small number of events due to their rare occurrence in historical observations. When the framework is applied to a different site with less data, the smaller sample size may result in higher uncertainty in modeling the upper tail of the NTR and RF distributions. This limitation can be addressed by combining the proposed framework with synthetic flood driver information derived from physics-based models (e.g., Gori et al., 2020)."