

This manuscript presents the MSTM-TE framework for compound tropical cyclone-induced coastal flooding risk estimation. The study shows potential value for improving the modeling of extreme total water levels. However, there are still significant issues regarding data source characterization, validation of key modeling assumptions, physical consistency of the TWL formulation, justification of threshold selection in the extreme value analysis, fairness in benchmark model comparisons, completeness of quantitative performance assessment, and overall clarity and readability of figures and presentation. Therefore, I recommend a major revision.

1. Lines 42 – 50, 57 – 64: The manuscript states that traditional statistical frameworks implicitly force co-occurrence of component extremes. This statement is overly strong. It is recommended to distinguish between independent extreme value models, copula/conditional extremes approaches, and event-based dynamical models, rather than generalizing all conventional approaches as enforcing simultaneous peak occurrence.
2. Lines 107–112; Figure 1 and Figure 2: The study uses only four nearshore points to represent coastal flood risk in Guadeloupe, but the rationale for their selection is not clearly explained. The authors should clarify why these four locations are representative, whether they capture different exposure directions or coastal morphologies, and whether any sensitivity analysis of site selection has been performed.
3. In Equation 1,  $S_{sh}$  is described as including both tidal and meteorological surge components. However, later sections treat  $S_{sh}$  primarily as storm surge. It should be clarified whether  $S_{sh}$  includes astronomical tide, storm surge, mean sea level anomalies, or a specific water level reference, and whether different interpretations affect extreme value estimation.
4. In Equation 8, if  $\cos(\theta - D_p)$  becomes negative,  $H_0$  may also become negative, which would lead to non-physical or numerically invalid values in the run-up formulation.
5. Lines 344–347, Table 1: A threshold corresponding to the 0.6 quantile is relatively low and may include non-extreme values, potentially biasing tail estimation. It is recommended that the manuscript include diagnostic plots in the main text rather than only in supplementary materials.
6. Lines 169–174, 218–230, 312–325: Although the authors acknowledge that full independence does not hold and apply east–west clustering to mitigate dependence, the manuscript only states that significant correlations decrease sharply without providing

quantitative results. The authors should report correlation statistics,  $\tau$  distributions, p-values or FDR-adjusted results, and include key diagnostic figures.

7. The partitioning of storms based on whether the longitude of maximum Hs lies east or west of  $61.5^\circ\text{W}$  lacks a physical or statistical justification.