Referee Comment 2 (EGUSPHERE-2026-1899)

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.

Response: We thank the reviewer for the careful evaluation of our manuscript and for the constructive comments, as well as the positive assessment of our methodology. We will revise the manuscript accordingly to address the comments raised. Our detailed responses to each comment are provided below.

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.

Response: Thank you for this comment. We agree that the original statement was broad and should be further clarified by distinguishing among the different classes of existing approaches. We will revise and expand the Introduction (in particular, lines 49-50; lines 57-62) to clearly distinguish between (1) independent extreme value models, which neglect the dependence among contributing processes (Coles, 2001); (2) multivariate statistical approaches (e.g., copula-based and conditional extremes models), which characterize extremal dependence among event-level variables without explicitly preserving their temporal evolution (Salvadori et al., 2016); and (3) event-based hydrodynamic models, which preserve the physical evolution of storm events but are computationally demanding and not designed for statistical extrapolation to long return periods (Dietrich et al., 2011).

Importantly, we will also revise the corresponding discussion leading to the motivation of this study (lines 62-64) to clarify that the limitation regarding implicit co-occurrence of component extremes applies only to specific classes of statistical frameworks rather than to all conventional approaches. This clarification will improve the logical progression of the Introduction by more accurately identifying the methodological gap addressed in this study

study: namely, the need for a framework that combines statistical effieicny with the preservation of temporal storm evolution.

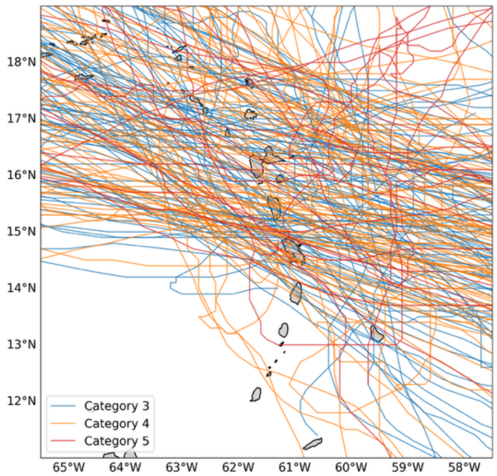
Additional references:

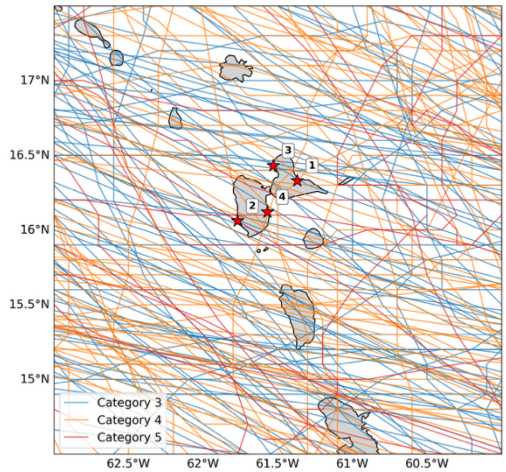
Dietrich, J. C., Westerink, J. J., Kennedy, A. B., Smith, J. M., Jensen, R. E., Zijlema, M., Holthuijsen, L. H., Dawson, C., Luettich, R. A., Powell, M. D., Cardone, V. J., Cox, A. T., Stone, G. W., Pourtaheri, H., Hope, M. E., Tanaka, S., Westerink, L. G., Westerink, H. J., and Cobell, Z.: Hurricane Gustav (2008) waves and storm surge: Hindcast, synoptic analysis, and validation in southern Louisiana, Mon. Weather Rev., 139, 2488–2522, https://doi.org/10.1175/2011MWR3611.1, 2011.

Salvadori, G., De Michele, C., Kottegoda, N. T., and Rosso, R.: Extremes in Nature: An Approach Using Copulas, Water Science and Technology Library, Vol. 56, Springer, Dordrecht, https://doi.org/10.1007/1-4020-4415-1, 2007.

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.

Response: Thank you for the comment. It should first be noted that the sites were not intended to provide a comprehensive representation of the entire Guadeloupe coastline, but rather to demonstrate the applicability of the proposed framework across representative coastal environments. Specifically, the four locations were selected because (1) they are adjacent to sandy beaches with significant tourism activities (beach names and locations indicated in lines 108-109), where coastal flooding and wave-induced hazards are of particular societal concern, and (2) they represent contrasting coastal settings, including both windward and leeward environments, which are exposed to different wave climates and tropical cyclone approaches. While these differences are partly illustrated by the wave rose diagrams (Figure 2), we will further support this rationale by adding supplementary figures showing the trajectories of Category 3-5 tropical cyclones relative to the study sites in a new Supplementary Section describing the dataset characteristics:





We acknowledge that a sensitivity analysis of site selection could be valuable when the objective is to quantify flood risk for the entire coastline or to identify representative locations for regional risk assessment. However, the objective of the present study is to demonstrate the proposed methodology to key sites exposed to cyclone-induced impacts, across several contrasting coastal environments rather than to estimate island-wide flood risk. Accordingly, a comprehensive site-selection sensitivity analysis is beyond the scope of the present work. We will revise the manuscript to expand the discussion of the site-selection rationale (lines 107-112), and explicitly discuss in Section 5.4 (Methodological implications and limitations) how a site-selection sensitivity analysis could be incorporated in future applications of the proposed methodology aimed at regional-scale flood risk assessment.

3. In Equation 1, Ssh is described as including both tidal and meteorological surge components. However, later sections treat Ssh primarily as storm surge. It should be clarified whether Ssh includes astronomical tide, storm surge, mean sea level anomalies, or a specific water level reference, and whether different interpretations affect extreme value estimation.

Response: Thank you for this valuable comment. In this study, Ssh represents only the atmospherically driven component of coastal water level obtained from the hydrodynamic simulations, i.e., the surge, and does not include astronomical tide nor mean sea level variability. We will revise the definition of Ssh in Section 2 (Equation 1; lines 119-121) to explicitly state this as follows:

"The sea surface height represents the atmospherically driven component of coastal water level at the shoreline (i.e., the surge)."

Accordingly, we will also revise the definition of the flood risk metric (lines 113-116) to clarify that the quantity analyzed in this study is computed from the atmospherically driven component of coastal water level and wave run-up, rater than from the complete coastal water level:

"To quantify the impact of TC-induced coastal flooding, this study focuses on a flood risk metric *FRM* that depends on the atmospherically driven component of coastal total water level (i.e., the surge) and the wave run-up as in Equation (1).

[..]

The additional components (e.g., astronomical tide nor mean sea level variability) of coastal total water levels are not included here due to the relatively lower amplitude in the region considered."

We will also clarify in the revised manuscript text that, in general, total coastal water level may comprise additional components (e.g., astronomical tide and wave-related contributions) by referring to specific studies, whereas only the atmospheric component is analyzed in this study.

Since the input dataset and all statistical analyses were performed using this atmospheric component only, this clarification does not affect the methodology, extreme value estimation, or the reported results.

4. In Equation 8, if cos(θ–D_p) becomes negative, *H0* may also become negative, which would lead to non-physical or numerically invalid values in the run-up formulation.

Response: Thank you for this comment. We acknowledge that Equation (8), as written, could imply non-physical negative values of the projected significant wave height when $\cos(\theta - D_p) < 0$. In the numerical implementation, however, negative projected wave heights were truncated to zero before calculating wave run-up. This truncation will be explicitly stated in Section 3.6.1 when the detail of the projection is provided (lines 265-267).

Since the present study focuses on extreme coastal flooding generated by shoreward-propagating storm waves, truncating offshore-directed waves (i.e., when $|\theta - D_p| > 90°$ and the projected wave height becomes negative) does not affect the estimation of extreme run-up or the resulting TWL return levels. Nevertheless, we acknowledge that this simplified projection does not explicitly account for more complex nearshore wave transformation processes. While the broader limitations associated with the analytical run-up paramterization are already discussed in Section 5.4, we will expand the discussion to explicitly state that the shore-normal projection is intended as a practical approximation within the present framework.

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.

Response: Thank you for the suggestion. While we agree that the threshold selection should be clearly justified, we respectfully believe that the threshold diagnostics plots are more appropriately presented in the Supplmentary Materials, as including these figures in the main manuscript would substantianlly increase its length without contributing directly to the principal findings of the study. To address this concern, we will instead strengthen the corresponding section in the main text by explicitly summarizing the findings of the diagnostics:

"Across all variables and both clusters, a marginal threshold corresponding to the $u_d = 0.6$ quantile was selected as it represented the highest threshold that provided stable parameter estimates; higher thresholds resulted in wider bootstrap confidence intervals and greater fluctuations in the estimated parameters. The full set of threshold diagnostics is provided in the Supplementary Material (see Figures S.2.1 and S.2.2)."

We believe that this additional explanation directly addresses the concern regarding potential bias.

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, τ distributions, p-values or FDR-adjusted results, and include key diagnostic figures.

Response: Thank you for the suggestion. We agree that the reduction in dependence resulting from the east-west clustering should be demonstrated quantitatively. We will therefore report, for each variable and threshold, the number and percentage of spatial locations (out of the 18,890 grid nodes) exhibiting significant Kendall's rank correlations ($p<0.05$) before and after clustering. The complete results will be provided in Section S.3 of the Supplementary Materials together with the corresponding spatial correlation maps. In addition, we will expand the discussion in the main manuscript in Section 4.1 (lines 322-325) to include a concise quantitative comparison highlighting representative results before and after clustering (i.e.,summarizing how the proportion of significant correlations is substantially reduced).

7. The partitioning of storms based on whether the longitude of maximum *Hs* lies east or west of 61.5°W lacks a physical or statistical justification.

Response: Thank you for this comment. The partition at 61.5°W was originally established by Sando et al. (2024) using the same synthetic tropical cyclone dataset analyzed in the present study. In that work, the boundary was motivated by evidence of mixed tropical cyclone populations over the broader Leeward Islands (i.e., bimodality shown in the distribution of MSTM values of wave (significant wave height; Hs) and wind (wind speed at 10 m; U10) variables) and physical considerations related to island-induced wave blocking and shadowing. More importantly, Kendall's rank-correlation diagnostics demonstrated that partitioning the storms substantially reduced the dependence between the MSTM and exposure sets, which was the primary motivation for introducing the clustering. Since the present study is built upon the same dataset and MSTM–TE framework, we retain this established partition and further verify its suitability through Kendall's rank-correlation analysis using the variables of our study. We will revise the explanation of the partitioning longitude (lines 320-321) to more clearly describe the rationale for the 61.5°W boundary by referencing Sando et al. (2024) and summarizing the physical basis for the partition. A more detailed explanation of the physical basis will also be added to the Supplementary Material.