Referee Comment 1 (EGUSPHERE-2026-1899)

This paper proposes the application of the MSTM-TE (Multivariate Spatio-Temporal Maxima with Temporal Exposure) framework to tropical cyclone-induced compound flooding risk assessment, with Guadeloupe as a case study, and systematically compares the statistical performance of trivariate and bivariate configurations. The research topic holds clear scientific significance, particularly in the practice of compound flooding risk assessment in tropical cyclone-prone regions. Overall, the research methodology framework of this paper is innovative, with the core contribution being the demonstration of the importance of preserving intra-storm temporal coherence for compound extreme value estimation. This manuscript fits within the scope of the journal and is worthy of consideration for publication after major revisions.

Response: We thank the reviewer for the encouraging assessment and constructive comments. We appreciate the recognition of the novelty and significance of our work. We will revise the manuscript accordingly, as detailed in the responses below.

1.Abstract Section: It is recommended to supplement quantitative descriptions of key results and quantify the core findings.

Response: Thank you for the suggestion. We agree that the abstract would benefit from a more quantitiative summary of the principal findings. Accordingly, we will revise the abstract to include representative numerical results that better convey the performance of the proposed framework and the significance of the main scientific findings by reporting (1) the reductions in bias and ensemble variance in return level estimation relative to the conventional approaches, as well as (2) the percentage of empirical and simulated storms exhibiting temporal alignment between the TWL peak and the compound wave-energy variable:

"Results show that MSTM-TE reduces bias and ensemble variance in TWL return-level estimation across all locations and modeling configurations, up to 88% and 79%, respectively, relative to conventional multivariate and location-specific approaches, highlighting the importance of preserving intra-storm temporal structure. The reconstructed storm evolutions reveal systematic temporal alignment between peaks in TWL and a compound wave-energy variable, observed in 96.3% of empirical storms and reproduced in 94.2% of simulated storms, indicating that wave-induced run-up dominates design-level flooding in the study domain."

2.Introduction Section: Further optimization is suggested regarding problem focus, logical coherence, and terminology transition.

Response: Thank you for this suggestion. We agree that the Introduction would benefit from a clearer problem focus and a more logical progression of ideas.

We will revise the Introduction in the following ways. First, we will expand the discussion of conventional methodologies (lines 49-50) by distinguishing among (1) independent extreme value models, which neglect the dependence among contributing processes (Coles, 2001); (2) multivariate statistical models (e.g., copula-based and conditional extremes models), which characterize the dependence structure among event-level extremes but generally describe storms through event-level maxima or threshold exceedances rather than their temporal evolution (Heffernan and Tawn, 2004; 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).

Second, following this discussion, we will incorporate recent developments that seek to incorporate temporal information into statistical extreme value analysis, including functional extreme value models (e.g., Gorse et al., 2026) and statistical models for the temporal evolution of extreme excursions (e.g., Tendijck et al., 2024), and clarify how these approaches extend beyond conventional methodologies while remaining distinct from the MSTM-TE framework.

Finally, we will revise the transition into the motivation for the use of MSTM-TE (lines 62-64) by explicitly highlighting the remaining methodological gap: existing statistical approaches effectively characterize extremal dependence, whereas process-based models preserve temporal evolution, but few frameworks simultaneously preserve both temporal coherence and statistical efficiency for extreme value analysis. This revised organization will provide a clearer progression from conventional approaches, to recent methodological developments, and finally to the motivation of this study's methodological framework.

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.

Gorse, N., Roustant, O., Rohmer, J., and Idier, D.: Simulation of extreme functionals in meteoceanic data: Application to surge evolution over tidal cycles, Adv. Stat. Clim. Meteorol. Oceanogr., 12, 123–148, https://doi.org/10.5194/ascmo-12-123-2026, 2026.

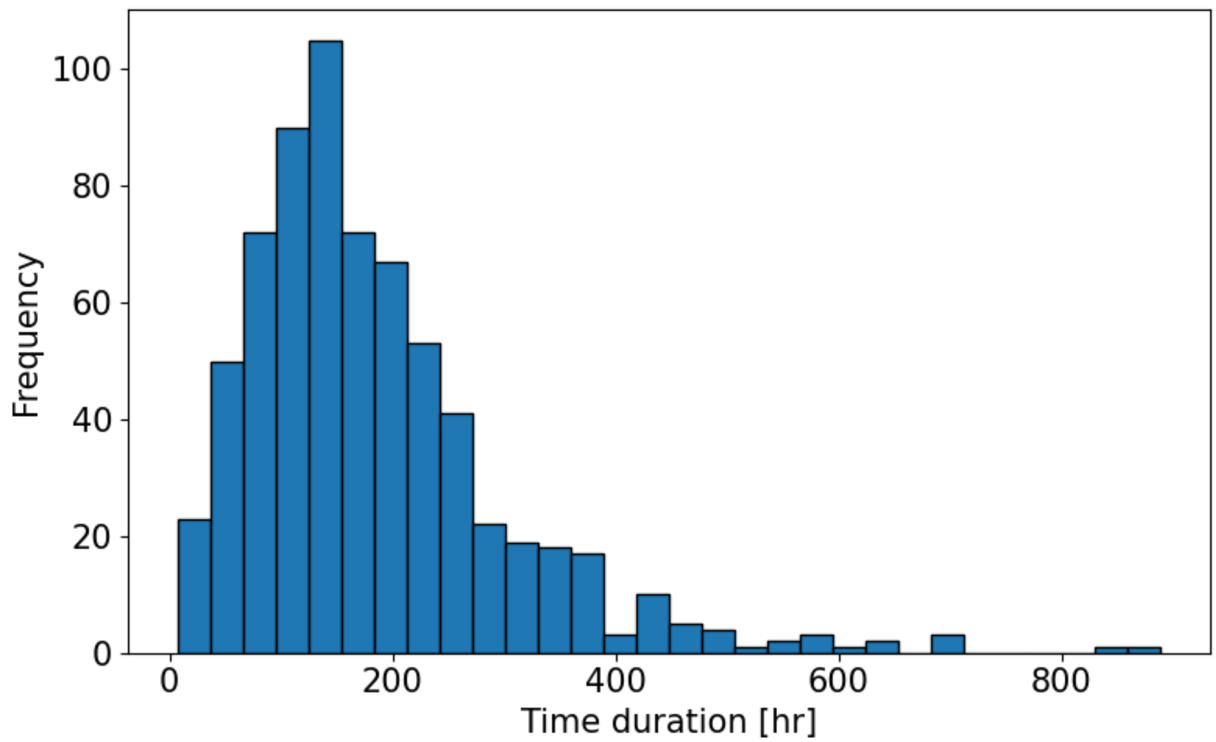
Salvadori, G., De Michele, C., Kottegoda, N. T., and Rosso, R.: Extremes in Nature: An Approach Using Copulas, Water Science and Technology Library, Vol. 56, Springer, Dordrecht, https://doi.org/10.1007/1-4020-4415-1, 2007.

3.Motivating Application Section: It is recommended to clarify the temporal resolution of the BRGM synthetic TC hindcast database, elaborate on the length of the time series, and explain how this resolution meets the dynamic simulation requirements of tropical cyclone processes.

Response: Thank you for the suggestion. We agree that the temporal characteristics of the database should be described more clearly. We will revise and supplement the database description (Lines 103-106) as follows:

"Each storm event is accompanied by hourly time series of the metocean variables mapped across an 18,890-node spatial mesh surrounding Guadeloupe. Storm durations range from 7 to 889 hours (approximately 37 days), with a mean duration of 182.1 hours (approximately 7.5 days). The majority of storms (594 out of 685; 86.7%) last between 50 and 400 hours (approximately 2-17 days). The hourly temporal resolution is considered sufficient to resolve the temporal evolution and interactions among the metocean variables governing total water level throughout the lifecycle of tropical cyclones (Haigh et al., 2011)."

In addition, we will add a histogram of storm durations to the Supplementary Materials:



Additional reference:

Haigh, I. D., Eliot, M., and Pattiaratchi, C.: Global influences of the 18.61 year nodal cycle and 8.85 year cycle of lunar perigee on high tidal levels, *Journal of Geophysical Research: Oceans*, 116, C06025, https://doi.org/10.1029/2010JC006645, 2011.

4.Methodology Section

(1) In Section 3.1 Overview, there is a spelling error in "（3） performance evalaution" (should be "evaluation").

Response: Thank you for noticing the error. This will be corrected accordingly in the revised manuscript.

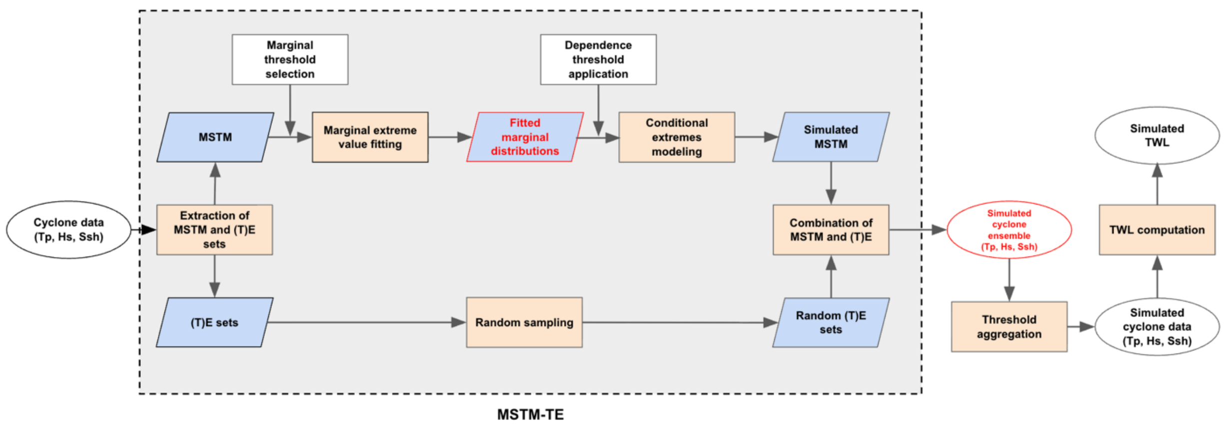
(2) In Section 3.2 Overview of the MSTM-TE Framework, there is a spelling error in "4. Each simulated MSTM vector is paired with an empircial TE set" (should be "empirical").

Response: Thank you for noticing the error. This will be corrected accordingly in the revised manuscript.

(3) For the MSTM-TE framework flowchart (Figure 3, methodology flowchart), it is recommended to supplement input/output data types for each step and label key parameters (e.g., quantiles for threshold selection).

Response: Thank you for this helpful suggestion. We have revised Figure 3 to provide clearer information on the data types (inputs and outputs) and key parameters used at each step. The revisions are as follows (indicated in red in the figure):

(1) We have added the output of the "Marginal extreme value fitting" step as "Fitted marginal distributions".
(2) We have added the output of the "Combination of MSTM and (T)E" step as "Simulated cyclone ensemble (Tp, Hs, Ssh)".



(4) The discussion on the validation of the independence assumption should be further elaborated.

Response: Thank you for this comment. We agree that the discussion supporting the independence assumption can be further strengthened. To address this, we will expand the corresponding lines in Section 4.1 by (1) providing a clearer justification for the east–west clustering strategy and (2) explaining how it improves the validity of the independence assumption used in the proposed framework.

We will specifically reference the analysis of Sando et al. (2024) when describing the partitioning at 61.5°W (lines 320-321), clarifying that the boundary was originally established for the same synthetic tropical cyclone dataset based on evidence of mixed tropical cyclone populations and the physical effects of island-induced wave blocking. To further support this discussion, we will include a more detailed explanation of the physical basis for this partition in the Supplementary Material.

We will also supplement the discussion on how the clustering was validated (lines 322-325) with representative quantitative results from the Kendall's rank correlation analysis before and after clustering, while providing the complete diagnostics (including the number and percentage of spatial locations out of the 18,890 grid nodes exhibiting significant correlations ($p<0.05$) and the corresponding spatial correlation maps) in Section S.3 of the Supplementary Materials.

(5) In the details of multi-threshold aggregation, it is recommended to specify the exact proportion for "equal proportion of storms from each threshold" and explain how to address conflicts in simulation results under different thresholds.

Response: Thank you for the recommendation. We agree that the multi-threshold aggregation procedure should be described more explicitly. We will therefore clarify and mention specifically that each threshold-specific ensemble is generated independently, after which the final ensemble is constructed by randomly sampling an equal proportion of retained storms from each threshold-specific ensemble (i.e., one-third from each of the three thresholds considered in this study).

We will also clarify the rationale for this aggregation strategy. Specifically, the aggregation is restricted to thresholds within the parameter stability region identified through the threshold sensitivity analysis. Consequently, each threshold-specific ensemble is regarded as an alternative realization of the same underlying dependence structure rather than representing fundamentally different statistical models. Under this assumption, the aggregated ensemble reflects sampling variability associated with threshold selection rather than conflicting simulation results across thresholds.

5.Application Section

(1) It is recommended to provide explanatory analysis for cases such as large standard deviations of some parameters in Tables 1 and 2.

Response: Thank you for the valuable suggestion. We agree that the relatively large standard deviations reported for several parameters warrant further discussion. We will therefore expand the corresponding sections to clarify how these large standard deviations reflect characteristics of the underlying dataset and the sensitivity of estimating particular tail properties.

For Table 1, we will provide additional discussion of the estimated shape parameter for the west cluster peak period (Tp), which exhibits the largest standard deviation among the marginal GPD parameters. Our additional analyses indicate that the upper tail of this variable is tightly bounded, with observations approaching closely the estimated upper endpoint (i.e., $u - \frac{\sigma}{\xi}$). Combined with the relatively small number of exceedances (88), this makes the estimation of the tail shape particularly sensitive to the largest observations. To support this interpretation, we will include (1) the figures for the MSTM distributions, (2) the estimated upper endpoints, and (3) spatial maps of the MSTM values for each variable in the Supplementary Materials, together with approriate reference in the main manuscript.

For Table 2, we will expand the discussion of the conditional dependence parameter (α). Additional threshold sensitivity analyses indicate that the relatively large standard deviations occur primarily for variable pairs exhibiting weak extremal dependence (i.e., $\alpha \approx 0$), where small changes in the conditioning sample can substantially influence the estimated parameter values. We will therefore include the corresponding figures for sensitivity analyses in the Supplementary Materials and reference them in the manuscript. We will also provide additional discussion of the physical interpretation of these weaker dependence relationships (e.g., between Ssh and Tp) to better explain both the estimated α values and their associated standard deviations.

(2) It is recommended to supplement the physical interpretation of "timing-offset cases account for 1.93%" and discuss with specific storm cases.

Response: Thank you for the suggestion. We agree that further physical interpretation of the timing-offset cases would strengthen the discussion. We will therefore expand Section 4.3 (lines 399–417) by examining the physical mechanisms responsible for the small influence of these cases on the resulting extreme total water levels.

Specifically, our analyses show that the limited impact of the timing offsets arises from two complementary mechanisms. First, for timing-offset events, the total water level typically evolves gradually around its maximum, resulting in a relatively broad peak rather than a sharp spike. Consequently, evaluating the total water level several hours before or after the maximum generally produces a very similar value. Second, the surge and wave run-up components partially compensate for one another, such that reductions in one component are offset by increases in the other, thereby maintaining a nearly unchanged total water level despite differences in the timing of the individual maxima.

To support this interpretation, we will include in the corresponding section representative storm time series illustrating the evolution of total water level during timing-offset events, together with additional analyses quantifying the compensation between the surge and wave run-up contributions. Specifically, we will include the figure comparing the changes in sea surface height and wave run-up between the times of maximum TWL and maximum wave run-up, and quantify the resulting differences in TWL.

6.References Section

(1) For the reference: Filippini, A., Bouvier, C., Laigre, T., Pedreros, R., Rohmer, J., Balouin, Y., and De La Torre, Y.: High resolution database of cyclone-induced waves for extreme analysis on the (French) Lesser Antilles, Scientific Data. In preparation., 2026. The status is marked as "In preparation"; please confirm if there is an updated status for this reference.

Response: Thank you for the comment. We have confirmed that this reference remains *in preparation*.

(2) For the reference: Bevacqua, E., Vousdoukas, M. I., Zappa, G., Hodges, K., Shepherd, T. G., Maraun, D., Mentaschi, L., and Feyen, L.: More meteorological events that drive compound coastal flooding are projected under climate change, Communications Earth & Environment, 1, https://doi.org/10.1038/s43247-020-00044-z, 2020. The journal name "Communications Earth amp; Environment" contains an error (should be "Communications Earth & Environment").

Response: Thank you for noticing the error. This will be corrected in the revised manuscript.

7.Supplementary Materials Section

(1) Tables S.5.8, S.5.9, and S.5.10 are titled "East cluster" but appear in Section S.5 (Results for the West Cluster); please confirm if this is an error.

Response: Thank you for noticing the error. The captions for Tables S.5.8, S.5.9, and S.5.10 will be corrected to “West cluster” in the revised Supplementary Material.

(2) In Table S.5.7, some standard deviations (values in parentheses) are missing in the row "Tp|Hs"; it is recommended to supplement them.

Response: Thank you for noticing this. The missing standard deviations in Table S.5.7 will be added in the revised Supplementary Material.

(3) The captions for Figures S.7.1 and S.7.2 are identical; please confirm if this is an error.

Response: Thank you for noticing the error. The caption for Figure S.7.2 will be corrected in the revised Supplementary Material as:

"Residual tail dependence coefficient $\bar{\chi}(u)$ as a function of threshold quantile u ..."