Temporally Coherent Modeling of Tropical Cyclone Compound Flooding for Reliable Coastal Hazard Estimation

Min Chung[1], Ryota Wada[1], Jeremy Rohmer[2], and Philip Jonathan[3]
[1]Graduate School of Frontier Sciences, The University of Tokyo, Tokyo, Japan
[2]BRGM, Orleans, France
[3]School of Mathematical Sciences, Lancaster University LA1 4YF, United Kingdom.
Correspondence: Ryota Wada (`r_wada@k.u-tokyo.ac.jp`)

Abstract.

Tropical-cyclone-induced coastal flooding results from the combined action of multiple interdependent processes, whose temporal phasing and magnitude govern the severity of inundation. Conventional statistical frameworks typically model these processes using static event maxima, neglecting their temporal evolution and implicitly forcing co-occurrence of component extremes, which can lead to biased or physically inconsistent design-level estimates. To address this limitation, this study applies the Multivariate Spatio-Temporal Maxima with Temporal Exposure (MSTM-TE) framework, which explicitly embeds temporal coherence in the simulation of compound extremes. The framework is applied to a long-term tropical cyclone dataset for Guadeloupe in the French Antilles, reconstructing time series of significant wave height, peak wave period, and sea surface height to generate physically consistent synthetic storm events. From these simulations, total water level (TWL) is computed as the sum of sea surface height and wave run-up, providing an integrated metric of coastal flooding hazard.

Results show that MSTM-TE reduces bias and ensemble variance in TWL return-level estimation 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, indicating that wave-induced run-up dominates design-level flooding in the study domain. These findings establish MSTM-TE as a physics-informed statistical framework that bridges data efficiency, spatio-temporal realism, and methodological efficiency, offering a robust pathway for compound coastal flood risk assessment under limited observational data.

KEYWORDS: tropical cyclone; spatial extremes; synthetic storm; return value; multivariate analysis

1 Introduction

Tropical cyclones (TCs) are among the most destructive geophysical hazards affecting coastal and island communities worldwide. Over recent decades, TCs have caused substantial global losses, with nearly 2,000 cyclone-related disasters since the 1970s resulting in approximately 779,000 fatalities and economic damages exceeding USD $1.4 trillion worldwide (World Meteorological Organization 2023). Individual events frequently produce catastrophic impacts. Typhoon Haiyan (2013) reached exceptional intensity and caused widespread devastation across the central Philippines, while Hurricane Dorian (2019) intensified to Category 5 strength and produced prolonged extreme winds and severe damage over the northern Bahamas due to its unusually slow translation speed (Mori et al. 2014; Avila et al. 2020). These events highlight both the destructive potential of intense TCs and the persistent vulnerability of densely populated coastal regions.

Among the diverse impacts associated with TCs, coastal flooding consistently ranks as one of the most catastrophic and economically costly. For instance, in the United States, storm surge has been identified as the leading cause of hurricane-related deaths over multi-decadal periods (Rappaport 2014). Notable examples include Hurricane Ike (2008), which generated storm surge exceeding 5 m along portions of the Texas coast, and Cyclone Nargis (2008), which produced water levels exceeding 4 m across the Irrawaddy Delta and resulted in more than 138,000 fatalities in Myanmar (Fritz et al. 2009; Kennedy et al. 2011). Even in events with lower mortality, flood-induced damage to ports, power systems, and transportation networks can disrupt regional economies for months or years (Hallegatte et al. 2013). As global coastal populations continue to grow, reducing the impacts of TC-driven flooding remains an urgent societal challenge.

To mitigate cyclone-induced flood risk, engineers and planners rely on probabilistic design standards that translate the statistical behavior of extreme events into quantifiable design criteria. For instance, coastal defenses such as seawalls, breakwaters, and levees are typically designed for return periods of 50, 100, or 500 years. However, in cyclone-exposed regions, observational records of tides, surge, and waves often span only a few decades, far shorter than the return periods required for design (Menéndez and Woodworth 2010; Wahl et al. 2017). As a result, analysts must rely on statistical extrapolation to infer the behavior of the unobserved tail of the distribution (Coles 2001). The reliability of such extrapolations depends critically on how well statistical models represent the physical processes that generate extremes.

For TC-induced flooding, these processes are inherently compound and involve the dynamic co-evolution of multiple interrelated components, principally wave run-up and storm surge. Wave run-up responds on much shorter timescales to local wave conditions driven by wind and pressure fields, whereas storm surge typically develops more gradually through large-scale oceanic setup and shelf response (Elahi et al. 2023; Mori et al. 2014). The relative timing between these components (i.e., whether surge peaks coincide with, precede, or lag periods of maximum wave activity) largely determines the extent and severity of coastal inundation (Bevacqua et al. 2020). Ignoring this timing effectively forces wave and surge maxima to coincide, creating compound extremes that may never arise from the underlying storm dynamics and consequently leading to biased or physically inconsistent estimates of flood hazard. Despite this, most conventional statistical frameworks either treat these processes independently or rely on static event maxima, failing to represent their temporal evolution and mutual dependence.

Recent advances in conditional extremes modeling and spatio-temporal extensions have sought to address these limitations by incorporating temporal structure into the analysis of joint extremes. The Multivariate Spatio-Temporal Maxima with Temporal Exposure (MSTM-TE) framework represents one such development, reconstructing the full time-series evolution of each metocean variable so that simulated events preserve realistic temporal coherence throughout the storm lifecycle (Sando et al. 2024). Rather than treating time as a secondary attribute, MSTM-TE places temporal evolution at the core of extreme event simulation, enabling the generation of physically consistent compound sequences even under sparse observational conditions.

While a wide range of spatio-temporal statistical models exists, most conventional formulations are designed to describe mean behavior using Gaussian, additive, hierarchical, or lag-based structures. These approaches excel at capturing average spatio-temporal patterns but are not constructed to reproduce tail behavior or the joint structure of rare events. Their reliance on separable covariance structures, linear dependence assumptions, or smoothing-based components limits their ability to represent extremal dependence pathways or the timing alignment that governs compound extremes (e.g., Habereder et al. 2025). In contrast, the MSTM-TE framework is explicitly designed for extreme dynamics, reconstructing full cyclone-time trajectories within a multivariate conditional extremes formulation. This makes it particularly well suited to compound TC hazards, where the interaction and timing of extremes control the resulting impact.

Despite its conceptual appeal, the MSTM-TE framework has not yet been comprehensively validated for real TC-induced flooding hazards. In particular, two questions remain unresolved: (1) whether temporally coherent simulations can produce statistically robust and unbiased return-level estimates under limited observational data, and (2) whether the simulated storm evolutions reproduce realistic wave–surge interaction mechanisms and reveal how extreme compound flooding arises from the joint dynamics of storm components. Addressing these questions is essential for establishing MSTM-TE as a reliable, physics-informed tool rather than a purely statistical generator of extremes.

Objectives and Layout

This study evaluates the capability of the MSTM-TE framework to improve both the statistical reliability and physical interpretability of TC-induced coastal flooding analysis. By explicitly accounting for temporal coherence among compound processes, the study seeks to bridge statistical extreme value modeling and process-based understanding. Specifically, the objectives are twofold:

1. to assess the framework's reliability in reproducing unbiased design-level extremes, with a focus on return levels associated with long return periods under limited data availability; and
2. to use the temporally coherent, physically consistent simulations generated by the framework to investigate the mechanisms governing extreme compound flooding, with particular emphasis on the timing alignment and relative contributions of wave-driven and surge-driven components.

To achieve these objectives, the MSTM-TE framework is applied to a dataset of historical TC events, leveraging its capacity to generate multivariate simulations that preserve both statistical dependence and temporal evolution. From the reconstructed storm time series, a compound coastal flooding risk metric is computed and analyzed, enabling direct comparison between modeled and observed extreme behavior. Framework performance is evaluated against benchmark approaches in terms of bias and variance, while the simulated storm evolutions are examined to identify dominant pathways and co-occurrence mechanisms leading to extreme flooding.

The remainder of this paper is structured as follows. Section 2 introduces the study domain, dataset, and coastal flooding risk metric. Section 3 outlines the modeling methodology, including marginal and dependence modeling, simulation, and validation procedures. Section 4 presents results on the statistical and physical performance of the framework. Section 5 discusses methodological implications and broader insights into compound coastal extremes.

2 Motivating Application

This study focuses on the island of Guadeloupe, located in the Lesser Antilles at the boundary between the Caribbean Sea and the Atlantic Ocean (extending from 60.5°W to 62.5°W longitude and from 15°N to 17°N latitude). Guadeloupe is frequently affected by TCs originating in the North Atlantic and propagating westward along the main hurricane corridor. The island's steep topography, narrow insular shelf, and exposure to energetic Atlantic waves amplify hydrodynamic responses to cyclone forcing, making it particularly vulnerable to compound coastal flooding (Walcker et al. 2019). This combination of coastal morphology and recurrent storm exposure provides an ideal testbed for evaluating statistical frameworks that aim to reconstruct physically consistent storm-induced extremes.

During major TCs such as Hugo (1989), Maria (2017), and Irma (2017), Guadeloupe experienced destructive storm surges, offshore significant wave heights exceeding 10 m, and widespread shoreline erosion (Walcker et al. 2019). These impacts, together with the island's densely developed coastal fringe, highlight the need for hazard-assessment tools capable of translating storm forcing into spatially and temporally coherent estimates of a risk metric.

The analysis uses the 1000-year synthetic TC hindcast database developed by the French Geological Survey (BRGM). The database contains 685 individual storm events that pass within a 300 km radius of the archipelago. Each storm event is accompanied by high-resolution time series of metocean variables mapped across a 18,890-node spatial mesh surrounding Guadeloupe.

From the full model grid, four coastal nodes were selected along Guadeloupe's shoreline to characterize TC-driven coastal flooding risk (Figure 1). The selection targeted grid points located immediately offshore of major sandy beaches (Plage de L'Autre Bord, Plage de Simaho, Anse du Souffleur, and Plage de Sainte-Claire, respectively), where wave-driven processes are of primary concern. Their corresponding latitude–longitude coordinates, in order from 1 to 4, are: (61.37°W, 16.33°N), (61.77°W, 16.06°N), (61.53°W, 16.43°N), and (61.57°W, 16.12°N). Figure 2 further presents a diagram to illustrate the directional distribution and relative contribution of wave energy at each location.

To quantify the impact of TC-induced coastal flooding, this study adopts total water level (TWL) as the principal risk metric. TWL provides a physically interpretable measure of the maximum water elevation at the shoreline during a storm, incorporating both the quasi-steady rise of the sea surface and the dynamic amplification due to breaking waves. This integrated quantity directly relates to inundation potential: the higher the TWL, the greater the likelihood of coastal overtopping and flooding.

In this study, TWL (in meters) is defined as the sum of two components as follows:

$$TWL = S_{sh} + R, \tag{1}$$

where S_{sh} denotes the sea surface height (in meters) and R the wave run-up (in meters). The sea surface height represents the low-frequency, hydrostatic elevation of the water surface at the shoreline, encompassing both tidal and meteorological surge effects. The wave run-up represents the high-frequency, dynamic component associated with wave breaking and uprush, parameterized here as a function of offshore wave characteristics following Stockdon et al. 2006:

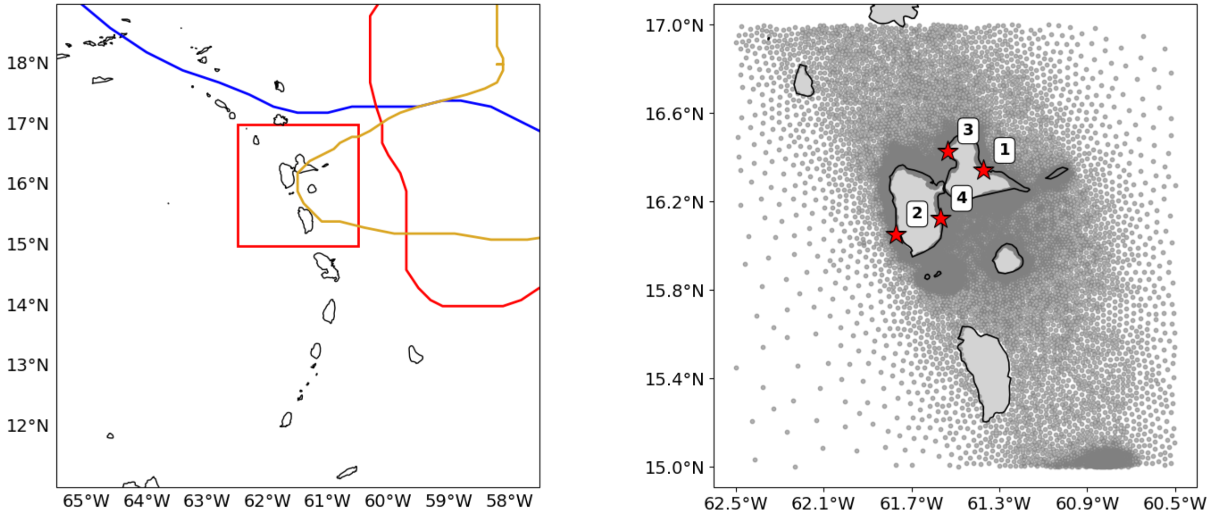
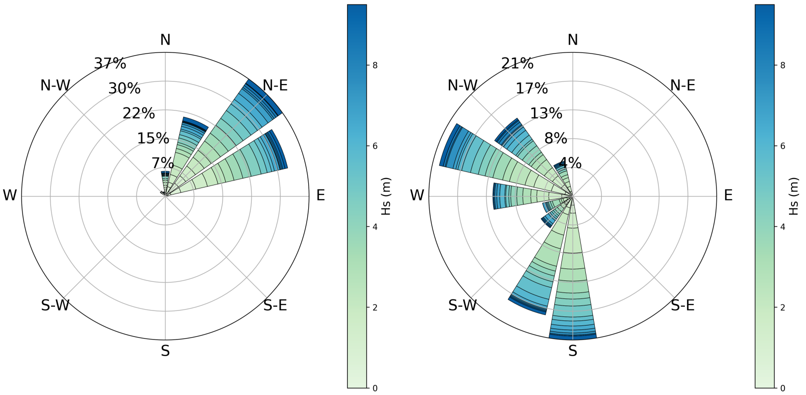


Figure 1. Map of the Guadeloupe region. Left panel: the Lesser Antilles, showing the tracks of storms from which temporal extremes (TEs) were extracted to generate the highlighted events in Fig. 8b. The red box delineates the study domain. Right panel: grid locations from which the datasets were sourced (grey dots) and the four study locations used in the analysis (red stars).

$$R = 1.1\left[0.35\,\beta_f\sqrt{\frac{9.81 H_0 T_p^2}{2\pi}} + \sqrt{\frac{9.81 H_0 T_p^2}{2\pi}\left(0.563\,\beta_f^2 + 0.004\right)}\right], \tag{2}$$

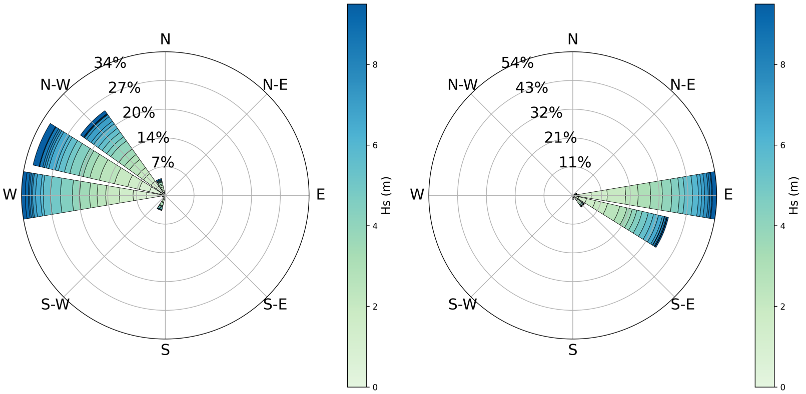
where H_0 is the significant wave height H_s projected orthogonally to the coastline (in meters) and T_p is the peak wave period (in seconds). The foreshore slope β_f is a dimensionless parameter that varies depending on shoreline morphology. In this study, the selected locations are open sandy beaches, for which $\beta_f = 0.1$, a representative sandy-beach slope, is adopted (Melet et al. 2018). Defining TWL in such manner, the multivariate analysis of this research involves the peak values of sea surface height (S_{sh}), significant wave height (H_s), and peak wave period (T_p).

Consistent with this physical interpretation, the three components are modeled from different spatial origins: S_{sh} is evaluated at the shoreline grid point, representing local water-level response, while H_s and T_p (representing wave run-up) are extracted along the 100 m depth contour, where waves remain in the deep-water regime and unaffected by bathymetric transformation. Specifically, for each coastal location, three associated reference points were identified along the 100 m depth contour line to approximate offshore wave characteristics before nearshore transformations occur. One reference point is selected as the nearest grid node on the contour line to the study location, while the other two points are chosen 1 km along the contour line on either side of the first point, following geodesic paths.



(a) Location 1

(b) Location 2



(c) Location 3

(d) Location 4

Figure 2. Directional distribution of H_s at the four study locations. Each sector represents waves arriving from a specific direction, with the radial extent indicating their relative frequency of occurrence. Colors correspond to the magnitude of H_s.

3 Methodology

This section outlines the methodology used to characterize and simulate TC-induced extremes and to translate these simulations into the coastal flooding metric. It details the implementation of the MSTM-TE framework and describes each procedural stage. The description proceeds sequentially, first summarizing the overall workflow (Section 3.1), then elaborating key model components and assumptions (Sections 3.2, 3.3, and 3.4), followed by simulation, reconstruction, and performance evaluation procedures (Sections 3.5, 3.6, and 3.7).

3.1 Overview

The analysis follows a multi-stage approach (Figure 3). The workflow consists of three main stages: (1) application of the MSTM-TE framework (details outlined in Sando et al. 2024), (2) computation of TWL, and (3) performance evalaution.

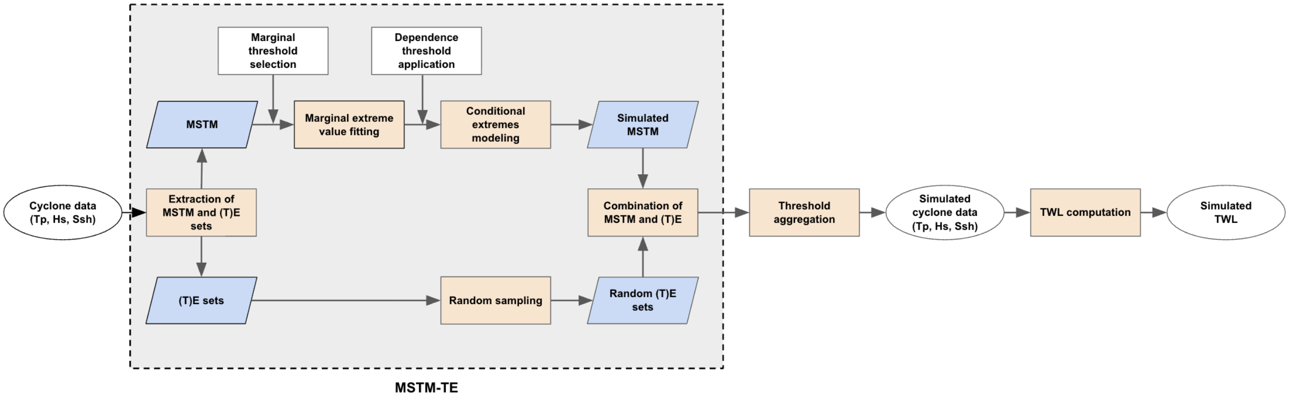


Figure 3. Overview of the methodology used for this study. The grey box indicates the application of the MSTM-TE framework. White circles denote raw input and final data, orange boxes indicate processing or modeling steps, blue parallelograms represent intermediate datasets, and white squares represent specification of key model tuning parameters.

First, storm-level maxima and temporal exposure patterns are extracted from the cyclone dataset (Section 3.2). Their extreme value behavior is then modeled through a combination of marginal tail estimation and multivariate conditional extremes modeling (Section 3.3). This produces synthetic multivariate storm peaks that preserve the dependence structure and extremal characteristics of the historical dataset (Section 3.5). In the second stage, the simulated storm realizations are translated into TWL time series by combining wave- and surge-related components (Section 3.6). Finally, the third stage evaluates model performance against benchmark approaches (Section 3.7), quantifying improvements in accuracy.

3.2 Overview of the MSTM-TE Framework

The MSTM-TE framework provides an integrated approach for modeling and simulating TC extremes by converting raw storm data (comprising multivariate, time-varying spatial fields) into compact descriptors of each event's magnitude and structure, and then using these descriptors to generate new synthetic storms (Sando et al. 2024).

Two components form the core of the framework: Multivariate Spatio-Temporal Maxima (MSTM) and Temporal Exposure (TE). The MSTM summarizes each storm's extremity. For every event, the maximum value of each key metocean variable across the time and location are identified and compiled into a multivariate vector of storm-level peaks.

The TE component captures how the storm evolves in time and space. For each variable, the full temporal sequence at every spatial grid point is normalized by its storm-level maximum. This produces dimensionless exposure fields ranging between 0 and 1, enabling consistent comparison across variables and storms.

The MSTM-TE procedure proceeds through four sequential stages:

1. For each event, MSTM vectors and TE fields are extracted.
2. The marginal tail behavior of each variable is modeled using a generalized Pareto distribution, and the dependence structure among variables is captured using the Heffernan-Tawn (HT) conditional extremes model (Section 3.3).
3. The fitted models are used to generate synthetic MSTM vectors representing plausible combinations of extreme storm conditions (Section 3.5).
4. Each simulated MSTM vector is paired with an empircial TE set, reconstructing full spatio-temporal storm trajectories (Section 3.5).

The framework relies on two principal assumptions (Sando et al. 2024). First, MSTM and TE are assumed statistically independent, allowing synthetic storms to be created by recombining independently sampled MSTM and TE realizations. This assumption is validated through a rank-correlation analysis (Section 3.4).

Second, storm behavior across the dataset is assumed to be statistically stationary in time, enabling resampling from the historical record to yield valid long-term representations of extreme conditions. Exploratory assessment of storm-level maxima does not indicate systematic temporal drift over the observation period.

Together, these components allow the framework to represent both the statistical extremity of storms and the physically realistic temporal structure governing compound coastal hazards.

3.3 Extreme Value Modeling

The MSTM–TE framework models TC extremes through two components: (1) the marginal distribution of each variable, and (2) the dependence structure between variables when extreme conditions occur. Together, these components define the joint probability of compound storm behavior.

The marginal modeling stage focuses on describing the tail behavior of each variable independently. Let S_d, for $d = 1, ..., D$, be the random variable denoting the storm-level maximum of variable d in a set of D variables, with observations $s_{d,n}$ for $n = 1, 2, 3, ...$ For each variable, exceedances above a threshold u_d are modeled using the generalized Pareto distribution (GPD):

$$P(Y_d \leq y | Y_d > 0) = 1 - \left(1 + \frac{\xi_d y}{\sigma_d}\right)^{-1/\xi_d}, \quad y > 0, \tag{3}$$

where $Y_d = S_d - u_d$, $y = s_{d,n}$, ξ_d is the shape parameter, and σ_d is the scale parameter.

To construct the full marginal distribution, the empirical cumulative distribution function (CDF) $F_d^*(s_{d,n})$ is used for values below the threshold, and the GPD tail replaces the exceedance portion:

$$F_d(s) = \begin{cases} F_d^*(s), & s \leq u_d, \\ 1-(1-F_d^*(u_d))\left(1+\frac{\xi_d(s-u_d)}{\sigma_d}\right)^{-1/\xi_d}, & s > u_d. \end{cases} \tag{4}$$

Thresholds are selected using a stability diagnostic in which candidate quantiles of S_d are tested, and the estimated shape parameter ξ_d is inspected for stable regions. To quantify uncertainty and avoid threshold-induced bias, a nonparametric bootstrap is applied: exceedances are resampled, GPD parameters re-estimated, and confidence intervals for ξ_d obtained. Details of the resampling and implementation are provided in the Supplementary Materials (see Section S.1). A final threshold is chosen to balance variance and tail stability.

After the full CDFs are defined for all variables, the storm maxima are transformed to a common standardized Laplace scale before dependence modeling, as each variable has different units and tail shapes. This is achieved by mapping each $F_d(s_{d,n})$ through the inverse Laplace CDF. For each value $s_{d,n}$:

$$s_{L,d,n} = \begin{cases} \log(2F_d(s_{d,n})), & F_d(s_{d,n}) < \frac{1}{2}, \\ -\log(2(1-F_d(s_{d,n}))), & F_d(s_{d,n}) \geq \frac{1}{2}, \end{cases} \tag{5}$$

yielding a variable centered at zero with symmetric exponential tails. This facilitates consistent dependence modeling.

Joint dependence among storm maxima is captured using the HT conditional extremes model (Heffernan and Tawn 2004). Let $Y = (Y_1, Y_2, ..., Y_D)$ be random variables representing the Laplace-transformed maxima. For a chosen conditioning variable Y_d exceeding a dependence threshold ψ_d, the remaining variables satisfy:

$$Y_{d'} \mid (Y_d = y) \approx \alpha_{d'|d}\, y + y^{\beta_{d'|d}} Z_{d'|d}, \quad \text{for } d \neq d', \tag{6}$$

where $\alpha_{d'|d} \in [-1, 1]$ captures linear scaling, $\beta_{d'|d} \in (-\infty, 1]$ controls how dispersion grows with the extremeness of Y_d, and $Z_{d'|d}$ is a residual random variable modeled using a flexible distribution with location $\mu_{d'|d}$, scale $\sigma_{d'|d}$, and shape $\delta_{d'|d}$. The residual distribution corresponds to a generalized Gaussian distribution (also referred to as the delta-Laplace distribution), allowing the residual tail behavior to deviate from a fixed Laplace form.

Each variable alternates as the conditioning variable, producing a complete set of conditional models that characterize the multivariate extremal dependence structure. Taken together, these conditional models define the joint distribution of $Y_{-d}|Y_d$, where Y_{-d} denotes the vector of all components of Y except the conditioning variable Y_d, under the assumption that, given an extreme realization of the conditioning variable Y_d, the remaining components $Y_{d'} : d' \neq d$ are conditionally independent through their respective residual terms $Z_{d'|d}$.

Model performance can be sensitive to the choice of dependence threshold ψ_d. To mitigate this sensitivity, the conditional model is fitted using a set of plausible threshold values (above the marginal threshold), producing distinct parameter sets for each threshold level. Synthetic storms are generated independently for each threshold, and resulting ensembles are later aggregated with equal weighting (Section 3.6.2). This yields a more robust ensemble reflecting a realistic spectrum of extremeness and reduces bias arising from any single threshold choice.

3.4 Independence Validation Test

A key assumption of the MSTM-TE framework is that MSTM and TE are statistically independent. This independence ensures that simulated MSTM values may be validly paired with exposure patterns from any historical storm without introducing structural bias.

To test this assumption, Kendall's rank correlation coefficient (τ) is used. For each variable d and spatial location m, all pairs of TC events (n, n') are compared in terms of their MSTM magnitude and exposure value E. Kendall's τ measures the monotonic association between these quantities without requiring linearity or distributional assumptions.

To assess whether the observed correlation could arise by chance, a permutation-based significance test is performed. Exposure values are randomly reassigned among storms to generate a null distribution for τ. Independence is accepted if the empirical τ lies within the central 95% interval (2.5th–97.5th percentiles) of the null distribution.

This procedure provides a rigorous, distribution-free validation that pairing MSTM and TE samples does not introduce systematic dependence into the synthetic storm reconstruction. This validation is conducted under the assumption of temporal stationarity of storm characteristics.

3.5 Synthetic Tropical Cyclone Simulation

A structured simulation framework is used to generate synthetic TC events that reflect both the statistical and physical characteristics of the historical dataset. The procedure consists of three stages: (1) partitioning of the Laplace-transformed maxima into dominant-component classes, (2) conditional simulation of multivariate extremes, and (3) reconstruction of full spatio-temporal storm fields using empirical exposure patterns.

For each Laplace-transformed storm vector $Y = (Y_1, Y_2, \ldots, Y_D)$, an event is assigned to the partition C_d of variable d when Y_d exceeds its dependence threshold ψ_d, and all other components remain smaller:

$$C_d = \{Y : Y_d > \psi_d \text{ and } \max(Y_{-d}) < Y_d\}, \tag{7}$$

where Y_{-d} represents the full vector of standard Laplace-scale random variables, with the exception of variable Y_d. This ensures that each extreme event has a clearly identified dominant conditioning variable, avoiding overlap among partitions. The empirical occurrence rate ρ_d for dominant component is then calculated as the proportion of storms belonging to C_d.

Given the estimated occurrence rates, synthetic storm are generated iteratively. For each simulated event, a dominant component is sampled according to its empirical probability ρ_d. An extreme value of the conditioning variable Y_d is then drawn from its tail distribution such that $Y_d > \psi_d$. The remaining components Y_{-d} are simulated using the Heffernan-Tawn conditional extremes model (Section 3.3), ensuring that the dependence structure observed in historical storms is preserved. The resulting synthetic vectors are then mapped back from the Laplace scale to the physical domain using the inverse of the marginal transformations, producing complete MSTM vectors representing storm-level peak intensities.

To restore realistic time evolution, each simulated MSTM vector is paired with an empirical temporal exposure sequence from the historical dataset. The selected TE set must originate from a storm with the same dominant component d, ensuring internal physical consistency between the simulated peak structure and its temporal unfolding. This pairing converts synthetic MSTM values into full spatio-temporal time series for all metocean variables.

3.6 Computation of TWL

Following exposure attachment, the simulated extreme values are expanded into complete time series for each variable, allowing TWL to be computed continuously throughout each event using Equation 1. Offshore wave characteristics are derived at three reference points along the 100 m depth contour and averaged to obtain a representative time series for H_s and T_p.

Two additional considerations guide this computation. Section 3.6.1 describes the estimation of the deep-water significant wave height H_0 through shore-normal projection, while Section 3.6.2 outlines the multi-threshold aggregation procedure used to stabilize dependence-based simulations across varying extremeness levels.

3.6.1 Deep-water Significant Wave Height

To isolate the component of offshore wave energy that contributes directly to shoreline run-up, the significant wave height is projected onto the shore-normal direction:

$$H_0 = \cos(\theta - D_p)\, H_s, \tag{8}$$

where H_s is the offshore significant wave height, D_p is the peak wave direction, and θ is the bearing angle from each offshore contour point to the target coastal location. The bearing angle is computed using geodesic azimuths.

This projection removes the influence of obliquely approaching waves, ensuring that only the shoreward-directed component contributes to the run-up term. Because the MSTM–TE simulations provide full time series for both H_s and D_p, the resulting H_0 also evolves dynamically throughout each synthetic storm.

3.6.2 Multi-Threshold Aggregation

Conditional extremes modeling may be sensitive to the choice of dependence threshold ψ_d. To reduce this sensitivity, simulations are conducted independently at several threshold levels above the marginal threshold. For each threshold, the conditional model is fitted, synthetic storms are generated, and only the upper tail of events representing the most extreme conditions is retained. A fixed, equal proportion of storms from each threshold-specific ensemble is then sampled and combined to form a single aggregated dataset. Aggregation is applied across thresholds within a range exhibiting stable HT parameter behavior (i.e., no excessively large uncertainty or marked inconsistency with neighboring thresholds). This ensures balanced representation of extreme behavior across thresholds and yields a more robust ensemble for TWL estimation.

3.7 Performance Evaluation

Model performance is evaluated by comparing the return values estimated from simulated storms with those obtained using less sophisticated models, and by direct evaluation from the TC sample. The return period T is defined through the annual exceedance probability:

$$P(X > x_T) = \frac{1}{T p_a p_e}, \tag{9}$$

where p_a is the annual TC occurrence rate (estimated here as 0.685 from the 685 events over a 1000-year dataset) and p_e is the probability that a storm exceeds the dependence threshold. The return value x_T is the quantile satisfying this exceedance relationship.

For the MSTM–TE framework, the peak TWL for each simulated storm is defined as the maximum value along its reconstructed time series, ensuring consistency with the empirical reference.

Two benchmark configurations are used for comparison against MSTM-TE: MSTM-E and Location-specific estimate (LSE) frameworks.

The MSTM–E framework serves as an intermediate configuration that retains spatial dependence but omits temporal evolution. Instead of reconstructing time series, each storm is represented by the event-wise static maxima of all metocean variables, which are modeled jointly through the multivariate dependence structure and then combined with the corresponding spatial exposure field E. The E field characterizes the relative spatial distribution of storm intensity across the model domain, allowing the framework to preserve spatial variability while neglecting intra-storm temporal coherence. By contrast, the LSE provides a purely local benchmark, in which maxima for each variable and location are extracted and modeled independently, without accounting for cross-variable, spatial dependence, or temporal structure. For both MSTM–E and LSE, return value estimation is based on these static maxima rather than on temporally resolved peaks, meaning that any intra-storm phase information was inherently neglected. To evaluate performance, bias (systematic deviation from ground-truth return values) and variance (dispersion among simulations) were computed for each framework.

4 Application

This section applies the methodology to TC events affecting Guadeloupe and presents the resulting statistical and physical characterization of extreme TWL values. Section 4.1 describes the application setup, including spatial partitioning introduced to satisfy the MSTM–TE independence assumption. Section 4.2 details the extreme value modeling and simulation procedures, including marginal threshold selection and conditional dependence fitting. Section 4.3 examines the reconstruction of synthetic storm time series and quantifies temporal alignment patterns among the metocean variables and TWL. Finally, Section 4.4 evaluates the performance of the MSTM–TE framework relative to benchmark configurations in terms of bias and variance of design-level return values.

4.1 Details of the application

To quantify variability in both individual metocean variables and derived TWL estimates, a large ensemble of synthetic TC events was generated by repeatedly resampling from the original MSTM dataset and applying the conditional simulation procedure. Each resampling iteration draws a subset of storm events equivalent to approximately 50 years of data from the 1000-year synthetic dataset. This window provides a statistically meaningful sample of extremes while maintaining computational feasibility and consistency with typical engineering design horizons.

As described in Section 3.2, the MSTM–TE framework relies on the assumption that MSTM vectors and their corresponding TE fields are statistically independent. Initial diagnostics using Kendall's rank correlation analysis (Section 3.4) indicate that this assumption does not hold uniformly across the full spatial domain, with statistically significant dependence detected between MSTM magnitude and spatial exposure patterns. Following the approach from Sando et al. 2024, the dataset is therefore partitioned into east–west clusters to reflect physically distinct regimes of TC forcing around Guadeloupe. This partitioning captures the contrast between the windward eastern coastline, which is directly exposed to incoming Atlantic swell, and the leeward western coastline, which is partially shielded by island geometry and experiences wave shadowing effects.

Storms are assigned to clusters based on the longitude of maximum significant wave height H_s: events peaking east of 61.5°W define the east cluster, while events peaking west of this boundary define the west cluster. Kendall's rank correlation analysis is then recomputed separately for each cluster to reassess the independence assumption. After east–west partitioning, the proportion of significant correlations decreases sharply (Supplementary Figures S.3.1-S.3.6), confirming that the clustering effectively minimizes residual dependence and allows MSTM and TE fields to be recombined without introducing structural bias. The resulting dataset comprises 466 storms in the east cluster and 219 storms in the west cluster, forming two physically and statistically coherent subsets for subsequent modeling.

In addition to spatial partitioning, two distinct modeling configurations are examined to represent the joint behavior of metocean extremes: a trivariate configuration and a bivariate configuration. In the trivariate configuration, the peak values of significant wave height (H_s), peak wave period (T_p), and sea surface height (S_{sh}) are modeled as three distinct variables within the multivariate conditional extremes framework. This formulation allows each component to vary independently, with their extremal dependence structure estimated explicitly through conditional modeling.

By contrast, the bivariate configuration replaces the separate H_s and T_p predictors with a single compound wave-energy variable, $H_sT_p^2$, paired with S_{sh}. This compound term represents the combined influence of wave height and period on deep-water wavelength and wave-run-up energy, which are the primary mechanisms through which offshore wave forcing contributes to shoreline water levels. Comparing the trivariate and bivariate configurations therefore provides a systematic means of assessing whether a reduced-dimensional, physically motivated representation can reproduce the statistical behavior and physical interpretation of extreme TWL with comparable accuracy while reducing model complexity and parameter uncertainty.

The motivation for introducing the bivariate configuration is rooted directly in the TWL formulation adopted in this study (Equation 1), in which H_s and T_p enter jointly through the run-up parameterization. Treating these variables as a single compound predictor is thus expected to capture their intrinsic co-variation under storm-driven conditions and to reflect the dominant wave-energy pathway governing extreme coastal flooding.

4.2 Extremes modeling and simulations

The extreme value modeling and simulation procedures follow the methodology described in Section 3.3. Marginal GPDs were fitted to exceedances of H_s, T_p, S_{sh}, and the compound variable $H_sT_p^2$ separately for the east and west clusters. Threshold selection was performed using stability diagnostics based on the behavior of the shape parameter and its bootstrap confidence intervals. Across all variables and both clusters, a marginal threshold corresponding to the $u_d = 0.6$ quantile was selected. The full set of threshold diagnostics is provided in the Supplementary Material (see Figures S.2.1 and S.2.2).

Table 1 summarizes the bootstrap mean estimates of the marginal GPD parameters, together with their associated uncertainty. All variables exhibit negative shape parameters ξ, indicating short-tailed distributions with finite upper end points, consistent with physical constraints on storm intensity. In particular, a rapidly decaying upper tail is observed for T_p in both clusters, as reflected by the markedly negative shape parameter estimates. The scale parameter σ is largest for H_s, reflecting the broader variability of wave forcing under TCs, while S_{sh} displays substantially smaller scale values due to its more limited physical range. $H_sT_p^2$ naturally exhibits a large scale parameter as a result of its multiplicative construction, while retaining a negative shape parameter inherited from its constituent variables.

Table 1. Bootstrap mean estimates of the marginal extreme value model parameters for each cluster and variable, based on $N_B = 1000$ resamples. Superscript B denotes bootstrap-based estimates. The table reports the mean threshold (u), scale (σ), and shape (ξ) parameter estimates, with standard deviations shown in parentheses.

Cluster	**Variable**	$\hat{\mathbf{u}}^{\mathbf{B}}$	$\hat{\sigma}^{\mathbf{B}}$	$\hat{\xi}^{\mathbf{B}}$
East	T_p	17.252 (0.444)	4.686 (0.729)	-0.641 (0.105)
	H_s	10.466 (0.374)	5.002 (0.511)	-0.335 (0.065)
	Ssh	0.507 (0.034)	0.317 (0.039)	-0.128 (0.067)
	$H_s(T_p)^2$	3211.879 (141.653)	2888.716 (294.911)	-0.257 (0.089)
West	T_p	14.389 (0.446)	8.404 (3.201)	-1.049 (0.445)
	H_s	6.854 (0.501)	5.298 (0.896)	-0.440 (0.105)
	Ssh	0.311 (0.020)	0.397 (0.054)	-0.303 (0.104)
	$H_s(T_p)^2$	1425.827 (167.897)	2327.514 (606.996)	-0.130 (0.175)

Following the multi-threshold procedure outlined in Section 3.6.2, the HT conditional extremes model was fitted at three dependence thresholds corresponding to 0.6, 0.7, and 0.8 quantiles of the Laplace-transformed conditioning variable. For each threshold, separate conditional models were estimated for all conditioning variables in both the trivariate and bivariate configurations. Across the three dependence levels, the estimated parameters $\alpha, \beta, \mu, \sigma, \delta$ remain broadly consistent, with no systematic drift observed as the dependence threshold increases. This stability indicates that the fitted dependence structure is robust to reasonable variations in threshold choice. For clarity, the results presented in this section focus on $\psi_d = 0.8$, which is representative of the behavior observed across all thresholds.

Figures 4 and 5 show the fitted conditional quantile curves for the trivariate and bivariate configurations, respectively, overlaid on the empirical exceedances at $\psi_d = 0.8$. For the trivariate configuration (Table 2), the estimated linear scaling parameters α are close to zero across all conditioning directions, indicating weak linear dependence in the joint tail. This observation, together with strictly positive estimates of the power parameter β and positive residual location parameters μ, implies increases in the conditional variable through sublinear scaling and residual variability rather than through proportional linear amplification with Y_d.

The power parameters β generally take values below unity, implying sublinear growth of conditional variability with increasing extremeness of the conditioning variable. In particular, conditional relationships involving T_p display small β values, reflecting the rapid saturation of wave period in the upper tail and its limited ability to extend the joint extremal structure independently of wave height.

Residual location and scale parameters exhibit moderate variation across conditioning directions, while the residual shape parameter δ remains bounded, indicating stable residual tail behavior. Together, these features suggest that the joint tail of (H_s, T_p, S_{sh}) is constrained and structured, with extreme variability dominated primarily by wave height and sea surface height rather than by independent excursions of wave period.

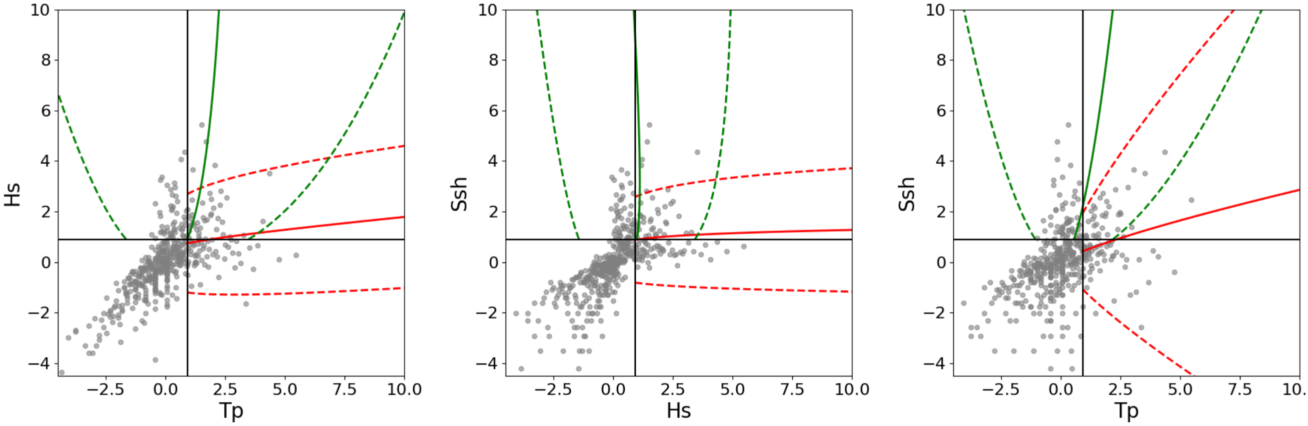


Figure 4. Fitted conditional quantile curves for each conditioning variable (conditioning on the x-axis variable in green and on the y-axis variable in red) for the trivariate configuration at the 0.8 dependence threshold, all shown on the standard Laplace scale. Solid lines represent conditional extremes models fitted using the bootstrap mean estimates of $(\alpha, \beta, \mu, \sigma, \delta)$ and residual correlation, and dashed lines show the associated 95% uncertainty bands (2.5–97.5% bootstrap quantiles) of residual distribution.

For the bivariate configuration, Table 3 reports the fitted parameters for the conditional relationships between $H_s T_p^2$ and S_{sh}. As in the trivariate case, linear scaling parameter α remains close to zero, while the power parameter β exhibits moderate asymmetry between conditioning directions. Conditioning on $H_s T_p^2$ yields a larger β value than conditioning on S_{sh}, indicating

Table 2. Bootstrap mean estimates of the HT parameters $(\alpha, \beta, \mu, \sigma, \delta)$ and residual correlation for the trivariate configuration at the 0.8 dependence threshold, based on $N_B = 1000$ resamples. Standard deviations are shown in parentheses.

Model form	$\hat{\alpha}^{\mathbf{B}}$	$\hat{\beta}^{\mathbf{B}}$	$\hat{\mu}^{\mathbf{B}}$	$\hat{\sigma}^{\mathbf{B}}$	$\hat{\delta}^{\mathbf{B}}$	**Correlation**
$H_s \mid T_p$	0.082 (0.270)	0.154 (0.207)	0.680 (0.936)	0.856 (0.710)	1.548 (0.836)	0.400 (0.001)
$S_{sh} \mid T_p$	0.000 (0.340)	0.793 (0.236)	0.461 (0.865)	0.701 (0.589)	1.403 (0.637)	0.400 (0.001)
$T_p \mid H_s$	-0.076 (0.627)	0.464 (0.348)	1.031 (0.051)	1.165 (0.339)	0.760 (0.241)	0.400 (0.001)
$S_{sh} \mid H_s$	0.000 (0.156)	0.153 (0.110)	0.896 (0.499)	0.745 (0.607)	1.774 (23.526)	0.400 (0.001)
$T_p \mid S_{sh}$	0.006 (0.217)	0.560 (0.156)	0.584 (0.688)	0.746 (0.416)	2.086 (1.580)	0.207 (0.187)
$H_s \mid S_{sh}$	-0.095 (0.599)	0.214 (0.356)	1.108 (2.091)	1.076 (5.412)	1.280 (73.379)	0.207 (0.187)

that variability in sea surface height increases more strongly under extreme wave-energy conditions than vice versa. When conditioning on S_{sh}, the combination of near-zero α and smaller β value implies weak conditional dependence on $H_sT_p^2$ on extreme sea surface height, resulting in increased bootstrap variability in the residual parameters (μ, σ, δ).

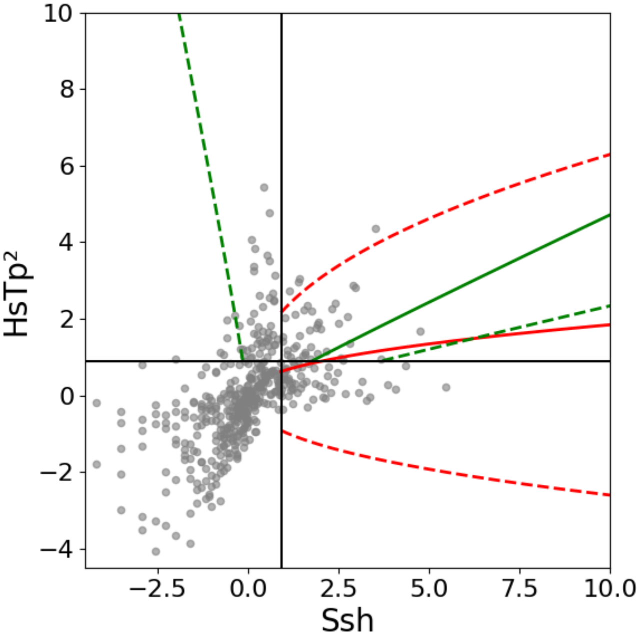


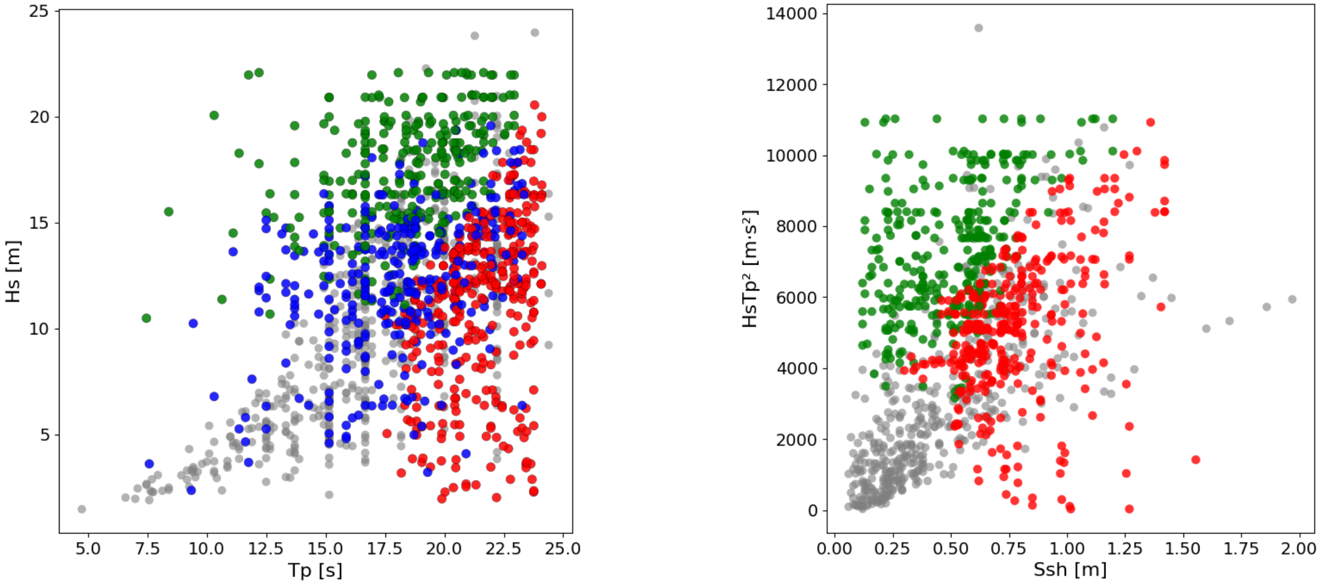
Figure 5. Fitted conditional quantile curves for each conditioning variable (conditioning on the x-axis variable in green and on the y-axis variable in red) for the bivariate configuration at the 0.8 dependence threshold, shown on the standard Laplace scale. Solid and dashed lines specifications are the same as in Fig. 4.

For each modeling configuration and cluster, a total of 100 independent subsets were generated, with each subset containing more than 1000 synthetic TCs. These ensembles form the basis for subsequent time-series reconstruction, temporal alignment

Table 3. Bootstrap mean estimates of the HT parameters $(\alpha, \beta, \mu, \sigma, \delta)$ for the bivariate configuration at the 0.8 dependence threshold, based on $N_B = 1000$ resamples. Standard deviations are shown in parentheses.

Model form	$\hat{\alpha}^{\mathbf{B}}$	$\hat{\beta}^{\mathbf{B}}$	$\hat{\mu}^{\mathbf{B}}$	$\hat{\sigma}^{\mathbf{B}}$	$\hat{\delta}^{\mathbf{B}}$
$H_sT_p^2 \mid S_{sh}$	0.005 (0.162)	0.443 (0.127)	0.648 (0.533)	0.696 (0.893)	1.391 (0.681)
$S_{sh} \mid H_sT_p^2$	0.082 (0.542)	1.050 (0.328)	1.886 (1.786)	0.925 (0.970)	0.769 (0.482)

analysis, and performance evaluation. Sample visualizations of simulated MSTM vectors for the east-cluster storms are shown in Figure 6 for both the trivariate and bivariate configurations. Equivalent visualizations for the west-cluster storms display similar dependence structure and variability (Supplementary Fig. 5.5).



(a) MSTM realizations for the trivariate configuration

(b) MSTM realizations for the bivariate configuration

Figure 6. Sample visualization of the simulated MSTM vectors for East cluster storms, using the (a) trivariate and (b) bivariate modeling configuration. Grey dots represent empirical TCs, For (a), red dots indicate T_p-conditioned simulations, green dots indicate H_s-conditioned simulations, and blue dots indicate S_{sh}-conditioned simulations. For (b) green dots indicate $H_sT_p^2$-conditioned simulations, and red dots indicate S_{sh}-conditioned simulations.

4.3 Time Series Reconstruction

Following the procedures outlined in Section 3.5, full time series are reconstructed for each metocean variable and for the derived TWL function. These reconstructions allow the temporal evolution of wave forcing, sea surface height, and their combined effect on coastal water levels to be examined explicitly. Figure 7 presents an example of reconstructed storm time series under the trivariate configuration at Location 1, illustrating the temporal evolution of H_s, S_{sh}, T_p, and the resulting TWL.

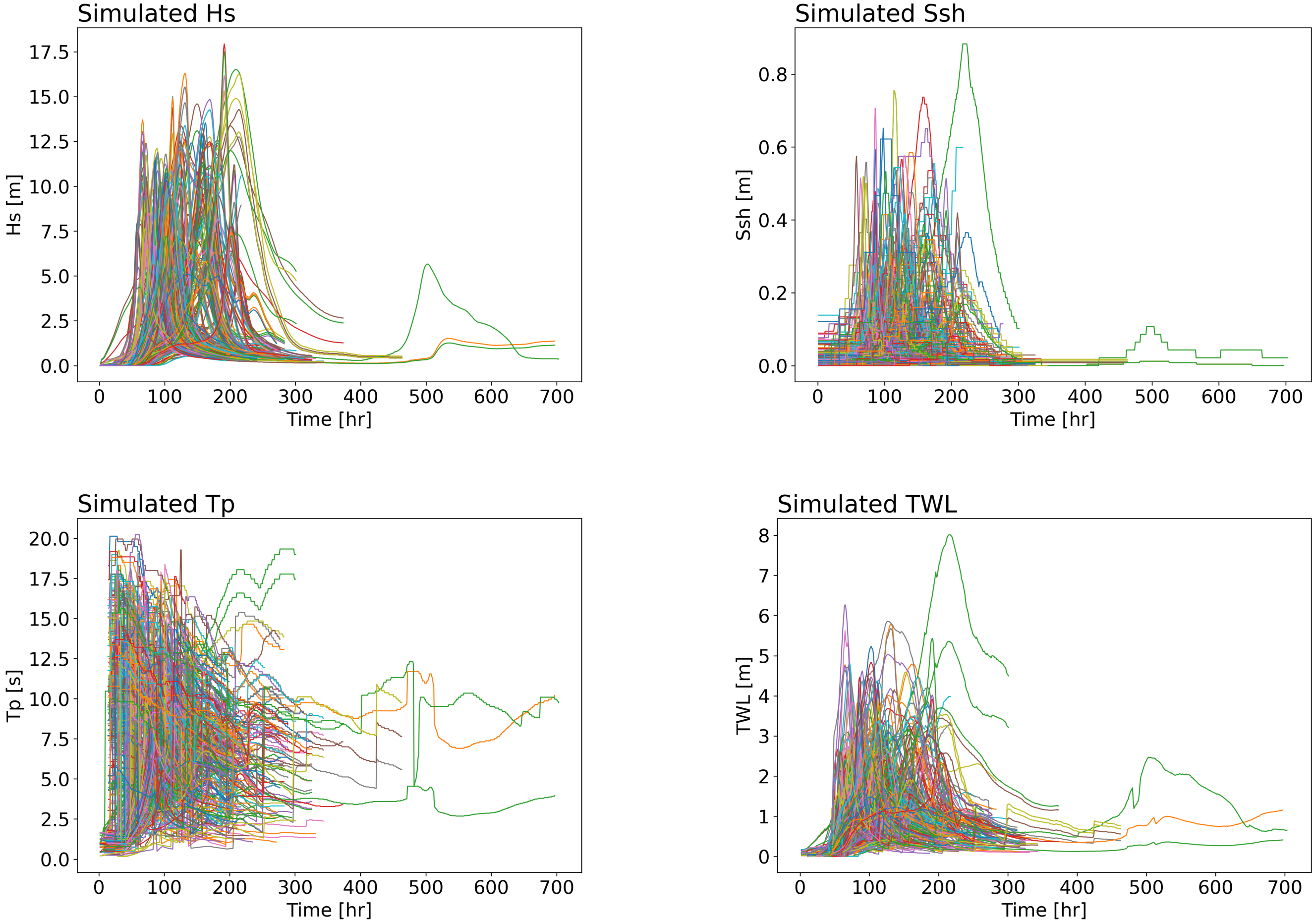
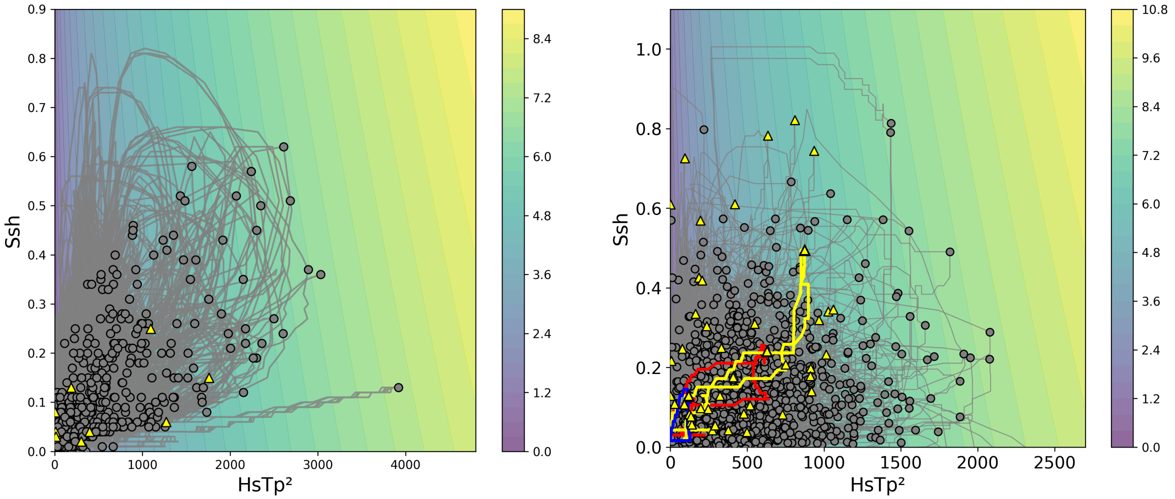


Figure 7. Time series of TC events simulated using MSTM-TE at Location 1 (trivariate configuration; East-cluster storms). Each curve represents the temporal evolution of each simulated cyclone.

To evaluate whether the simulated storms reproduce the joint temporal behavior observed in the historical record, storm trajectories are examined in the $(H_sT_p^2, S_{sh})$ phase space for the bivariate configuration. Each trajectory traces the co-evolution of wave-energy forcing and sea surface height throughout the duration of an individual storm, with markers indicating the timing of peak TWL. Equivalent trajectories are constructed for both empirical and simulated storms to facilitate direct comparison (Figure 8).

The empirical storm trajectories show a clear tendency for peak TWL to co-occur with the maximum of $H_sT_p^2$, indicated by grey discs in Figure 8a. Only a small subset of events deviates from this behavior, with peak TWL occurring at times offset



(a) Empirical storms

(b) Simulated storms

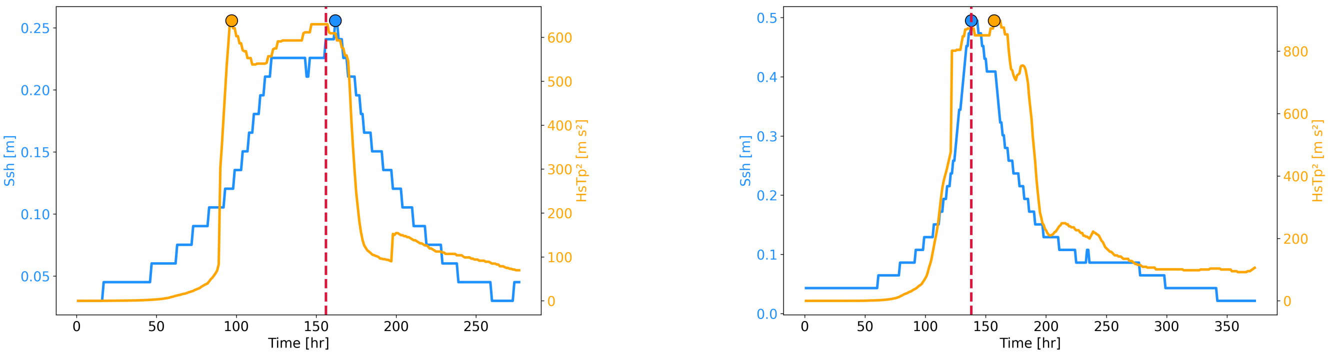
Figure 8. Joint time series trajectories of storms (East-cluster storms) for Location 1 in $H_sT_p^2$-S_{sh} space, with the colorscale representing the magnitude of the resulting TWL. Grey discs indicate the timing of peak TWL per storm that co-occurs with peak $H_sT_p^2$, and yellow triangles indicate the timing of peak TWL that does not co-occur with peak $H_sT_p^2$.

from the wave-energy maximum (indicated by yellow triangles). Quantitatively, such timing-offset cases account for 1.93% (9 out of 466 storms) of the empirical storm sample. The simulated storms generated by the MSTM–TE framework reproduce this characteristic structure closely (Figure 8b). In the simulated ensemble, timing-offset events similarly represent a small fraction of storms (3.22%; 42 out of 1304 storms), and while the synthetic ensemble contains a larger number of deviating cases owing to its greater size, the dominant pattern of temporal co-occurrence between peak TWL and peak $H_sT_p^2$ remains consistent.

To further illustrate the behaviors of the storms displaying deviant timing relationships between TWL and its contributing components, three representative simulated storms are selected to typify distinct peak-timing configurations: (1) peak TWL occurring prior to the maximum of S_{sh} (red-highlighted storm in Figure 8b), (2) peak TWL coinciding with the maximum of S_{sh} (yellow-highlighted), and (3) peak TWL occurring after the maximum of S_{sh} (blue-highlighted). Figure 9 presents the reconstructed time series of $H_sT_p^2$, S_{sh}, and TWL for these representative cases, with markers indicating the timing of each peak.

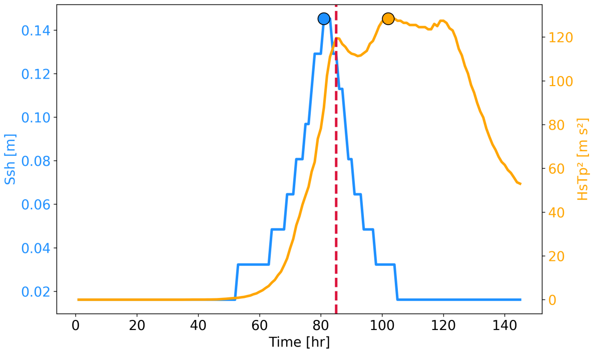
Across these examples, the timing of peak TWL is consistently controlled by the evolution of wave-energy forcing, while sea surface height modulates the magnitude of the resulting response. Even in cases where peak S_{sh} does not coincide with peak $H_sT_p^2$, extreme TWL emerges when elevated surge levels interact with high (but not necessarily maximal) wave energy.

It is also noteworthy that, across the full set of empirical storms and simulated ensembles, no event was identified in which the TWL maximum occurs outside a ± 10 hour window of both the S_{sh} maximum and the $H_sT_p^2$ maximum. Instead, peak



(a) Temporal evolution of each variable for the red-highlighted storm

(b) Temporal evolution of each variable for the yellow-highlighted storm



(c) Temporal evolution of each variable for the blue-highlighted storm

Figure 9. Temporal evolution of $H_sT_p^2$ (yellow) and S_{sh} (blue) for the highlighted storm ((a) red, (b) yellow, and (c) blue) from Fig. 8b. Discs mark peak value recorded for each variable. The vertical red dashed line represents the timing in which peak TWL occurs.

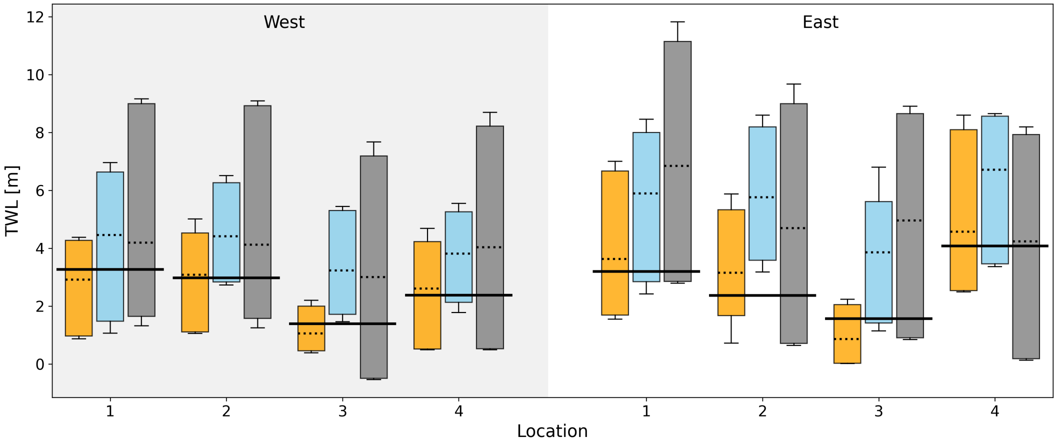
TWL remains temporally constrained to periods when at least one component is near its maximum while the other remains elevated.

4.4 Performance Assessment

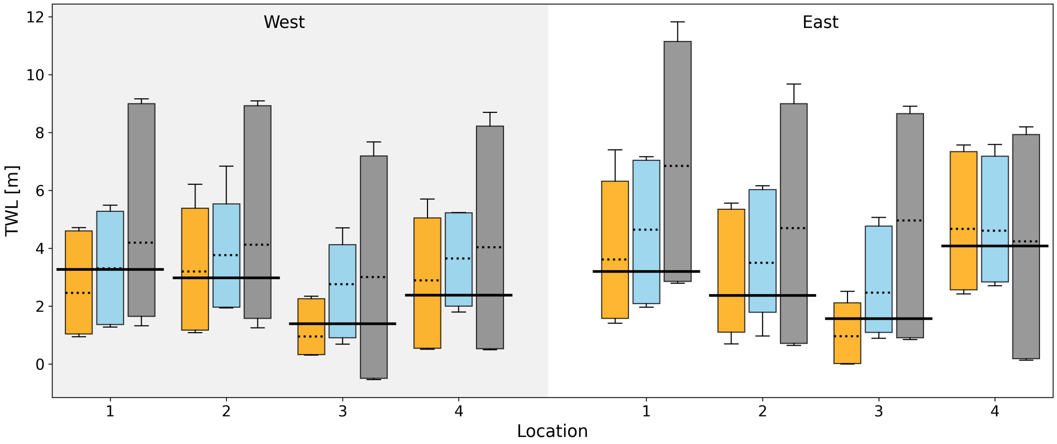
4.4.1 MSTM-TE

420 Following the procedures outlined in Section 3.7, return values of TWL are estimated for each ensemble subset under the three statistical frameworks considered in this study: MSTM–TE, MSTM–E, and LSE. The analysis focuses on a return period of $T = 100$ years and is conducted for all study locations under both the trivariate and bivariate modeling configurations. Results for an additional return period ($T = 300$ years) are provided in the Supplementary Materials (see Fig. S.6.1 - S.6.3). To

facilitate comparison, the resulting return values are summarized using box–whisker plots, which illustrate the central tendency and dispersion of estimates across ensemble subsets (Figure 10).



(a) Trivariate configuration



(b) Bivariate configuration

Figure 10. Box–whisker plots of 100-year return value estimates of TWL obtained from MSTM–TE (yellow), MSTM–E (blue), and LSE (grey), shown for west-cluster storms (left) and east-cluster storms (right). Estimates are based on an ensemble of 100 independent subsets, each comprising 1000 synthetic tropical cyclones. For each method, boxes represent the central 95% range (2.5th–97.5th percentiles) across ensemble subsets, the central line denotes the median, and whiskers extend to the minimum and maximum values within each ensemble. The black horizontal line indicates the ground-truth return value estimated directly from the full TC dataset.

Across all locations and configurations, a clear performance hierarchy emerges. The MSTM–TE framework consistently produces TWL return value estimates with reduced bias relative to MSTM–E, while MSTM–E in turn exhibits lower variance than the purely local LSE approach. This ordering reflects the progressive incorporation of spatial dependence and temporal coherence across the three frameworks. By reconstructing the full temporal evolution of storm components, MSTM–TE avoids the artificial inflation or suppression of extremes that arises when event-wise maxima are combined independently, without regard to intra-storm timing.

The trivariate and bivariate configurations exhibit nearly indistinguishable performance under all three frameworks. This similarity indicates that replacing the separate wave variables H_s and T_p with $H_sT_p^2$ does not degrade the accuracy or robustness of TWL return-level estimation. Instead, the reduced-dimensional representation preserves the dominant statistical structure governing extreme coastal water levels while simplifying the dependence model.

Spatial variability among study locations is also more clearly resolved by the MSTM–TE framework than by the benchmark approaches. As shown in Fig. 10, the ground truth TWL return values vary substantially across locations, with Location 3 exhibiting a markedly lower return value than the other sites. This spatial contrast is reflected in the MSTM-TE results, which show a correspondingly lower median TWL estimate at Location 3 and reduced bias relative to the ground truth. In contrast, return values estimated using LSE, which does not explicitly account for spatial variability, remain relatively uniform across locations, indicating limited sensitivity to spatial differences in storm evolution and exposure.

To further examine the physical drivers underlying extreme TWL values, the contributions of sea surface height and wave run-up to the 100-year TWL return level are analyzed under the bivariate configuration. Figure 11 presents the corresponding value ranges of S_{sh} and the run-up component R associated with extreme TWL events. Across all locations, wave run-up consistently contributes a larger fraction of the total response than storm surge, although the relative magnitude varies spatially. In particular, locations with reduced wave exposure exhibit lower run-up contributions and correspondingly lower TWL extremes.

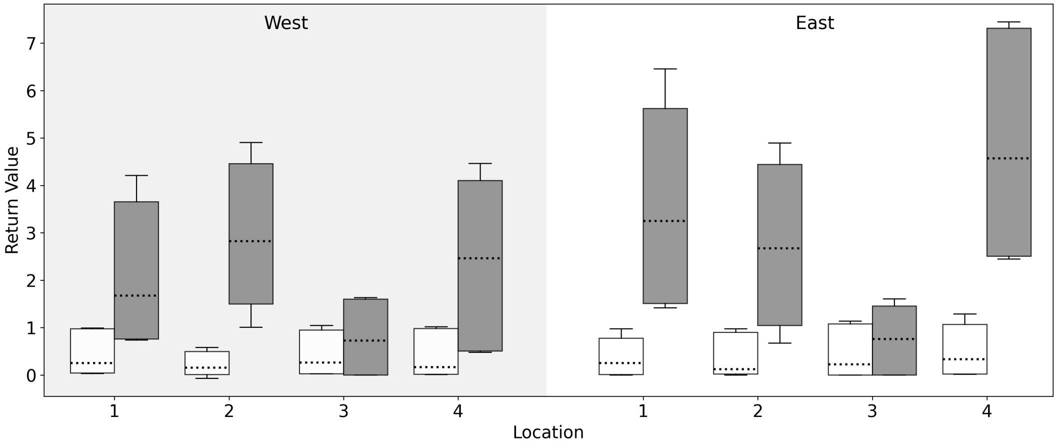


Figure 11. Box–whisker plots showing the contributions of sea surface height (white boxes) and wave run-up (grey boxes) to the 100-year return value of TWL across study locations. Box–whisker specifications are the same as in Figure 10.

4.4.2 Simplified representation

The systematic temporal alignment identified in Section 4.3 motivates exploration of a simplified alternative to the full MSTM–TE framework. Because peak TWL is observed to co-occur consistently with the peak of $H_sT_p^2$, a reduced approach is considered in which variable values are extracted at the time of maximum wave-energy forcing rather than through full time-series reconstruction. This timing-aligned framework retains the dominant co-occurrence mechanism while substantially reducing computational and data requirements.

Figure 12 compares TWL return values obtained using this simplified framework with those produced by the full MSTM–TE framework under both the trivariate and bivariate configurations. The simplified approach reproduces the overall range and distribution of return values closely, with only minor deviations from the full framework across locations and clusters.

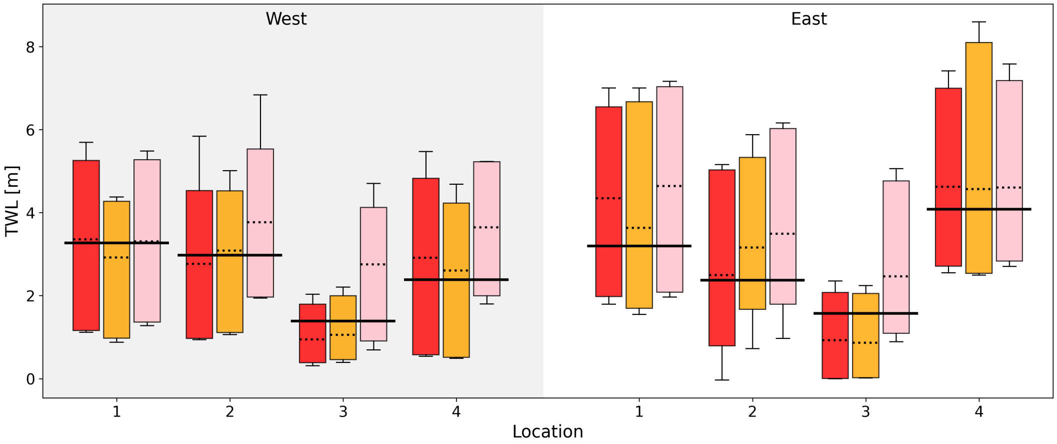


Figure 12. Box–whisker plots of 100-year return value estimates of TWL obtained from the simplified timing-aligned framework (red), the MSTM–TE framework under the trivariate configuration (yellow), and the bivariate configuration (pink), shown for west-cluster storms (left) and east-cluster storms (right). Box–whisker specifications are the same as in Figure 10.

5 Discussion

This study evaluates how multivariate extreme value frameworks reconstruct TC-driven TWL and whether explicitly retaining the temporal structure of storms improves the reliability of design-level estimates. Using the MSTM-TE framework, relevant metocean variables and their compound formulation are jointly modeled, and storm time series across multiple coastal locations are reconstructed.

Across all locations, MSTM-TE systematically reduces bias and variance relative to estimates obtained using joint extreme models based on storm-level maxima, which do not account for intra-storm temporal dependence or alignment. This improvement arises from preserving the full temporal evolution of storm components rather than combining isolated maxima, thereby avoiding the artificial inflation or suppression of extremes that can occur when temporal coherence is neglected. Notably, reconstructed TWL extremes are highly consistent between the trivariate and bivariate configurations, indicating that dimensionality reduction through the compound wave variable preserves the dominant physical and statistical structure of extreme sea states in this region.

5.1 Simplification of the dependence structure

The joint marginal and conditional extremes analysis further reveals that the upper-tail behavior of (H_s, T_p, S_{sh}) is considerably more constrained than a fully unconstrained trivariate system would suggest. The peak wave period T_p exhibits a short, tightly bounded tail, together with near-zero linear scaling and rapid decay in the conditional fits involving T_p. This behavior indicates that wave period does not extend the joint tail independently of wave height. Instead, extreme values of T_p occur only when H_s is already large, causing the wave-related component of the sea state to collapse onto a single dominant run-up axis governed primarily by H_s. In contrast, sea surface height S_{sh} exhibits a less restrictive tail shape and more asymmetric conditional behavior: it often increases alongside high values of $H_sT_p^2$, but can also attain unusually large values when wave energy is only moderate. This asymmetry, combined with weaker dependence when conditioning on S_{sh}, indicates that surge contributes a partially independent extremal direction, in the sense that while surge and wave energy often increase together, the most extreme wave run-up events are not systematically accompanied by equally extreme surge. This interpretation is further supported by diagnostics of extremal dependence through the tail correlation coefficient $\chi(u)$ and the tail dependence coefficient $\bar{\chi}(u)$ (see Supplementary Section S.7). Taken together, these results show that although the physical sea state is described by three variables, its effective dimensionality with respect to extreme TWL generation is reduced to two dominant axes: a wave-run-up axis, with T_p embedded within it, and a secondary surge axis that is not fully aligned with wave processes.

Within this reduced effective-dimensional structure, the behavior of the compound run-up proxy $H_sT_p^2$ follows naturally. Rather than introducing a new mode of extremeness, this transformation amplifies variability along the same dominant run-up axis that already governs the joint tail of (H_s, T_p). Because T_p does not define an independent extremal direction and instead saturates rapidly in the tail, multiplication by T_p^2 magnifies variations in H_s without altering the underlying geometry of the extremes. This explains why the bivariate configuration using $(H_sT_p^2, S_{sh})$ reproduces the trivariate results closely: the compound variable preserves the effective structure of the wave-driven tail while providing a clearer physical representation of the run-up process that drives the majority of extreme TWL events.

5.2 Role of spatial exposure

Observations from both the original storm record and the MSTM-TE simulations under the bivariate configuration consistently show that peaks in TWL coincide with peaks in the compound variable $H_sT_p^2$. This alignment indicates that wave run-up, as represented through the $H_sT_p^2$ formulation, is the predominant driver of extreme TWL events and therefore a primary contributor to coastal flooding risk. Motivated by this behavior, a physically informed statistical framework is developed in which the timing of peak run-up is first identified, and the corresponding value of S_{sh} is then simulated to produce a time-consistent estimate of peak TWL (Section 4.4.2). This timing-aligned approach embeds the physical co-occurrence mechanism observed in the data directly into the statistical reconstruction of extremes. When evaluated against the full MSTM-TE framework, the method demonstrates broadly comparable performance across locations, reinforcing the conclusion that extreme TWL response is governed primarily by the run-up component and that capturing its timing enables simplified yet accurate return-level estimation with reduced statistical complexity.

Spatial differences in storm exposure further play a significant role in shaping the dominant drivers of extreme TWL across the study locations. Although wave run-up consistently exceeds storm surge in magnitude at all sites, Location 3 exhibits noticeably lower run-up values. As a consequence, the resulting TWL extremes are also lower at this location, a pattern observed consistently in both the original and simulated storm sets. This finding emphasizes that TWL return levels are modulated not only by the statistical dependence among storm components but also by spatial variability in the wave field, which determines the extent to which run-up acts as the dominant mechanism of extreme loading.

In such locations with limited wave energy, rare combinations arise in which extreme storm surge coincides with only moderate run-up. Such events occur infrequently in the observational record, making their extreme pathways difficult to characterize using data alone. For instance, at Location 3, 12.70% of the original 685 storms (29 among the 219 West-cluster storms and 58 among the 466 East-cluster storms) correspond to cases where the maximum S_{sh} exceeds the 0.6 quantile threshold of the marginal distribution while the coinciding $H_sT_p^2$ remains at / or below the same quantile. The full time-series reconstruction enabled by the MSTM-TE framework generates additional realizations of these low-energy compound extremes, allowing their behavior to be examined more systematically. In contrast to run-up-dominated regimes, peak TWL in these environments does not coincide with the maximum run-up, nor does it necessarily align with the peak surge. Instead, the TWL maximum emerges from the interaction between the two components and their relative timing. This behavior indicates that, in low-energy settings, neither the magnitude nor the timing of an individual component is sufficient to describe the extremal response, and accurate characterization of TWL requires simulation of the joint temporal evolution of surge and waves.

5.3 Influence of local coastal morphology

The wave run-up component is also governed by local coastal morphology, particularly the foreshore slope β_f. In Equation 2, run-up increases monotonically with β_f, such that steeper beaches produce larger uprush amplitudes, whereas flatter, dissipative beaches attenuate wave energy. The foreshore slope adopted here ($\beta_f = 0.1$) is representative of gently sloping sandy beaches. Consequently, the finding that wave-driven run-up dominates extreme TWL formation in this domain should not be assumed universal. Even within sandy coastal settings, environments characterized by substantially flatter morphologies may exhibit a reduced relative contribution of run-up. For instance, when the MSTM-TE framework is applied using alternative representative slopes ($\beta_f = 0.01, 0.025, 0.05, 0.075$), the relative contribution of wave run-up to extreme TWL shifts systematically, with reduced slopes favoring a greater relative influence of surge (Fig. 13). Nevertheless, even at the lowest range of slopes commonly reported for sandy beaches ($\beta_f = 0.01$), wave run-up remains comparable to or exceeds surge contributions at most locations. Only at Location 3 does surge contribution exceed that of run-up under very gentle slopes ($\beta_f \leq 0.05$). This suggests that while morphological flattening influences the relative contribution of run-up, its capacity to shift the dominant TWL driver is conditional on the offshore energy regime. In wave-limited environments (e.g., Location 3), extreme TWL composition is more sensitive to changes in foreshore slope, whereas in more wave-exposed settings, run-up dominance remains robust across realistic sandy beach gradients.

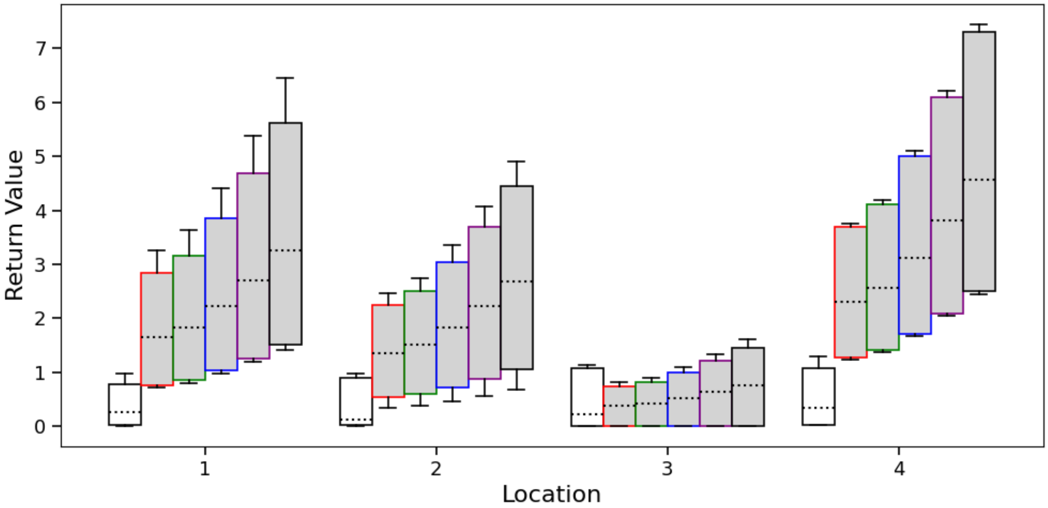


Figure 13. Box–whisker plots illustrating the relative contributions of wave run-up and sea surface height and the influence of alternative β_f values to the 100-year return value of TWL across study locations for East-cluster storms. Wave run-up contributions are shown as grey boxes, with border colors indicating alternative foreshore slopes ($\beta_f = 0.01$, red; $\beta_f = 0.025$, green; $\beta_f = 0.05$, blue; $\beta_f = 0.075$, purple; $\beta_f = 0.1$, black). Contributions from S_{sh} are shown as white boxes. Box–whisker specifications are the same as in Figure 10.

5.4 Methodological implications and limitations

It should be noted that the present formulation of wave run-up (Equation 2) is derived for open sandy beaches and therefore does not generalize to coastlines characterized by fundamentally different morphologies (e.g., rocky dikes, revetments, or armored shorelines). In such settings, wave run-up behavior is governed by distinct physical mechanisms, and alternative run-up formulations are therefore required (e.g., van der Meer and Stam 1992; Shankar and Jayaratne 2003).

More broadly, the use of an analytical run-up parameterization may introduce structural limitations, as it does not explicitly resolve nearshore wave-transformation processes (e.g., refraction, shoaling, infragravity dynamics, or the influence of complex bathymetry) that can alter run-up scaling behavior. Such limitations can be more pronounced in environments where the coupling between offshore forcing and shoreline response is strongly nonlinear, such as under high-energy wave conditions (Senechal et al. 2011) or in reef-lined/vegetated coastal settings (Laigre et al. 2023). More process-based numerical models, such as XBeach (Roelvink et al. 2009), FUNWAVE-TVD (Shi et al. 2013), or Boussinesq-type solvers (e.g., Wei et al. 1999), could better represent these mechanisms. Alternatively, surrogate and reduced-order modeling approaches (e.g., Gaussian process–based coastal flood emulators; Anderson et al. 2021) may offer computationally efficient pathways for incorporating improved run-up physics within extreme value frameworks. Integrating such approaches would help reduce structural uncertainty in TWL prediction and enhance physical fidelity in regions characterized by complex coastal morphodynamics.

Furthermore, application of the MSTM-TE framework to other datasets or regions may require explicit verification of temporal stationarity, as stability of marginal extremes and their dependence structure is a fundamental assumption of the resampling procedure. Formal evaluation could be implemented by extending rank-based diagnostics to test for temporal trends in marginal exceedances (e.g., Mann-Kendall trend tests; Yue et al. 2002) or by fitting non-stationary extreme value models with time-varying parameters (e.g., Dey and Patwary 2025). Within the conditional extremes framework, time-varying dependence parameters could likewise be assessed to detect systematic changes in extremal association (Eastoe and Tawn 2008). Such analyses would ensure that synthetic recombination remains statistically valid when applied to datasets exhibiting longer temporal coverage or differing storm climatology.

6 Conclusion

The results highlight the critical importance of temporal coherence for accurately estimating extreme TWL and coastal flooding hazards. This is consistent with recent extreme value studies demonstrating that neglecting intra-event temporal dependence can lead to systematic misrepresentation of compound extremes (Ali et al. 2025). In particular, approaches based on static maxima or rescaled historical trajectories have been shown to distort extremal dependence pathways and derived response variables, even when marginal extremes are well represented (Rinaldo et al. 2020; Tendijck et al. 2023). By preserving the relative timing of storm components through temporally coherent reconstruction, the MSTM–TE framework avoids these deficiencies and yields more physically consistent estimates of design-level TWL. For conditions typical of Guadeloupe, the effective dimensionality of sea-state extremes is low: the joint behavior of waves and surge collapses onto two dominant extremal axes, with wave run-up acting as the primary driver in most events. This is reflected in the consistent alignment between peak run-up and peak TWL across both observed and simulated storms. At the same time, the analysis demonstrates that this alignment can break down in locations with reduced wave exposure, where surge can interact with moderate run-up to produce extreme responses whose timing cannot be inferred from either component alone. These findings emphasize that robust design-level load estimation must account simultaneously for the dominant contribution of wave run-up, the full temporal evolution of storm components, and spatial variability in wave energy that modulates how these components combine.

Code availability. MATLAB code for the analysis is provided on GitHub at **?**.

Data availability. Sample wave data for the analysis are available on Zenodo at **?**.

Author contributions. RW, JR and PhJ originally conceived the study. YK and JR conducted cyclone and wave modelling. RW and PhJ worked on statistical methodology and modelling. All authors contributed to writing the paper.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. Numerical simulations were conducted using the computational resources of the C3I (Centre Commun de Calcul Intensif) in Guadeloupe. Jeremy Rohmer acknowledges the funding of the Carib-Coast Interreg project (https://www.interreg-caraibes.fr/carib-coast, number: 2014TC16RFTN008) as well as the BRGM-funded PDEV MODLIT. MATLAB code for the analysis is provided on GitHub at **?**. Sample wave data for the analysis are available on Zenodo at **?**.

References

Ali, J., Wahl, T., Morim, J., Enriquez, A., Gall, M., and Emrich, C. T.: Multivariate compound events drive historical floods and associated losses along the U.S. East and Gulf coasts, npj Natural Hazards, 2, https://doi.org/10.1038/s44304-025-00076-5, 2025.

Anderson, D. L., Ruggiero, P., Mendez, F. J., Barnard, P. L., Erikson, L. H., O'Neill, A. C., Merrifield, M., Rueda, A., Cagigal, L., and Marra, J.: Projecting climate dependent coastal flood risk with a hybrid statistical dynamical model, Earth's Future, 9, https://doi.org/10.1029/2021ef002285, 2021.

Avila, L. A., Stewart, S. R., Berg, R., and Hagen, A. B.: Tropical cyclone report: Hurricane Dorian (AL052019), 24 August – 7 September 2019, Tech. Rep. AL052019, National Hurricane Center, Miami, Florida, available at https://www.nhc.noaa.gov/data/tcr/AL052019_Dorian.pdf, 2020.

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 amp; Environment, 1, https://doi.org/10.1038/s43247-020-00044-z, 2020.

Coles, S.: An introduction to statistical modeling of extreme values, Springer London, https://doi.org/10.1007/978-1-4471-3675-0, 2001.

Dey, A. K. and Patwary, M. S. A.: Utilizing Non-Stationary Extreme Value Model to Quantify Extreme Rainfall in Two Major Cities in Bangladesh, Stochastic Environmental Research and Risk Assessment, 39, 2281–2296, https://doi.org/10.1007/s00477-025-02969-3, 2025.

Eastoe, E. F. and Tawn, J. A.: Modelling Non-Stationary Extremes with Application to Surface Level Ozone, Journal of the Royal Statistical Society Series C: Applied Statistics, 58, 25–45, https://doi.org/10.1111/j.1467-9876.2008.00638.x, 2008.

Elahi, M. W. E., Wang, X. H., Salcedo-Castro, J., and Ritchie, E. A.: Influence of wave–current interaction on a cyclone-induced storm surge event in the Ganges–Brahmaputra–Meghna Delta: Part 1—effects on water level, Journal of Marine Science and Engineering, 11, 328, https://doi.org/10.3390/jmse11020328, 2023.

Fritz, H. M., Blount, C. D., Thwin, S., Thu, M. K., and Chan, N.: Cyclone Nargis storm surge in Myanmar, Nature Geoscience, 2, 448–449, https://doi.org/10.1038/ngeo558, 2009.

Habereder, I., Kneib, T., Echizen, I., and Spinde, T.: A systematic review of spatio-temporal statistical models: theory, structure, and applications, https://doi.org/10.48550/ARXIV.2511.00422, 2025.

Hallegatte, S., Green, C., Nicholls, R. J., and Corfee-Morlot, J.: Future flood losses in major coastal cities, Nature Climate Change, 3, 802–806, https://doi.org/10.1038/nclimate1979, 2013.

Heffernan, J. E. and Tawn, J. A.: A conditional approach for multivariate extreme values (with discussion), Journal of the Royal Statistical Society Series B: Statistical Methodology, 66, 497–546, https://doi.org/10.1111/j.1467-9868.2004.02050.x, 2004.

Kennedy, A. B., Gravois, U., Zachry, B. C., Westerink, J. J., Hope, M. E., Dietrich, J. C., Powell, M. D., Cox, A. T., Luettich, R. A., and Dean, R. G.: Origin of the Hurricane Ike forerunner surge, Geophysical Research Letters, 38, https://doi.org/10.1029/2011gl047090, 2011.

Laigre, T., Balouin, Y., Nicolae-Lerma, A., Moisan, M., Valentini, N., Villarroel-Lamb, D., and De La Torre, Y.: Seasonal and Episodic Runup Variability on a Caribbean Reef-Lined Beach, Journal of Geophysical Research: Oceans, 128, https://doi.org/10.1029/2022jc019575, 2023.

Melet, A., Meyssignac, B., Almar, R., and Le Cozannet, G.: Under-estimated wave contribution to coastal sea-level rise, Nature Climate Change, 8, 234–239, https://doi.org/10.1038/s41558-018-0088-y, 2018.

Menéndez, M. and Woodworth, P. L.: Changes in extreme high water levels based on a quasi-global tide-gauge data set, Journal of Geophysical Research: Oceans, 115, https://doi.org/10.1029/2009jc005997, 2010.

Mori, N., Kato, M., Kim, S., Mase, H., Shibutani, Y., Takemi, T., Tsuboki, K., and Yasuda, T.: Local amplification of storm surge by Super Typhoon Haiyan in Leyte Gulf, Geophysical Research Letters, 41, 5106–5113, https://doi.org/10.1002/2014gl060689, 2014.

Rappaport, E. N.: Fatalities in the United States from Atlantic tropical cyclones: new data and interpretation, Bulletin of the American Meteorological Society, 95, 341–346, https://doi.org/10.1175/bams-d-12-00074.1, 2014.

Rinaldo, T., Ramakrishnan, K. A., Rodriguez-Iturbe, I., and Durán Vinent, O.: Probabilistic structure of events controlling the after-storm recovery of coastal dunes, Proceedings of the National Academy of Sciences, 118, https://doi.org/10.1073/pnas.2013254118, 2020.

Roelvink, D., Reniers, A., van Dongeren, A., van Thiel de Vries, J., McCall, R., and Lescinski, J.: Modelling storm impacts on beaches, dunes and barrier islands, Coastal Engineering, 56, 1133–1152, https://doi.org/10.1016/j.coastaleng.2009.08.006, 2009.

Sando, K., Wada, R., Rohmer, J., and Jonathan, P.: Multivariate spatial and spatio-temporal models for extreme tropical cyclone seas, Ocean Engineering, 309, 118 365, https://doi.org/10.1016/j.oceaneng.2024.118365, 2024.

Senechal, N., Coco, G., Bryan, K. R., and Holman, R. A.: Wave runup during extreme storm conditions, Journal of Geophysical Research: Oceans, 116, https://doi.org/10.1029/2010jc006819, 2011.

Shankar, N. and Jayaratne, M.: Wave run-up and overtopping on smooth and rough slopes of coastal structures, Ocean Engineering, 30, 221–238, https://doi.org/10.1016/s0029-8018(02)00016-1, 2003.

Shi, F., Tehranirad, B., and Kirby, J. T.: FUNWAVE-TVD fully nonlinear Boussinesq wave model with TVD solver (Version 2.1), Tech. rep., Center for Applied Coastal Research, University of Delaware, https://www1.udel.edu/kirby/papers/shi-etal-cacr-11-04-version2.1.pdf, 2013.

Stockdon, H. F., Holman, R. A., Howd, P. A., and Sallenger, A. H.: Empirical parameterization of setup, swash, and runup, Coastal Engineering, 53, 573–588, https://doi.org/10.1016/j.coastaleng.2005.12.005, 2006.

Tendijck, S., Jonathan, P., Randell, D., and Tawn, J.: Temporal evolution of the extreme excursions of multivariate kth order Markov processes with application to oceanographic data, Environmetrics, 35, https://doi.org/10.1002/env.2834, 2023.

van der Meer, J. W. and Stam, J. M.: Wave runup on smooth and rock slopes of coastal structures, Journal of Waterway, Port, Coastal, and Ocean Engineering, 118, 534–552, https://doi.org/10.1061/(ASCE)0733-950X(1992)118:5(534), 1992.

Wahl, T., Haigh, I. D., Nicholls, R. J., Arns, A., Dangendorf, S., Hinkel, J., and Slangen, A. B. A.: Understanding extreme sea levels for broad-scale coastal impact and adaptation analysis, Nature Communications, 8, https://doi.org/10.1038/ncomms16075, 2017.

Walcker, R., Laplanche, C., Herteman, M., Lambs, L., and Fromard, F.: Damages caused by hurricane Irma in the human-degraded mangroves of Saint Martin (Caribbean), Scientific Reports, 9, https://doi.org/10.1038/s41598-019-55393-3, 2019.

Wei, G., Kirby, J. T., and Sinha, A.: Generation of waves in Boussinesq models using a source function method, Coastal Engineering, 36, 271–299, https://doi.org/10.1016/s0378-3839(99)00009-5, 1999.

World Meteorological Organization: State of the global climate 2023, Tech. rep., World Meteorological Organization, Geneva, Switzerland, wMO, State of the Global Climate 2023 report, 2023.

Yue, S., Pilon, P., and Cavadias, G.: Power of the Mann–Kendall and Spearman's rho Tests for Detecting Monotonic Trends in Hydrological Series, Journal of Hydrology, 259, 254–271, https://doi.org/10.1016/s0022-1694(01)00594-7, 2002.