

Storm surge time series de-clustering using correlation analysis

Ariadna Martín ^{a,b,*}, Thomas Wahl ^{a,b}, Alejandra R. Enriquez ^{a,b,c}, Robert Jane ^{a,b}

^a Department of Civil, Environmental and Construction Engineering & National Central for Integrated Coastal Research, University of Central Florida, Orlando, USA

^b National Center for Integrated Coastal Research, University of Central Florida, Orlando, FL, 32816, USA

^c Institute for Environmental Studies (IVM), Vrije Universiteit Amsterdam, 1081, HV Amsterdam, the Netherlands

ARTICLE INFO

Keywords:

de-clustering

Time series

Independent events

Storm surge

ABSTRACT

The extraction of individual events from continuous time series is a common challenge in many extreme value studies. In the field of environmental science, various methods and algorithms for event identification (de-clustering) have been applied in the past. The distinctive features of extreme events, such as their temporal evolutions, durations, and inter-arrival times, vary significantly from one location to another making it difficult to identify independent events in the series. In this study, we propose a new automated approach to detect independent events from time series, by identifying the standard event duration across locations using event correlations. To account for the inherent variability at a given site, we incorporate the standard deviation of the event duration through a soft-margin approach. We apply the method to 1 485 tide gauge records from across the global coast to gain new insights into the typical durations of independent storm surges along different coastline stretches. The results highlight the effects of both local characteristics at a given tide gauge and seasonality on the derived storm durations. Additionally, we compare the results obtained with other commonly used de-clustering techniques showing that these methods are more sensitive to the chosen threshold.

1. Introduction

Extreme value theory is based on the hypothesis that all events are independent and identically distributed (iid). When working with sequential data, such as time series of environmental phenomena, where events can cluster in time, serial dependence is almost inevitable, violating the iid assumption. To satisfy the assumption of independence, the time series must be de-clustered. There is, however, no universal definition of independence between events, and improper sampling from the time series can lead to the misrepresentation of associated hazards.

Bock Maxima (BM) and Peaks Over Threshold (POT) are the two most used approaches to sample extreme events. BM is based on the selection of the maximum value within non-overlapping time intervals, where the time intervals must be long enough to ensure independence between events (e.g., annual, or monthly maxima). The BM approach, however, may result in the loss of valuable information; for instance, when dealing with short time series the number of extreme events is limited. In addition, within a single time block (e.g., one year), there could be two or more events that are larger than the maxima in other

blocks, thus neglecting relevant information (Arns et al., 2013). The POT method overcomes this problem, allowing the sampling of a larger number of extreme events, yet it raises the question of how to properly de-cluster the observations to ensure independence.

Different methodologies have been proposed over the years to ensure independence when applying the POT method. The most used in hydrologic and coastal applications are: (1) the runs de-clustering method, (2) the use of a standard event duration (abbreviated hereafter as SED), and (3) the extremal index. Method (1) defines a new event to have begun once there is a specified number (run length) of consecutive observations below the threshold. The maximum observation is then extracted for each event (Leadbetter et al., 1989; Acero et al., 2011). Method (2) uses a SED (also referred to as “Standard Storm Length” in storm analysis) (Tawn, 1988). In essence, it uses a specified event duration in such a way that any exceedance beyond this is considered independent from the preceding event (Mathiesen et al., 2010; Ward et al., 2018; Zachary et al., 1998). Both methods require a predefined threshold as well as the run length and the SED (for Method 1 and Method 2, respectively) which in many cases is chosen subjectively (Ferro and Segers, 2003), though it has been shown that it plays an

* Corresponding author. Department of Civil, Environmental and Construction Engineering & National Central for Integrated Coastal Research, University of Central Florida, Orlando, USA.

E-mail address: ariadna.martinoliva@ucf.edu (A. Martín).

<https://doi.org/10.1016/j.wace.2024.100701>

Received 18 January 2024; Received in revised form 5 April 2024; Accepted 30 May 2024

Available online 1 June 2024

2212-0947/© 2024 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

important role in the de-clustering results (Arns et al., 2013; Vanem, 2015). Additionally, these methods do not consider spatial variability and assume equally sized event durations (natural variability) (Arns et al., 2013).

To overcome this issue, Ferro and Segers (2003) proposed the use of the extremal index as an automated way of de-clustering extreme events without forcing a fixed parameter during the process (Method (3)). This method measures the degree of clustering of stationary extreme values as the proportion of inter-exceedance times that may be regarded as inter-cluster times. The hypothesis is that extreme events cluster together in a compound Poisson process (i.e., extremes cluster in time, therefore, they do not occur randomly as expected from a Poisson distribution; however, clusters do occur randomly). This method has been extensively used (Poon et al., 2004; Rueda et al., 2016), especially when sampling events with low annual exceedance probabilities from the data. However, the results are highly dependent on the threshold value (Della-Marta et al., 2009), which must be high enough to ensure asymptotic behavior.

Sequential events are getting more attention, especially in environmental science (Besio et al., 2017; Telesca, 2007; Wadey et al., 2014; Xi et al., 2023). For example, consecutive storms near the coast can lead to compounding effects, e.g., communities recovering from a shock are more vulnerable to future shocks than they otherwise would be. These consecutive events may all be extreme, or they could include multiple moderate events (with higher annual exceedance probabilities), but in any case, the clustering leads to more devastating impacts than isolated extreme events (de Ruiter et al., 2020; Zscheischler et al., 2020). For example, Hurricanes Irma and Maria caused strong winds, intense precipitation, and storm surges with associated impacts in September 2017 in Puerto Rico within less than two weeks of each other. Another example of consecutive events was the remarkable sequence of storms that struck the United Kingdom (UK) during the winter of 2013/2014. From December to February a storm arrived on average every 2.5 days (Jenkins et al., 2022) and not all of them were equally extreme; in fact, some were moderate but still caused significant impacts as flood protection measures were already weakened. More recently, Hurricane Ian struck the coast of Florida and became the costliest natural disaster of the 2023 Atlantic Hurricane season. Only 45 days later, a weaker storm, Hurricane Nicole, with a much smaller storm surge destroyed infrastructure and beach property along the Florida east coast, where beaches and dunes were already eroded from Ian. In certain circumstances, having consecutive events can lead to less detrimental impacts. For instance, when tropical cyclones occur in closer succession, coral reefs have more time to recover between events and remain in a healthy state for longer than when the cyclones occurred randomly in time (Mumby et al., 2011). These examples highlight the importance of sampling the continuum from moderate to extreme events from underlying time series. However, the methods outlined above were developed for the most extreme events, ignoring that more moderate events, especially when occurring in close succession, can also lead to large impacts.

In this paper, we propose a new automated and more objective method of identifying independent events (extending Method (2) outlined above) by defining the SED based on the correlation between events in the same time series. As pointed out by Soukissian (2011), the consideration of a unique SED for the entire population is often debatable since different events exhibit different temporal evolutions and durations because of natural variability. As mentioned previously, most of the approaches found in the literature do not consider local variability and assume equally sized event durations or fixed parameters (Arns et al., 2013). To overcome this challenge and based on the concept of soft margins (Cortes et al., 1995), we propose an extra step in the de-clustering process, allowing the events to be part of mixed populations, i.e., we allow the events from the same time series to have different durations. This concept introduces a degree of freedom to the de-clustering window, enabling events to be merged instead of (wrongly) being considered as individual events.

In the following, we first develop the new correlation-based framework for time series event de-clustering. Then, we apply it to storm surge data from 1 485 tide gauges located along the global coasts, offering valuable insights into the spatial variability of derived window lengths for de-clustering purposes. Next, we explore the implications of time series non-stationarity (decadal and seasonal variations) in the de-clustering process. Finally, we compare our results with those derived from other widely used methods.

2. Data

Sea level data are obtained from tide gauge observations from the Global Extreme Sea Level Analysis database - version 3 (GESLA-3) (Haigh et al., 2022; Woodworth et al., 2016), <https://gesla.org/>. Most of the time series have hourly data but some have higher temporal resolution; those were interpolated to hourly to ensure consistency. Next, we remove the influence of mean sea level variation by subtracting the annual mean sea level time series from each record. Then, the astronomical tide is calculated from the detrended sea level records using the MATLAB Unified Tidal Analysis and Prediction Functions (U-Tide, Codiga, 2023) on a year-by-year basis, for years with more than 70% of data. We kept the seasonality of the series (i.e., SA and SSA tidal constituents were set to zero in the tidal analysis); this did not affect our overall results in terms of the derived SED values. We consider the resulting non-tidal residual time series (referred to here as storm surge) for our analysis.

From a total of 5 119 tide gauges included in GESLA-3, we select a subset of 1 485 time series based on the following criteria:

1. Coastal tide gauges: GESLA-3 includes river, lake, and coastal tide gauges. Only coastal tide gauges were included in the analysis.
2. Record length and completeness: to ensure robust results from the tidal analysis only years with at least 70% data completeness are analyzed. In addition, we only use stations with record lengths of more than 10 years. A total of 1 920 stations remain after applying these two filters.
3. Quality check: after removing the tidal signal, storm surge time series are visually inspected and physically unrealistic jumps are removed. Also, tide gauges with visible changes in the datum were removed, leaving 1 629 tide gauges.
4. Duplications: for some stations, the same water level information is retrieved from different providers, creating duplications. In such cases, we retain the longest record available. After removing duplicates, 1 485 time series remain for the final analysis.

3. Methods

We propose a two-step approach to identify independent events in time series. The first step is an updated version of Method 2 (use of a SED). While in most cases the definition of the SED is the same for all the time series (i.e., for storm surges usually 3 days), we propose the use of a site-specific SED, thus accounting for local characteristics such as bathymetry, coast orientation, storm type, etc. To find the SED for each time series, we extended an approach outlined in (Tawn, 1988) and use the correlation between events. A shortcoming of most approaches to de-cluster time series is the assumption of a uniform SED value for all events within a series, disregarding potential natural variability (i.e., different types of storms affecting that area, seasonal changes, changes in the wind orientation, etc.). To address this, we introduce a soft margin (as in Cortes et al., 1995) in the second step of our proposed method. In the following we outline how the SED and soft margins are computed and how the method is implemented, exemplarily, to identify independent storm surge events from tide gauge time series.

Fig. 1 shows how the SED and soft margin values are derived based on a storm surge time series from a single tide gauge. To help characterize the average duration of the events at each location, we select

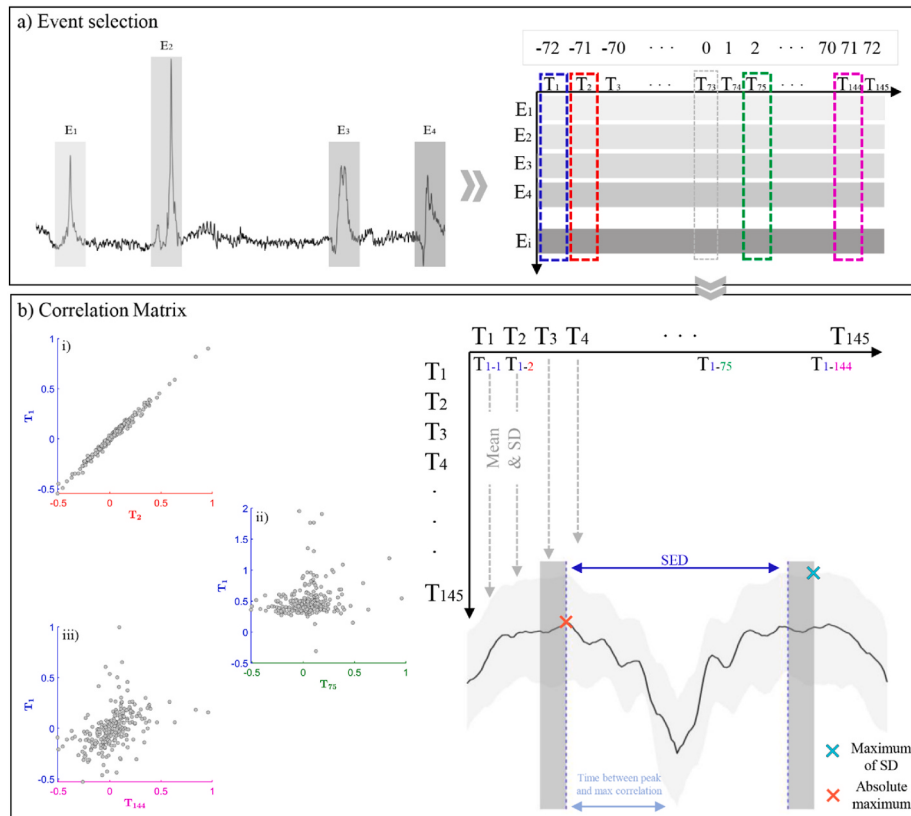


Fig. 1. Scheme of the method using the Apalachicola tide gauges as example. Panel a) illustrates the event selection process and the construction of the event matrix. Each row in the matrix represents an event, and the columns denote the time steps. Panel b) shows the construction of the correlation matrix used to derive the average correlation function (black line) and standard deviation (grey light area). The crosses on the correlation function indicate points of interest for the method. The color scheme in both panels a) and b) corresponds to the scatter plots on panel b) (i, ii and iii). These scatter plots illustrate the values used to calculate the correlation between specific time steps, each scatter plot corresponds to one cell of the correlation matrix.

events above the 99th percentile (based on Wahl et al. (2017) and others) and a window wide enough to capture all possible event durations (in our case we use 6 days, double than the usual 3-day independence window). These parameters, while not pivotal to the final SED selection, serve as an initial step to characterize typical storm surge behavior at a location. Note that these parameters can be changed depending on the type of event and time series used. We found that, in our case, the method is not very sensitive to changes in these parameters; we tested the method using lower thresholds (98th and 98.5th) and found the same values of the parameters in more than half (55% and 64 %, respectively) of the stations. At tide gauges where the SED is not the same differences reach 40% on average for both thresholds. Including the soft margin (measuring the difference of the SED + soft margin value) reduces the average difference to 20%. This shows that while the method is not entirely independent of the threshold, (the effect on the model parameters and ultimately the selected events (from moderate to more extreme events) is small.

For each station, we identify the exceedances above the threshold, in descending order. For each of those exceedances, we capture the storm surge levels occurring 3 days before and after, which we will refer to as events hereinafter. Note, that storm surge levels belonging to one event, will no longer be included in any other event. Then, we build a matrix consisting of the 6-day time series of those events, with the peaks centered at time zero (see Fig. 1a). We obtain a matrix where the number of rows matches the number of identified events, and each row has 145 columns including information about the temporal evolution of the storm surge during the 6 days (145 h, or time steps). We then calculate the correlation between all time steps, obtaining a symmetric 145×145 matrix. The scatter plots in Fig. 1b exemplify this process; each point within the scatter plot represents the time steps being correlated on each

cell of the matrix. For example, panel i) shows all events on T1 (blue) vs T2 (red) that would correspond to the T_{1-2} cell on the matrix, and so forth. The correlation function (black line in Fig. 1b) is then obtained taking the mean of each column (i.e. mean for each time step across all events). We also derive the standard deviation (light grey area in Fig. 1b). The SED and soft margins are computed based on the shape of the resulting correlation function. Overall, a minimum correlation is found during the peak of the event, while correlation levels increase as the surge levels approach the mean regime before and after the peak (Fig. 1b shows an example for the Apalachicola tide gauge, in Florida, see also Fig. 2f–j).

The SED (blue line in Figs. 1b and 2f–j) is defined as twice the difference between the peak of the event (minimum of correlation) and the time step at which the correlation function reaches an absolute maximum (orange cross in Fig. 1b), ensuring a symmetric independence window. This signifies the moment at which all time steps are most similar to both the extreme and the “regular” surge values, indicating a transition from regular (at the sides of the function) to extreme conditions (center). Other selection techniques can be used (i.e. “change point” or the coefficient of variation), but we found the absolute maximum to be more conservative and to better represent the physical aspect of the SED.

Individual events show slightly different event durations; the soft margin helps capture this variability across events for a single site. The soft margin is computed as the absolute difference (in hours) between the SED (dashed blue lines in Figs. 1b and 2f–j) and the time step of maximum standard deviation (blue cross in Fig. 1b). The soft margin is always considered outside the SED limit margins (grey-shaded areas in Figs. 1b and 2b f–j following the SED limits, dashed blue lines).

Fig. 3 shows an example of the method implementation to de-cluster

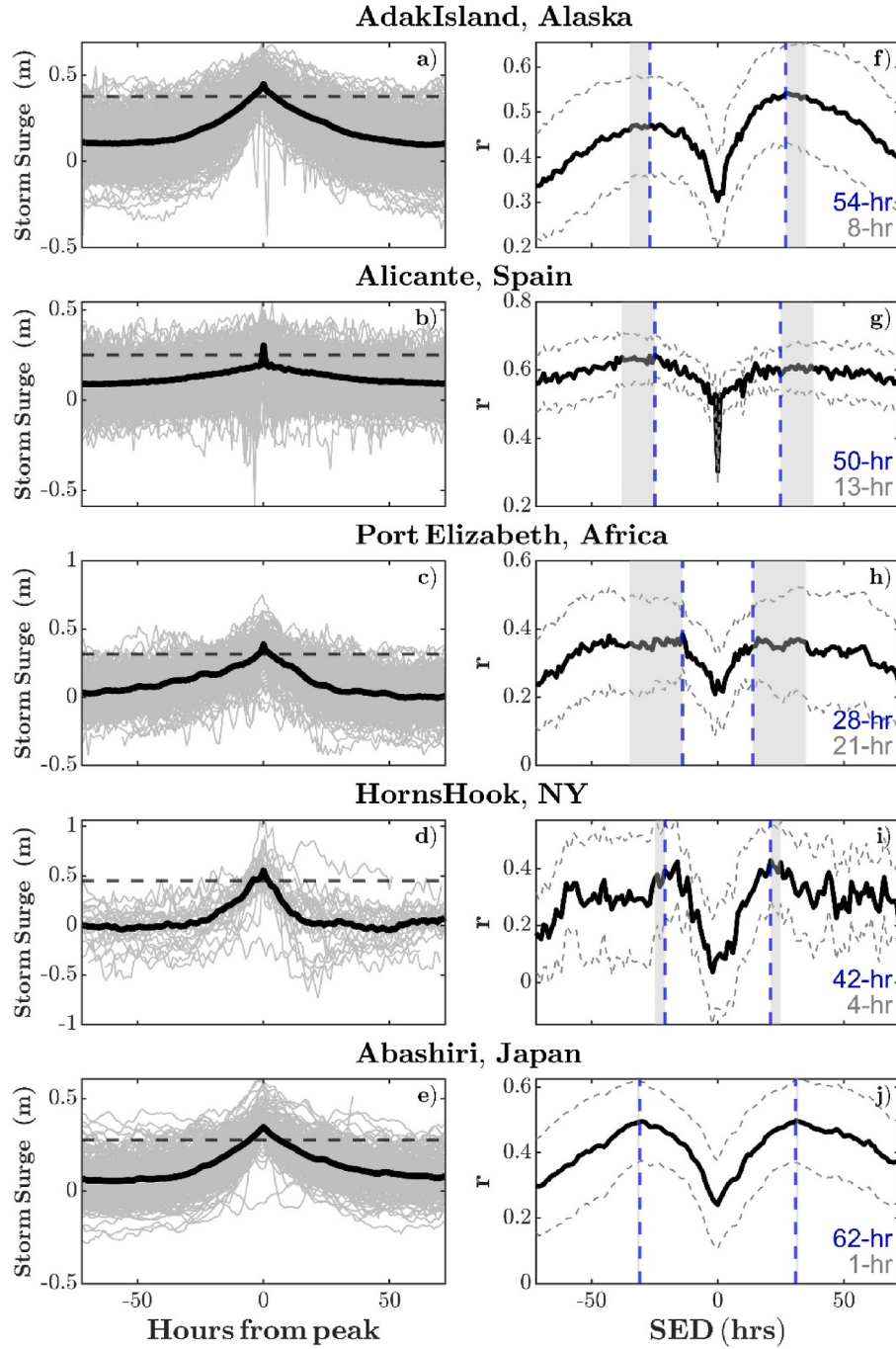


Fig. 2. Panels on the left show 6-day periods of storm surge time series at selected tide gauges (indicated in the title) with events above the 99th percentile (dark grey dashed line) centered at time zero. Black line represents the mean of all events at that tide gauge. The right-side panels show the average correlation function (black line) and standard deviation (grey dashed lines) of those events (on a 6-day window). Dashed blue line marks the SED, and the soft margin is marked with the grey shaded area (see text for explanation).

the time series at Aburatsu, Japan for events above the 1-yr RL (note that from now on we use return levels as thresholds). Using the events identified after applying the proposed method, we fit a Generalized Pareto Distribution to obtain the return levels. The first step consists of de-clustering the time series using Method 2, with a SED of 58 h (for this site), instead of the usual 3-day window. The blue dots represent the initial independent events identified based on the estimated SED (1st step).

The second step is applied to the (blue) events that fall within the soft margin boundaries (grey area). This is, if the difference between the beginning of one event (b) and the end of the previous event (e) is less

than the soft margin found for that time series (see the rule below), these events are merged and considered part of the same event, regardless of whether it has values below the threshold or not, allowing for increased variability in event durations. Both the SED and the soft margin are applied on an event basis. This results in the final de-clustered time series (red dots).

$$\begin{cases} b - e > \text{soft margin} & \text{remain independent events} \\ b - e \leq \text{soft margin} & \text{merged into a single event} \end{cases}$$

Panels b) and c) in Fig. 3 show two examples of this situation, where events longer than the SED incorrectly span more than one

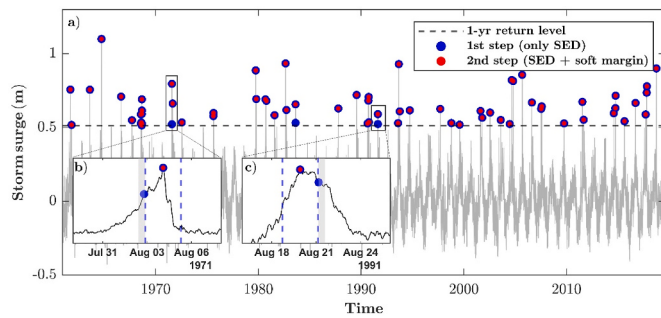


Fig. 3. Example of de-clustered time series. a) Complete time series shown in grey, blue dots mark the peaks found after applying the first step of the de-clustering (SED), red dots are the final independent events found after applying the second step (soft margin), for events above the 1-yr return level. b) and c) Show two instances where events are merged after applying the soft margin. Dashed blue lines mark the 58-hr SED and the grey lines are the 11-hr soft-margins.

"independent" event. The soft margin (in this case 11 h) serves as a corrective measure, ensuring the identification of true independent storms, as depicted by the red dots in Fig. 3 (second step of our method).

4. Results

We apply the method to 1 485 storm surge time series (see Sec. 2). First, we derive SED and soft-margin values at each station (Sec. 4.1). Then, we perform a non-stationarity analysis of the SED (Sec. 4.2). Lastly, we compare our results with other common de-clustering methods (Sec. 4.3).

4.1. Global implementation of the new framework

Fig. 4 shows the values of SED (panel a) and the soft margins (panel b) for all 1 485 tide gauges. The mean value for the SED across all stations is 65 h (the median is 62 h), which is similar to the 72 h commonly used to de-cluster storm surge data (Enríquez et al., 2020; Harter et al., 2022; Rashid et al., 2022). However, values notably vary spatially ranging from 20 to 140 h (Fig. 4a). Most tide gauges along the U.S. East Coast, the North Sea, and Japan have relatively longer SED values (~90

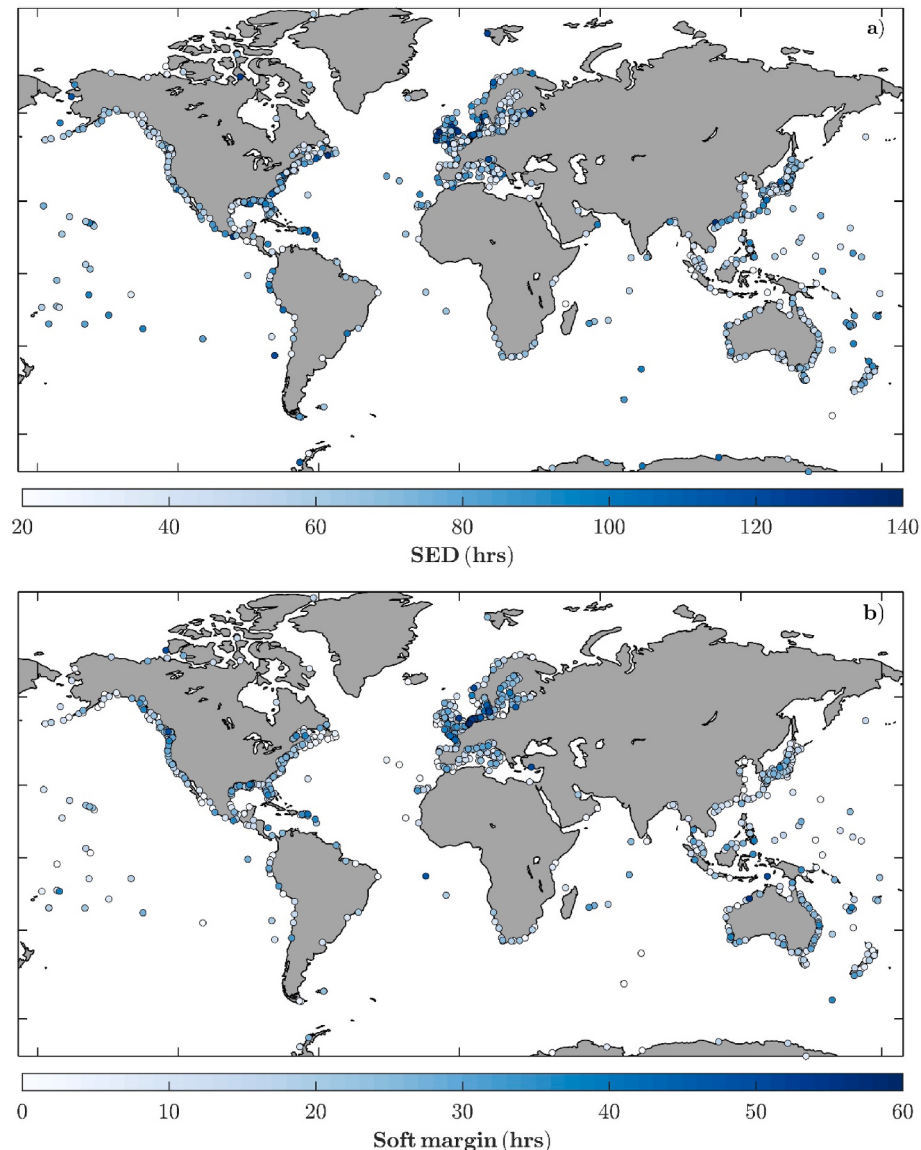


Fig. 4. Values of a) SED and b) soft margins for 1 485 tide gauges.

h) (see [Supplementary Fig. 1](#)), compared for example to Australia, Africa, or the U.S. West Coast where SED tends to be shorter (~ 50 h). In several instances, we also find large differences between nearby stations. We hypothesize that this is due to local characteristics such as bathymetry, coastline orientation, and/or tide gauge placement. [Fig. 5](#) compares mean normalized storm surges (and their SED) for nearby stations in three regions. Despite the short distance between tide gauges, the surge elevations differ across locations, resulting in longer SED values at the tide gauges where storm surges tend to be larger, taking more time to return to the mean regime.

[Fig. 4b](#) shows the values of the soft margin found for all the stations (with average values of 12 h). Values in the northern part of Europe stand out (~ 60 h), as well as in some parts of the U.S. coasts, indicating large variability in storm surge durations in these areas. This variability may be related to changes in storm patterns, as a result of seasonal or larger-scale changes. Other large values can be observed scattered across the global coast, especially in tide gauges located near bays or inlets where local effects such as freshwater inflow can lead to variability.

4.2. Non-stationarity analysis

So far in our analysis, as well as in previous studies, SED is assumed stationary, but changes in storminess can modulate those values. In this section, we study the evolution of the SED in time and how much it is influenced by seasonal to decadal variations. [Fig. 6](#) shows the effect of decadal variations on SED. We estimate the SED using 10-yr overlapping moving windows shifted 1-yr each time step and display the standard deviation of the non-stationary SED values. Some places such as the North Sea exhibit larger temporal variability, up to 66 h (33 h in each direction), while others like Japan are more consistent over time. The results indicate that the variability is not dependent on the record length (indicated by circle sizes in [Fig. 6](#)).

The stationary ([Fig. 4a](#)) and non-stationary SED values show little spatial coherence, indicating that they are more likely attributable to local conditions rather than being the result of global or regional climate patterns.

Seasonality in storminess, where different types of storms (tropical vs. extra-tropical) occur at different times of the year, can also influence SED. We derive the SED for the tropical season, defined as the North Atlantic hurricane season (from June 1st to November 30th), and for the extra-tropical season, i.e., the rest of the year. [Fig. 7a](#) shows the ratio between the number of events (above the 1-yr return level) in both seasons. The northern hemisphere experiences more events in the extra-tropical season during the boreal winter and the same happens in the southern hemisphere because of the storm season definition (note that we do not imply that events that occur during the tropical season in

those areas are actually tropical storm events, but rather split the year into two seasons). In lower latitudes, most of the events occur in the tropical season.

[Fig. 7b](#) shows the differences in the SED computed using events from the tropical and extra-tropical seasons (in absolute values). Different types of events can generate different SED values, for example, extra-tropical cyclones generally generate longer storm surges (SEDs in the tropical cyclone season are 4 h longer on average). Major differences can be noticed in south Japan, the Gulf of Mexico, the East Coast of the U.S., and northern Europe (see [Supplementary Fig. 2](#)). This indicates that the SED in those areas is highly dependent on the type of events that are considered. This variability is partially captured in our proposed method through the soft margin parameter; the soft margin ([Fig. 4b](#)) is larger in areas where seasonal differences are also larger (i.e., Gulf of Mexico, West Coast of Canada, and North Sea).

4.3. Comparison between methods

The two most used techniques in the literature for storm surge de-clustering are the extremal index (Method 3) ([Coles, 2001](#); [Poon et al., 2004](#)) and the use of a fixed 3-day window (Method 2) (e.g., [Enríquez et al., 2020](#); [Harter et al., 2022](#); [Rashid et al., 2022](#)). These methods have some inherent disadvantages, in particular when one is not only interested in the most extreme events, but also more moderate ones, which, if they occur in close succession, can still lead to large impacts.

Here we compare the results between the proposed method and these commonly used methods by applying the extremal index (Method 3) and a fixed 3-day window (Method 2) to the same 1 485 tide gauges. The differences in the identified number of events among the methods are larger at lower thresholds and decrease as higher thresholds are applied. For instance, when using a 1-yr return level threshold, the extremal index and the 3-day window identify, on average, 10.4 and 2.4 events more than the method proposed here ([Table 1](#)), respectively. When using a threshold equivalent to the 5-yr return level, differences decrease to 2.4 and 1.3 additional events found with the other two methods. Note that the thresholds used in all three methods are the same for direct comparison.

The differences vary spatially; at 1-yr return period (5-yr return period), differences in the number of events between our proposed method and Method 2 are found at 49% (8%) of the tide gauges, while the comparison against the extremal index reveals differences in approximately 89% (40 %) of the sites. In almost all those cases ($>90\%$) the extremal index identifies more events compared to the proposed method, this is not the case for Method 2 where the difference in the number of events can be positive or negative ([Supplementary Figs. 3 and 4](#) show the spatial distribution on the difference of events using the

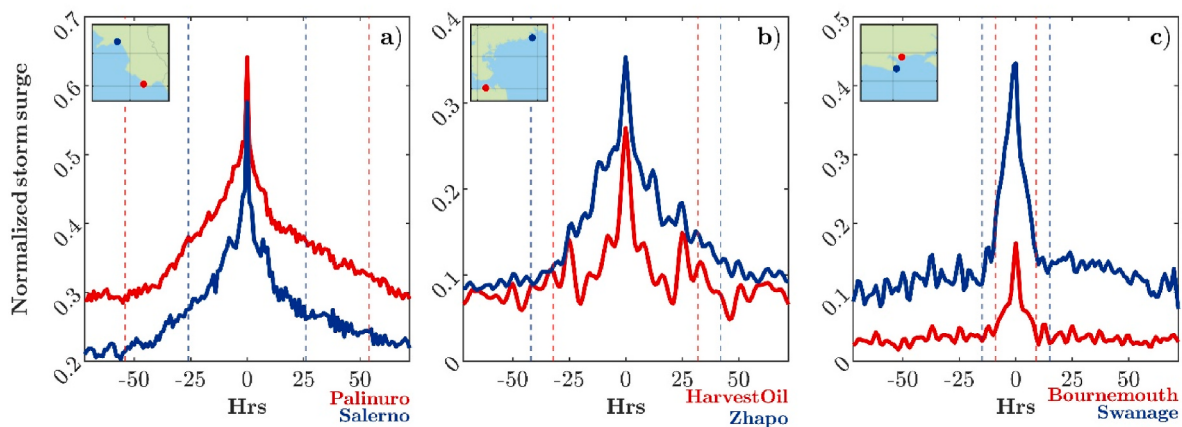


Fig. 5. Mean normalized storm surges for nearby tide gauges with different SED values located in a) Italy, b) Hong Kong, and c) the UK. The names of the stations are under each panel with the same font color. Dashed lines represent the SED values of each tide gauge. Insets in the panels show the locations of the tide gauge stations.

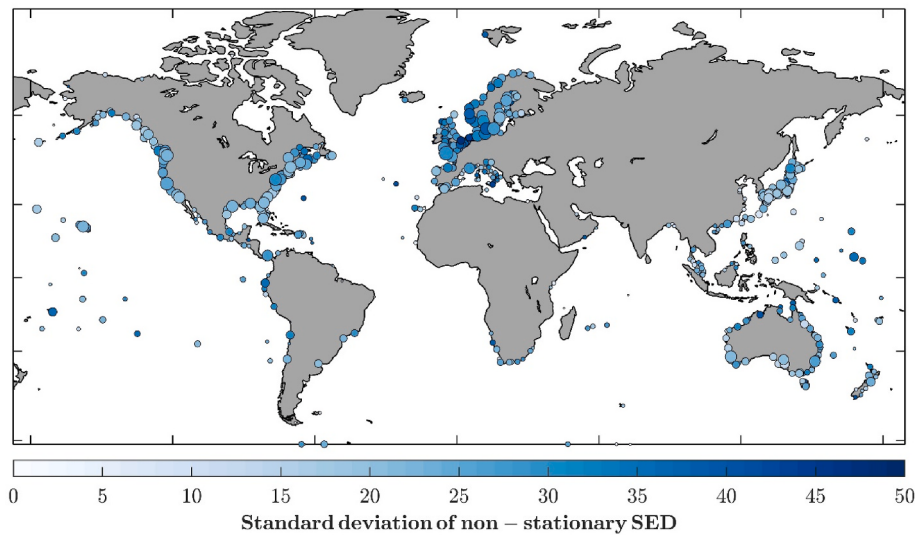


Fig. 6. Standard deviation of the non-stationary (decadal variations) SED values at each tide gauge; the size of the circles represents the record length (larger circles indicate longer records, the smaller being 10-yr).

proposed method versus the 3-day window and the extremal index, respectively).

Differences in the number of events can be due to the definition of the SED and/or the implementation of the soft margin. However, the soft margin is a corrective measure only used if the criteria explained in Sec. 3 is met, i.e., for events longer than the standard (SED) definition. For example, for the lowest threshold (1-yr RL) the soft margin is used at 668 tide gauges, and for those, a mean of four events are merged, which represents on average 12% of the events on the time series. As the threshold increases (5-yr RL) less sites are affected (130 tide gauges) with an average of 1.7 events merged, which corresponds approximately to 21% of the total events.

Fig. 8 shows the events identified by the three methods when applied to the time series of Saigo, Japan (the values found for this station are 66 h SED and 14 h of soft margin). Markers in Fig. 8 represent the independent events identified by each method. In Fig. 8b and c two instances are zoomed in to allow a closer examination of the differences. Fig. 8b shows that for a single event, the extremal index identifies four different events (blue triangles), while the 3-day window (black crosses) identifies two events, and our proposed method identifies only one event based on the highest peak. The same can be seen in Fig. 8c for three different events happening between August and September of 1984, where only the correlation method proposed here correctly identifies the peaks of independent events. Additional examples are shown in Supplementary Fig. 5.

5. Discussion

De-clustering time series is an important step when studying extreme events. An overview of the most used techniques shows that there are some disadvantages mostly due assuming constant de-clustering values along the globe as well as a tendency toward poor performance when considering more moderate rather than only very extreme events (see Table 1 and Supplementary Figs. 3 and 4). Here we propose a new method that reduces subjectivity compared to existing methods, using the correlation function to find the de-clustering parameters (SED and soft margin). Furthermore, unlike other methods, our proposed approach imposes no constraints on the level of extremeness of the events that can be included in the study, thereby enhancing its flexibility. The combination of SED and the soft margin allows different event durations within the same time series and provides a more objective and automated way to de-cluster many different types of

(environmental) time series.

To exemplify the procedure, we applied the method to storm surge time series from 1 485 tide gauge stations across the globe. The results are compared to those derived from applying the extremal index (Method 3) and the 3-day window (Method 2). The extremal index is identified as sub-optimal when interested in a broader range of events (moderate to extreme), since the threshold must be high enough to ensure asymptotic behavior. The 3-day window approach proved to be more effective in such cases, however, there are still difference in event counts when compared to the proposed method (Fig. S2). Therefore, the choice of the de-clustering method can also affect other applications such as the results of extreme value analysis. This effect diminishes when higher thresholds are used to include only the most extreme events.

The values of the SED parameter found in our analysis (Fig. 4) are in general agreement with values derived and/or used in previous local studies. Zachary et al. (1998), for example, used values between 24 and 72 h when analyzing wave and wind data in the North Sea. Mathiesen et al. (2010) used higher values between 120 and 168 h for higher latitudes in the northern hemisphere, where we also found relatively higher SED values. For a specific site in Lowestoft, UK, Tawn (1988) found a value of 30 h whereas we derive a SED of 38 h with a soft margin at this location of 5 h.

We also assess changes in the SED over time and show that some tide gauges, in particular in northern Europe, exhibit relatively strong decadal variability. In addition, seasonality (here associated with different storm types occurring during and outside the Atlantic Hurricane season) modulates the SED in many regions. Including the soft margin parameter in the de-clustering process can accommodate some of that variability. In essence, this parameter identifies the areas where it is less suitable to use a single constant de-clustering parameter (Arns et al., 2013), which at the same time leads to a more conservative method in those areas since the de-clustering window becomes longer. Therefore, when conducting localized studies, we recommend computing the relevant SED and soft margin for the time series de-clustering, from the data observed during the season of interest (e.g., when storms only occur in winter then only consider data from that period).

The results of applying the method to storm surge time series show an overall improvement over other traditionally used techniques, but there are still some limitations. It is very important to have reliable time series data that has gone through rigorous quality control. The proposed method can accommodate gaps in the time series offering the

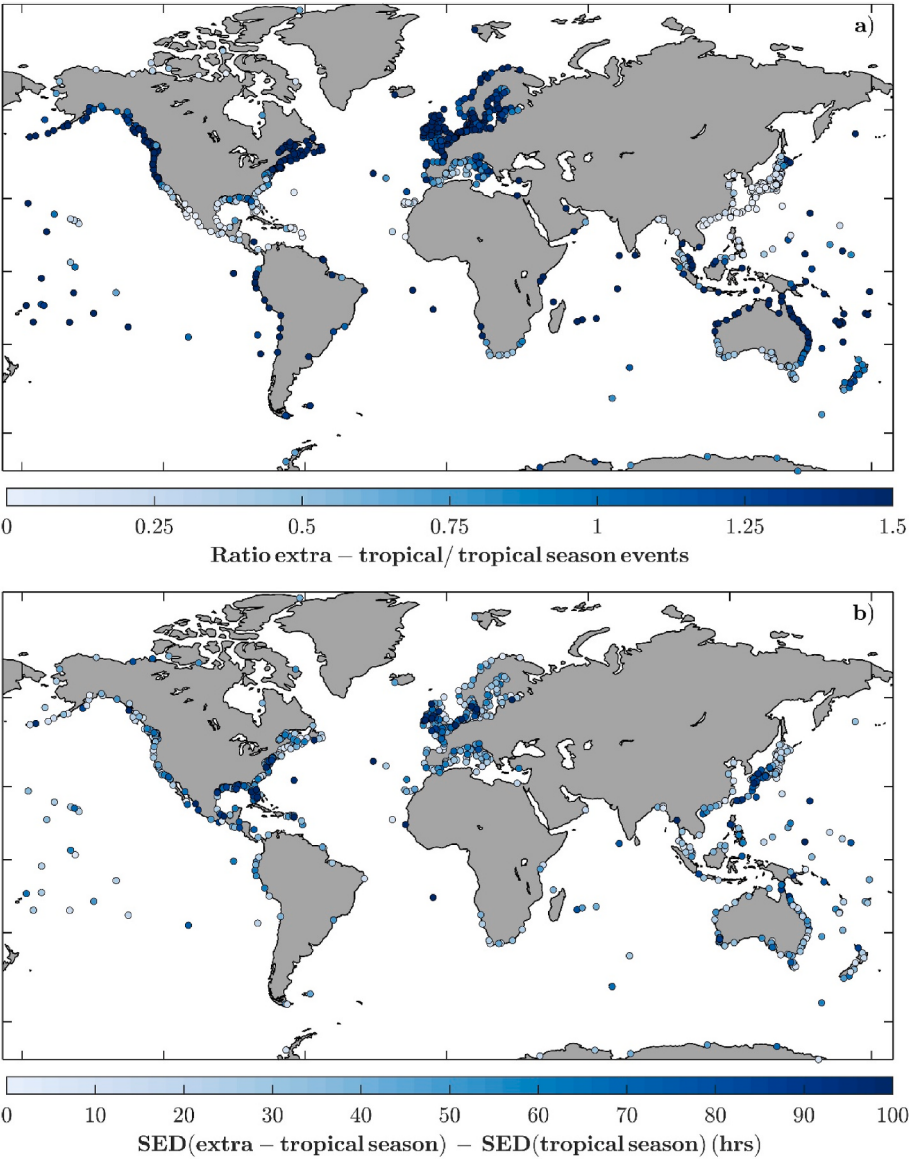


Fig. 7. Seasonal variability of the SED parameter. a) Ratio of extra-tropical vs tropical season events (tropical season defined as the north Atlantic hurricane season), for events above the 1-yr return level. b) Absolute difference between the tropical and extra-tropical SED.

Table 1
Comparison between methodologies using the proposed method as baseline, for 1 and 5-yr return level thresholds. The first row indicates the % of sites where differences across methodologies are found, while the second row shows the difference in the average number of independent events, the last row indicates the percentage of stations where other methods identified more events than the proposed method.

	Method 2: 3-day window		Method 3: Extremal index	
	1-yr RL	5-yr RL	1-yr RL	5-yr RL
% of tide gauges with different number of events	49.1	8.4	89.0	39.7
Average difference in the number of events compared to new method	2.4	1.3	10.7	2.4
% of tide gauges with more events identified compared to new method	73.5	46.4	98.8	99.0

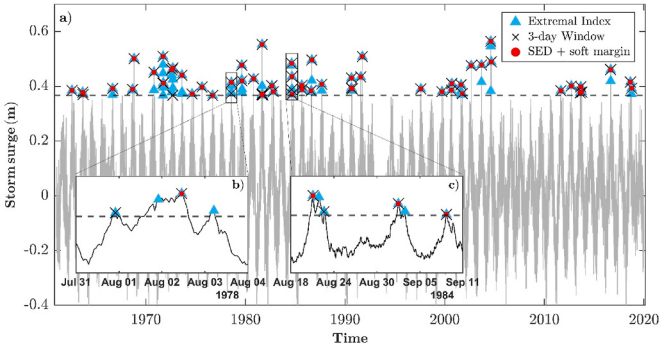


Fig. 8. a) Comparison of the de-clustered time series using the proposed method (SED + soft margin), the extremal index, and the 3-day window. b) and c) show details for different specific events.

opportunity to remove suspicious data, however, the presence of outliers can still generate issues in the correlation function. Another important limitation is that we are not accounting for the characteristics of each individual event but a SED; therefore, the method can still potentially falsely count a single event as two or merge two events into one. However, we introduce varying SED values across stations as well as a soft margin to account for variability in each time series making it more likely to correctly identify independent events. Other anomaly detection algorithms (Kalman filter, Random Forest regression, K-means, etc.) have been developed over the last decades for different applications including the identification of individual extreme events (Schmidl et al., 2022). Some of the most innovative algorithms based on deep learning, data mining, or stochastic learning are theoretically capable of identifying all events, however, they also require fine-tuning of the parameters and substantial amount of data. Schmidl et al. (2022) evaluated 71 anomaly detection algorithms falling into seven different categories (statistics, outlier detections, data mining, Classic Machine learning, etc.). They found that the most complex algorithms are very sensitive to the parameter settings, as well as time consuming and very demanding in terms of memory space. Consequently, the challenge lies in striking a balance between algorithmic complexity and practical feasibility.

6. Conclusion

We introduce a flexible de-clustering process that allows the inclusion of both moderate and extreme events that applies to different types of time series. It automatically identifies the optimal parameter settings (SED and soft margins) through an objective and streamlined implementation, without imposing restrictions on the level of extremeness of the events to be identified from the continuous time series. Adding the soft margin accounts for the natural variability within the time series.

Here, we apply the method to storm surge time series from tide gauges across the global coast. We do not find strong spatial coherence indicating that both SED and the soft margin are more impacted by local characteristics than the large-scale storm climate (Fig. 4). We assess non-stationarity of the parameter values related to decadal variability and seasonality and show that the effects are more pronounced in parts of northern Europe, south Japan, and the Gulf of Mexico (Figs. 6 and 7). The results highlight that local studies should consider these variations and focus, for example, only on the stormy season or only on the last few decades of data (when time series are very long) when identifying suitable values for SED and the soft margin; of course, it depends on the goal of the analysis.

To discern the reliability of the proposed method it is compared to other existing methods, the extremal index, and the 3-day fixed window, two of the most commonly applied methods in storm surge analysis. In both cases, a strong relation of the results with the chosen threshold was noticed. The extremal index is intended to identify only the most extreme events, which leads to larger differences in the number of identified events compared to our proposed method, when the threshold is relatively low to include moderate events. These differences decrease as the threshold increases. The same occurs with the 3-day window method but differences are generally smaller.

Overall, we show that the proposed method is robust even in the presence of variability, seasonality, and gaps in the data. The method, designed to mitigate subjectivity in time series analysis, eliminates restrictions related to event extremeness, enabling its application to both moderate and extreme events. While applied here to storm surge time series, the method is transferable to diverse time series datasets.

CRedit authorship contribution statement

Ariadna Martín: Writing – original draft, Formal analysis, Data curation, Conceptualization. **Thomas Wahl:** Writing – review & editing, Supervision, Project administration, Conceptualization. **Alejandra R. Enriquez:** Writing – review & editing, Writing – original draft, Data

curation. **Robert Jane:** Writing – review & editing, Formal analysis.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Ariadna Martín reports financial support was provided by National Science Foundation. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

All the data and code used in this study is publicly available and can be downloaded from a Zenodo data repository: <https://zenodo.org/records/10513063>, which also includes an interactive map and table with all the used stations as well as the model parameters (SED and soft margin) found for each tide gauge. The codes used to implement this method are also available at this GitHub repository: https://github.com/AriadnaMartin98/De-clustering_method.

Acknowledgments

This work was supported by the United States National Science Foundation (grant number: 2141461). A.R.E. acknowledges support by the European Union Horizon 2020 EXCELLENT SCIENCE – Marie Skłodowska-Curie Actions (grant number: 101019470).

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.wace.2024.100701>.

References

- Acero, F.J., García, J.A., Gallego, M.C., 2011. Peaks-over-Threshold study of trends in extreme rainfall over the Iberian Peninsula. *J. Clim.* 24 (4), 1089–1105. <https://doi.org/10.1175/2010JCLI3627.1>.
- Arns, A., Wahl, T., Haigh, I.D., Jensen, J., Pattiaratchi, C., 2013. Estimating extreme water level probabilities: a comparison of the direct methods and recommendations for best practise. *Coast. Eng.* 81, 51–66. <https://doi.org/10.1016/J.COASTALENG.2013.07.003>.
- Besio, G., Briganti, R., Romano, A., Mentaschi, L., De Girolamo, P., 2017. Time clustering of wave storms in the Mediterranean Sea. *Nat. Hazards Earth Syst. Sci.* 17 (3), 505–514. <https://doi.org/10.5194/nhess-17-505-2017>.
- Coles, S., 2001. An introduction to statistical modeling of extreme values. <https://doi.org/10.1007/978-1-4471-3675-0>.
- Cortes, C., Vapnik, V., Saitta, L., 1995. Support-vector networks. *Mach. Learn.* 20 (3), 273–297. <https://doi.org/10.1007/BF00994018>, 1995 20:3.
- Codiga, Daniel, 2023. UTide unified tidal analysis and prediction functions. MATLAB Cent. File Exchange. <https://www.mathworks.com/matlabcentral/fileexchange/46523-utide-unified-tidal-analysis-and-prediction-functions>.
- de Ruiter, M.C., Couason, A., van den Homberg, M.J.C., Daniell, J.E., Gill, J.C., Ward, P. J., 2020. Why we can no longer ignore consecutive disasters. In: *Earth's Future*, vol. 8. John Wiley and Sons Inc. <https://doi.org/10.1029/2019EF001425>, Issue 3.
- Della-Marta, P.M., Mathis, H., Frei, C., Liniger, M.A., Klein, J., Appenzeller, C., 2009. The return period of wind storms over Europe. *Int. J. Climatol.* 29 (3), 437–459. <https://doi.org/10.1002/JOC.1794>.
- Enriquez, A.R., Wahl, T., Marcos, M., Haigh, I.D., 2020. Spatial footprints of storm surges along the global coastlines. *J. Geophys. Res.: Oceans* 125 (9), e2020JC016367. <https://doi.org/10.1029/2020JC016367>.
- Ferro, C.A.T., Segers, J., 2003. Inference for clusters of extreme values. *J. Roy. Stat. Soc. B* 65 (2), 545–556. <https://doi.org/10.1111/1467-9868.00401>.
- Haigh, I.D., Marcos, M., Talke, S.A., Woodworth, P.L., Hunter, J.R., Hague, B.S., Arns, A., Bradshaw, E., Thompson, P., 2022. GESLA Version 3: a major update to the global higher-frequency sea-level dataset. *Geosci. Data J.* 00, 1–22. <https://doi.org/10.1002/GDJ3.174>.
- Harter, C., McLaughlin, P., Price, D., Carter, J., Maristany, L., 2022. Joint probability extremal analysis of combined storm surge and precipitation. *Ports 2022: Port Engineering - Papers from Sessions of the 16th Triennial International Conference 1*, 301–307. <https://doi.org/10.1061/9780784484395.030>.
- Jenkins, L.J., Haigh, I., Camus, P., Pender, D., Lamb, R., Trust, J., Kassem, H., 2022. The temporal clustering of storm surge, wave height, and high Sea Level exceedances around the UK coastline. <https://doi.org/10.21203/rs.3.rs-1412525/v1>.

- Leadbetter, M.R., Weissman, I., De Haan, L., Rootzen, H., 1989. On clustering of high values in statistically stationary series. Centre for Stochastic Processes. Department of Statistics, University of North Carolina at Chapel Hill, Tech. Rep. 253.
- Mathiesen, M., Goda, Y., Hawkes, P.J., Mansard, E., Martín, M.J., Peltier, E., Thompson, E.F., Van Vledder, G., 2010. Recommended practice for extreme wave analysis. <https://doi.org/10.1080/00221689409498691>.
- Mumby, Peter J., Vitolo, Renato, Stephenson, David B., 2011. Temporal clustering of tropical cyclones and its ecosystem impacts, 108 (43), 17626–17630. <https://doi.org/10.1073/pnas.1100436108>.
- Poon, S.H., Rockinger, M., Tawn, J., 2004. Extreme value dependence in financial markets: diagnostics, models, and financial implications. *Rev. Financ. Stud.* 17 (2), 581–610. <https://doi.org/10.1093/RFS/HHG058>.
- Rashid, M.M., Wahl, T., Nasr, A., Camus, P., Haigh, I.D., Rashid, M.M., Wahl, T., Nasr, A., Camus, P., Haigh, I.D., 2022. Leveraging Machine Learning for Predicting Compound Flooding Potential from Extreme Surges and Precipitation, vol. 2022. AGUFM. NH41A-03. <https://ui.adsabs.harvard.edu/abs/2022AGUFMNH41A..03R/abstract>.
- Rueda, A., Camus, P., Tomás, A., Vitousek, S., Méndez, F.J., 2016. A multivariate extreme wave and storm surge climate emulator based on weather patterns. *Ocean Model.* 104, 242–251. <https://doi.org/10.1016/J.OCEMOD.2016.06.008>.
- Schmidl, S., Wenig, P., Papenbrock, T., 2022. Anomaly detection in time series. *Proc. VLDB Endow.* 15 (9), 1779–1797. <https://doi.org/10.14778/3538598.3538602>.
- Soukissian, T.H., 2011. The effect of declustering in the r-largest maxima model for the estimation of HS-design values. *Open Ocean Eng. J.* 4 (1), 34–43. <https://doi.org/10.2174/1874835X01104010034>.
- Tawn, J.A., 1988. An extreme-value theory model for dependent observations. *J. Hydrol.* 101 (1–4), 227–250. [https://doi.org/10.1016/0022-1694\(88\)90037-6](https://doi.org/10.1016/0022-1694(88)90037-6).
- Telesca, L., 2007. Time-clustering of natural hazards. *Nat. Hazards* 40 (3), 593–601. <https://doi.org/10.1007/s1069-006-9023-z>.
- Vanem, E., 2015. Uncertainties in extreme value modelling of wave data in a climate change perspective. *J. Ocean Eng. Mar. Energy* 1 (4), 339–359. <https://doi.org/10.1007/S40722-015-0025-3/FIGURES/12>.
- Wadey, M.P., Haigh, I.D., Brown, J.M., 2014. A century of sea level data and the UK's 2013/14 storm surges: an assessment of extremes and clustering using the Newlyn tide gauge record. *Ocean Sci.* 10 (6), 1031–1045. <https://doi.org/10.5194/os-10-1031-2014>.
- Wahl, T., Haigh, I.D., Nicholls, R.J., Arns, A., Dangendorf, S., Hinkel, J., Slangen, A.B.A., 2017. Understanding extreme sea levels for broad-scale coastal impact and adaptation analysis. *Nat. Commun.* 8 (1), 1–12. <https://doi.org/10.1038/ncomms16075>, 2017 8:1.
- Ward, P.J., Couasnon, A., Eilander, D., Haigh, I.D., Hendry, A., Muis, S., Veldkamp, T.I.E., Winsemius, H.C., Wahl, T., 2018. Dependence between high sea-level and high river discharge increases flood hazard in global deltas and estuaries. *Environ. Res. Lett.* 13 (8), 084012 <https://doi.org/10.1088/1748-9326/AAD400>.
- Woodworth, P.L., Hunter, J.R., Marcos, M., Caldwell, P., Menéndez, M., Haigh, I., 2016. Towards a global higher-frequency sea level dataset. *Geosci. Data J.* 3 (2), 50–59. <https://doi.org/10.1002/GDJ3.42>.
- Xi, D., Lin, N., Gori, A., 2023. Increasing sequential tropical cyclone hazards along the US East and Gulf coasts. *Nat. Clim. Change* 13 (3), 258–265. <https://doi.org/10.1038/s41558-023-01595-7>, 2023 13:3.
- Zachary, S., Feld, G., Ward, G., Wolfram, J., 1998. Multivariate extrapolation in the offshore environment. *Appl. Ocean Res.* 20 (5), 273–295. [https://doi.org/10.1016/S0141-1187\(98\)00027-3](https://doi.org/10.1016/S0141-1187(98)00027-3).
- Zscheischler, J., Martius, O., Westra, S., Bevacqua, E., Raymond, C., Horton, R.M., van den Hurk, B., AghaKouchak, A., Jézéquel, A., Mahecha, M.D., Maraun, D., Ramos, A. M., Ridder, N.N., Thiery, W., Vignotto, E., 2020. A typology of compound weather and climate events. *Nat. Rev. Earth Environ.* 1 (7), 333–347. <https://doi.org/10.1038/S43017-020-0060-Z>.