The Annals of Applied Statistics
2017, Vol. 11, No. 3, 1403–1428
DOI: 10.1214/17-AOAS1031

A SEMIPARAMETRIC METHOD TO SIMULATE BIVARIATE SPACE–TIME EXTREMES[1]

BY ROMAIN CHAILAN[*,†], GWLADYS TOULEMONDE[*] AND JEAN-NOEL BACRO[*]

University of Montpellier[*] and IBM France[†]

Coastal hazards raise many concerns, as their assessment involves extremely high economic and ecological stakes. In particular, studies on rarely observed but damaging events are quite numerous. In order to anticipate upcoming events of this kind, specialists need to extrapolate the results of their studies to events that have not yet occurred. Such events might be more extreme than those already observed and could therefore severely impact the coast. It is therefore paramount to propose methodologies to simulate such extreme conditions. Parametric and nonparametric statistical methods have already been used to assess environmental extreme quantities, from univariate framework to spatial context; however, they do not generally focus on the simulation of extreme environmental scenarios. This study introduces a semiparametric approach based on the Extreme Value Theory (EVT), dedicated to the simulation of extreme space–time processes. In the proposed application context, these processes describe near-shore hydraulic conditions. They nourish coastal impact models to assess hazards along the coast. The benefit of this approach is to be able to characterise coastal hazards on an event scale, meaning we can characterise the impact both in space and through time for a given extreme event. The usefulness of this space–time extreme modelling is illustrated by a risk analysis related to the long-shore impact of extreme wave events in the Gulf of Lions.

1. Introduction. Coastal hazards raise many concerns, as highly valuable economic and ecological assets are exposed along the world's coasts. Several studies demonstrate the significant benefits of understanding both littoral hydrodynamic and morphodynamic patterns in order to preserve them [e.g., Brunel et al. (2014), Gutierrez et al. (2015), Michaud et al. (2013)]. Some experts focus on extreme and devastating conditions, such as Campmas et al. (2014), who observes sediment transport patterns during the season of typhoons in Taiwan. Such a study helps preserve the littoral by enabling efficient beach nourishment.

An alternative to direct observations is the chaining of numerical models, which represent the physics from offshore to coastal areas. Typically, output data from

Received January 2016; revised February 2017.

[1]This work was supported by the french national program LEFE/INSU and by the labEx NUMEV.

Key words and phrases. Space-time extreme processes simulation, extreme value modelling, extreme waves, coastal hazards.

atmospheric and ocean circulation models force a wave model, which in turn feeds a littoral model [Bouchette et al. (2012), Michaud (2011)]. Refined output data of the latter are used to assess the hazard question.

In the case of observable extreme events, the reliability of physical models still holds. As soon as we consider very extreme events, their numerical simulation from physical models is generally unachievable. This is due to a lack of knowledge of boundary conditions and also of their physical reliability for such extreme quantities. As an alternative we propose to use statistical approaches, the main challenge being to extrapolate information from observations to simulate (very) extreme quantities.

From univariate to spatial approaches, analyses dealing with the understanding of extremes generally rely on the widely accepted Extreme Value Theory (EVT) [Beirlant et al. (2004), Coles (2001), Davison, Padoan and Ribatet (2012), Davison and Huser (2015)].

Various approaches have been presented to construct extreme scenarios of nearshore conditions like in Gouldby et al. (2014), but are generally not spatial. In the spatial context, Chailan et al. (2014) present an application of max-stable processes to analyse the spatial behaviour of extreme waves. The outputs of this study would be typical requirements to force physical hazard models in a coastal area. Indeed, max-stable processes are appealing in a spatial extreme context because they are the only possible nondegenerate limits for rescaled pointwise maxima of random processes [de Haan (1984)]. Inference of such max-stable processes is widely based on likelihood techniques, either in a frequentist approach [Engelke et al. (2015), Huser and Davison (2013), Padoan, Ribatet and Sisson (2010), Wadsworth and Tawn (2014)] or in a Bayesian one [Ribatet, Cooley and Davison (2012), Shaby (2014)]. Shaby and Reich (2012) present a Bayesian spatial extreme value analysis but interpreting the parameters in their hierarchical modelling is unfortunately not easy [for a possible interpretation as well as recent investigations on inference for spatial extremes, see Castruccio, Huser and Genton (2016)]. From a practical point of view, simulations of max-stable processes are of primary interest. They can be divided in two categories: unconditional and conditional simulations. For instance, Dieker and Mikosch (2015) propose exact simulations of the Brown–Resnick max-stable process at a finite number of locations. Their approach has been generalised by Dombry, Engelke and Oesting (2016) who also propose a more efficient algorithm. Wang and Stoev (2011) introduce a solution to construct a conditional process for max-linear processes. This work was extended by Bechler, Bel and Vrac (2015), Dombry and Eyi-Minko (2013), Dombry, Éyi-Minko and Ribatet (2013) in a less restrictive case. Nevertheless, the number of conditioning points remains limited and Lantuéjoul and Bel (2014) have recently remedied this weakness by introducing a new algorithm. However, since max-stable processes appear as natural for modelling block maxima (e.g., annual maxima), using simulations of such a process is more relevant in long-term questioning than in event-scale questioning due to the limited physical interpretation

of the simulated processes. This is clearly a limiting factor when questioning is more event-scale related (e.g., a submersion phenomenon along a coastline is an event-scale phenomenon and must be distinguished from a long-term problematic like the study of the decennial coastline dynamic). In the former case, not only the spatial information of an extreme process is needed, but also information characterising the time evolution of the analysed extreme event itself. For instance, this is essential in coastal engineering applications to compute dimensioning characteristics, such as the fatigue of seawalls through time when they are impacted by storm-waves.

In the following, we focus on space–time processes. Max-stable processes have also been developed and exemplified in a space–time context [Davis, Klüppelberg and Steinkohl (2013a, 2013b), Huser and Davison (2014), Embrechts, Koch and Robert (2016)] but are rarely alluded to in the literature. Their capacity to model complex dependence structures can still be questioned and the physical interpretation in any event-scale context of the simulated space–time processes issued by these models can be questioned as well.

Since these fully parametric methods do not directly answer the presented event-scale problematic and since it is unfeasible to model statistically the physical characteristics of storm events, we propose a methodology based on an empirical uplifting of real storms. This has the benefit of preserving the underlying physics of the considered processes. The idea is to exploit a peaks-over-threshold based approach and to propose a simulation scheme for extreme realisations. This does not assume any parametric model for the dependence structure. In the proposed methodology, we are focused on a semiparametric approach stemming from parts of the original work of Caires, de Haan and Smith (2011), de Haan and de Ronde (1998), Ferreira and de Haan (2014), Groeneweg, Caires and Roscoe (2012), summed up as follows.

Let $\{Z(s,t), s \in S, t \in \mathcal{T}_0\}$ be a space–time process, with $S \subset \mathbb{R}^2$ the area of interest and $\mathcal{T}_0 \subset \mathbb{R}^+$ the time dimension. In the sequel, such a process will represent an extreme event and will be named 'storm' for the sake of simplicity. The first step consists in selecting such a storm. To do so, the complete process is standardised in a preprocessing step. A combination between a preprocessing step and an extreme modelling has been proposed by Eastoe and Tawn (2009) but in a context of nonstationarity due to the presence of covariates. Here, a more extreme process is obtained by uplifting with a coefficient denoted $\zeta > 1$ the space–time process, which is initially transformed on a standard scale as $T(Z)$ where T is a marginal transformation detailed in Section 3.1. The process $T^{\leftarrow}(\zeta T(Z))$ becomes more extreme at the original scale.

Assuming that Z belongs to a max-stable domain of attraction, this approach is mathematically justified (see the Appendix). In practice, the space–time dependence structure of Z will be taken as constant in the extreme, leading to an asymptotic dependence context. Caires, de Haan and Smith (2011), Groeneweg, Caires and Roscoe (2012) use this methodology to simulate space–time extreme

processes. We leverage this approach to perform a bivariate simulation of such processes, with a view to better represent sea-states conditions at extreme levels. This leads us to develop a distinct strategy for the selection of storms and to use marginal distributions for the standardisation of the data as those used in Thibaud and Opitz (2015).

The behaviour of the produced storms is discussed around a case-study: the quantification of the long-shore mass flux of energy in a coastal area during extreme storms.

Since the presented methodology is applied to a large multidimensional volume of data, specific distributed algorithms are developed to process the data, which raises an additional technical dimension.

Section 2 introduces both the dataset used for this application and a preliminary study about the storms contained in it. Section 3 then presents in detail the statistical methodology and its justifications. The results are presented in Section 4.1 and then used for a risk analysis in Section 4.2. The final section provides a discussion about the introduced notions and their applications.

2. Data. Our region of interest is a semi-closed French coast area located in the northwestern Mediterranean sea, namely the Gulf of Lions (GOL) as presented in Figure 1. This study aims to simulate extreme space–time wave processes in order to use them as inputs for a littoral hazard model. For instance, a model of coastal submersion due to storm-waves, which is a physical process depending on near-shore hydrodynamic conditions. Such a model is forced by inputs describing

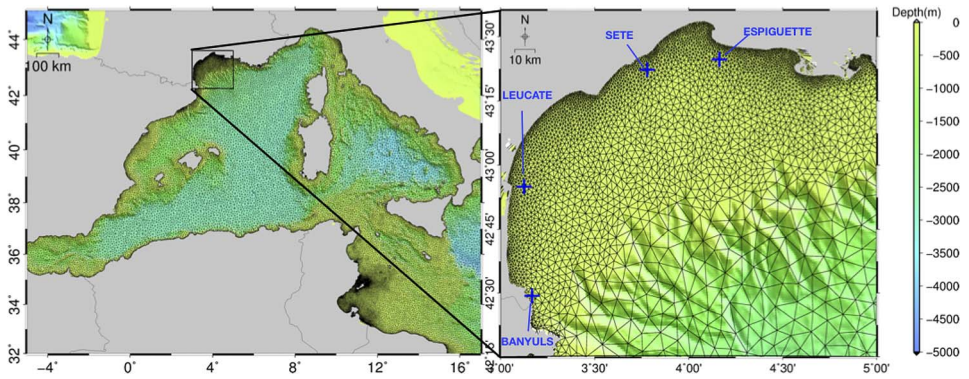


FIG. 1. *The left panel is the full extension of the domain considered for the hindcast. The right panel is the studied area*: *the Gulf of Lions* (*GOL*). *The crosses indicate the locations of surface buoys measuring waves features. The colour scale indicates the bathymetry, that is, underwater topography, of the northwestern Mediterranean sea. The computational mesh used for the hindcast is also overlaid. It is composed of* 47,086 *nodes with a spatial resolution ranging from* 1 *km to* 12 *km. The right panel is a zoom of the grid on the GOL. Computational nodes situated in the GOL form the set* $\mathcal{M}$.

the sea-states conditions at an instant t. Generally, these inputs are the mean wave direction $\psi(t)$, the significant wave height $\mathrm{Hs}(t)$ and the peak wave period $\mathrm{Tp}(t)$.

Three sources are principally considered in obtaining such data. The first is surface-buoys that monitor these three variables. In the GOL, there are four surface-buoys as illustrated in Figure 1. These observations are accurate but sparsely provided in our region of interest, in both space and time dimensions. Spatial scarcity would degrade the spatial modelling of the process whereas short time series would degrade the quality of the extrapolation to more extreme values.

An alternative is to use satellite-altimeter datasets. The major issue is that only $\mathrm{Hs}(t)$ can be observed from an altimeter. Satellites embedding Synthetic Aperture Radar (SAR) must be considered if wave direction and wave period are required, but the time series are shorter (first launch in the 1990s). Moreover, since satellites tracks are nonregular through time and space around the globe, any extreme statistical analysis considering such datasets [see, e.g., Raillard, Ailliot and Yao (2014)] becomes hard to handle, especially when the modelling concerns events in a fixed and relatively confined area.

The final way to observe wave data variables is the use of the numerical simulation. Chailan et al. (2014) proposed a 52-year hindcast of wave features over the north-western Mediterranean sea, extending from the Strait of Gibraltar to the south of Italy. This hindcast is obtained by the use of a widely recognised wave numerical model in ocean community. In the sequel, this hindcast—validated against in situ observations—is used since it provides the longest and refined wave time series for this area to the best of our knowledge [Chailan (2015), Chapter 3]. Details are given in the next subsection.

2.1. *A 52-year wave hindcast.* The 52-year hindcast is produced with the WAVEWATCHIII® v4.18 (WW3) wave model [Tolman (2014)]. Two regional reanalyses have been used as forcing conditions—meaning used as inputs of the numerical model: Herrmann and Somot (2008) for atmospheric conditions and Herrmann et al. (2010) for ocean conditions. The bathymetry used has a spatial resolution of 0.0083 degree. The physical time range of the simulations ranges from January 1961 to December 2012 at an hourly scale. Finally, the unstructured computational grid illustrated in Figure 1 is composed of 47,086 nodes—3,944 for the GOL only—with a spatial resolution ranging from 1 km to 12 km.

The former quantities of interests [i.e., $\psi(t)$, $\mathrm{Hs}(t)$ and $\mathrm{Tp}(t)$] derive from the computed wave spectral density at each node of the mesh. For the GOL only, these three variables are stored in a binary file of 19 GB.

The dataset produced is validated against the records of the four surface buoys, at a yearly scale in terms of the time series available. As it is often the case, the wave model shows a good performance but tends to slightly underestimate the extreme occurrences. One way to understand the performance of the hindcast is to look at both marginal and joint measures of validation.

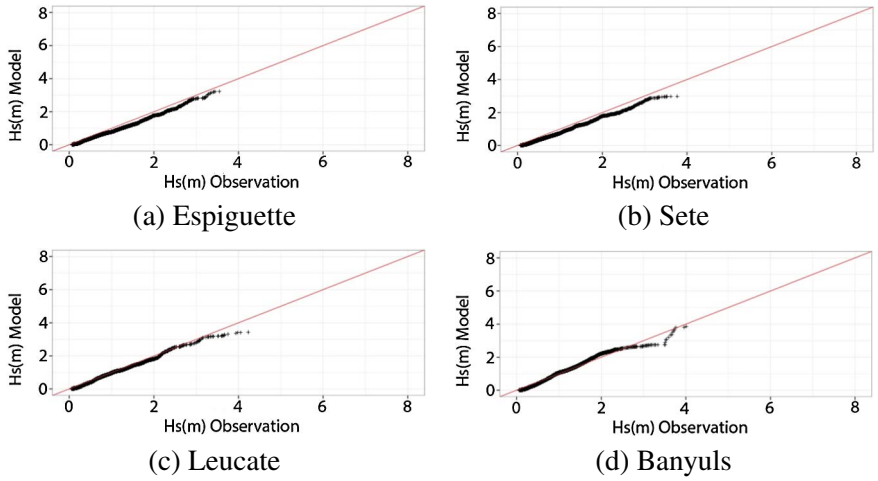


FIG. 2. *Quantile–quantile plots of the observed significant wave heights* (Hs) *against the modelled ones for* 2012. *Locations are the four littoral surface buoys of the GOL.*

For instance, the median over all buoys of the yearly correlation factors reaches 0.903 while the median of the root mean square errors is 0.272. Figure 2 illustrates quantiles of the observed significant wave heights (Hs) against the modelled ones for the year 2012. For this year, the former measures approximate their medians for each location, respectively. It makes 2012 a representative candidate to diagnose the overall hindcast quality [see Chailan (2015), Chapter 3, for additional measures].

The observed bias might not come from the wave model only [Rascle and Ardhuin (2013)]. Indeed the forcing re-analyses, especially the wind fields, are sometimes underestimated for instantaneous and abrupt wind gusts. Consequently, the generated wind-waves are underestimated as well. Despite these slight underestimations, the produced data are relatively satisfactory.

Insofar as it is a key feature of our study, the performance of the numerical model in regards to the spatial dependence structure must be presented as well. An analysis of joint survival probability is performed to this end. The purpose is to compare from each source—either buoys or numerical models—the joint probability of exceedance from various sets of sites corresponding to the locations of the buoys. In the empirical computations and for each sub-set of locations, the records taken into account are those that are simultaneously available at each site of the set. The thresholds quantiles are calculated, respectively, for each source of data. This limits misinterpretation due to bias from marginal intensities.

For the sake of clarity, only three out of the ten combinations available (four buoys) are presented in Figure 3, but similar results are observed regarding the other sets.

Those plots reveal a good match between survival joint probabilities from buoys compared to the ones from the numerical model, whatever the distance between the

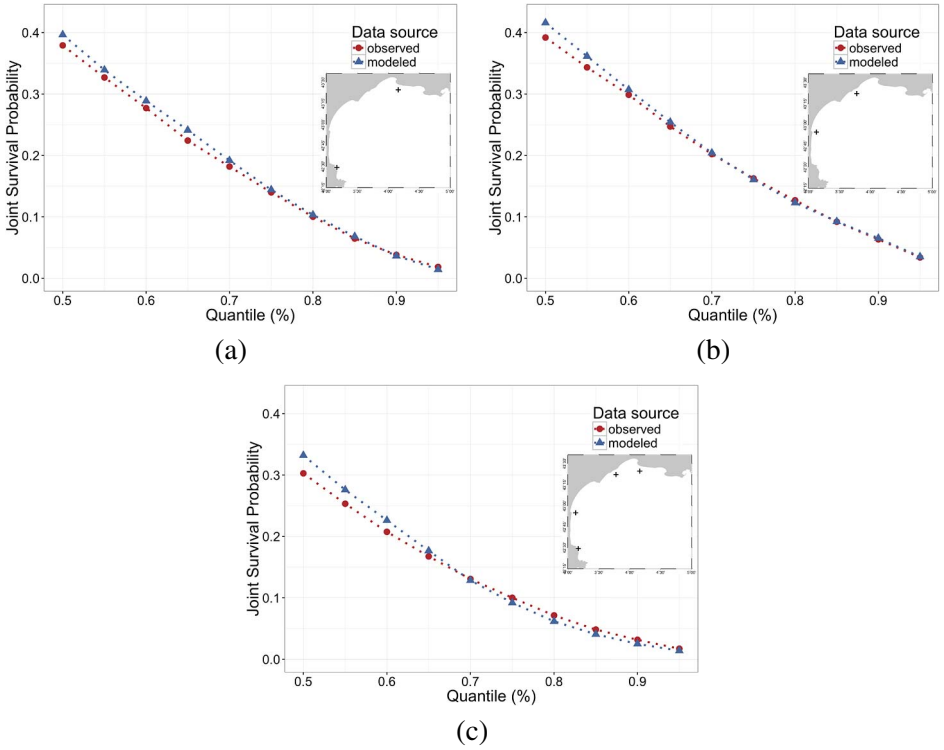


FIG. 3. *Joint survival probabilities of exceedance of significant wave heights* (Hs). *The empirical probabilities are computed from each data source* (*buoys or numerical model*). *Each sub-panel represents joint probabilities over various sets of sites corresponding to the buoys' locations. Selected sites are localised by the crosses on the map for each sub-panel.*

sites or their numbers. The adequacy is especially valid for joint probabilities of exceedance over high quantiles but with higher bias on smaller quantiles. It means that small waves are more spatially structured when observed from the numerical wave model but the spatial dependence structure is properly modelled for high waves. This remark reinforces the relevance of considering those produced data as observations in the sequel.

2.2. *Preliminary analysis*. A preliminary analysis is realised to develop our expertise on the wave data previously presented. As the reader may know, wind is the major factor of wave construction. The GOL is exposed to three dominant wind regimes. The first two are called *Tramontane* and *Mistral*. They come from the northwest and north, respectively. The last is called *Marin* and comes from the southeast. When the region is exposed to a Tramontane or Mistral episode or both, waves tend to propagate towards the southeast but are formed far from the coastline. This is due to a too short *fetch* zone—the zone where the wind stresses the sea-surface causing the growth of the waves. On the contrary, as soon as the

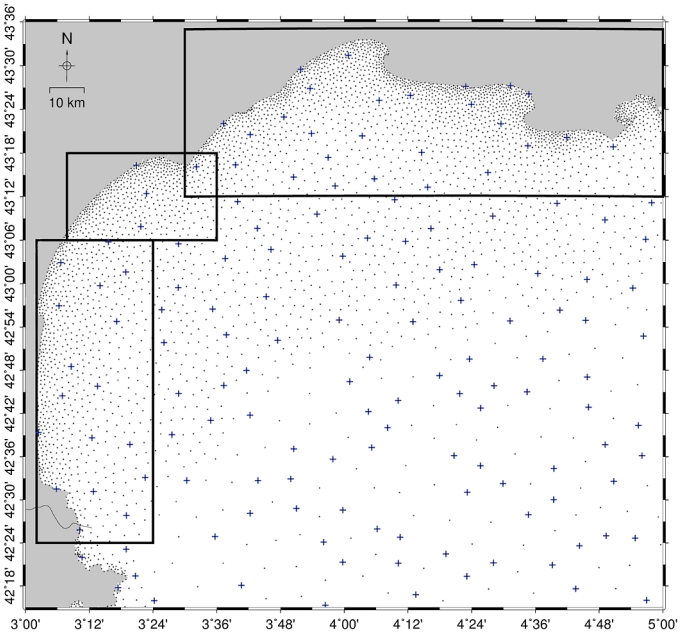


FIG. 4. *Spatial specification. Littoral area $S^* \subset S$ is the union of squared areas. From expert advice, if* Hs *is high in S^* the coastline is likely to be impacted. Wave data are available at the set of locations of the mesh nodes in this area, which is denoted $\mathcal{M}^* \subseteq S^*$. Cross points form a subset χ of* 140 *sites from the locations of the computational mesh nodes. χ is constructed in manner of spatially representing all observation locations.*

area is exposed to a Marin episode, waves are formed offshore and are propagated to the coasts. In such cases, the waves impact the coastline. Winds hitting the GOL are sometimes more complex and the resulting hydrodynamic is fairly modified: occasionally a southwest wave-flux is dominant in the GOL. Experts advise that the relevant storms to study the impact on the coastline are those in which the Hs variable reaches high values inside a very littoral area denoted S^*. For the GOL, we decided to choose the union of the determined areas (Figure 4).

Beside these physical characteristics, some statistical information can provide valuable information about the general behaviour of a wave-storm in the GOL. In particular, the extremal coefficient θ [Smith (1990), Schlather and Tawn (2003)] is a quantity that enables us to quantify the dependence in the context of extreme values.

This measure stems from the following reasoning. Without loss of generality, let us consider identically distributed random variables $Y^{(1)}, \ldots, Y^{(M)}$ with unit Fréchet distribution, that is, $P(Y^{(i)} \leq y) = e^{-1/y}, i = 1, \ldots, M, 0 < y < \infty$. If the joint distribution of $(Y^{(1)}, \ldots, Y^{(M)})$ is a multivariate extreme value distribution, it is well known that the joint probability $P(Y^{(1)} \leq y, \ldots, Y^{(M)} \leq y)$ can be expressed as $e^{-\theta/y}$. The so-called extremal coefficient $\theta = \theta(Y^{(1)}, \ldots, Y^{(M)})$,

$1 \leq \theta \leq M$ summarises the extremal dependence. The limiting case $\theta = 1$ represents the full dependence whereas $\theta = M$ represents the total independence.

In the context of threshold-based extreme value methods, realisations above a high threshold are considered as extreme. Assuming predetermined thresholds vectors $(u_j^{(1)}, \ldots, u_j^{(M)})$ and random vectors $(Y_j^{(1)}, \ldots, Y_j^{(M)})$, $1 \leq j \leq N$, the $Y_j^{(k)}$ are observed only if $Y_j^{(k)} > u_j^{(k)}$; otherwise, $Y_j^{(k)}$ is censored at $u_j^{(k)}$.

In this context, Smith in Caires, de Haan and Smith (2011) defines a natural estimator of the extremal coefficient function θ as

$$\hat{\theta} = m \Big/ \sum_{j=1}^{N} \frac{1}{\max(Y_j, u_j)}, \tag{2.1}$$

where Y_j and u_j are defined as $\max(Y_j^{(1)}, \ldots, Y_j^{(M)})$ and $\max(u_j^{(1)}, \ldots, u_j^{(M)})$, respectively; m is the number of excesses $Y_j > u_j$.

The pairwise extremal coefficient is commonly considered in statistical applications, meaning $Y_j = \max(Y_j^{(1)}, Y_j^{(2)})$ with $M = 2$ in (2.1). In the sequel, three extremal coefficients are introduced and estimated for the sea-states hindcast dataset. The first two are related to the dependence of the variable Hs through time and spatial distance, respectively. The time extremal coefficient $\theta^{\mathrm{tim}}(k)$ measures the dependence between pairs of observations of Hs separated by a time lag k, at a given location. The spatial extremal coefficient $\theta^{\mathrm{spa}}(h)$ measures the dependence between pairs of Hs observations separated by a spatial distance h, at a given time.

Figure 5 presents the extremal coefficients estimated for the full period (1961–2012) of the hindcast on a yearly block of data in order to monitor their fluctuations. Here, u_j in (2.1) is set as a 0.95-quantile to avoid issues stemming from a lack of data. Figure 5(a) presents the estimations $\widehat{\theta^{\mathrm{spa}}}(h)$ for two locations separated by a distance h. In this case, $(Y_j^{(1)}, Y_j^{(2)})$ in (2.1) corresponds to $(Y(t_j, s), Y(t_j, s + h))$. To compute $\widehat{\theta^{\mathrm{spa}}}(h)$, only a subset χ of 140 sites (Figure 4) from the computational mesh is considered. It limits the combinations of pairs available in the dataset. The selection of sites is optimised to fairly cover the entire area as described in Chailan et al. (2014). Estimations $\widehat{\theta^{\mathrm{spa}}}(h)$ are binned to 1,500 distinct distances h.

We observe that $\widehat{\theta^{\mathrm{spa}}}(h)$ is always strictly inferior to 2. More precisely, it is approximately 1.75 at the longest distance, exemplifying that dependence within a storm on the GOL, which is a relatively confined area, seems to be relatively persistent even at the longest distances. Beside the presented omnidirectional graphic, directions of pairs were considered and regrouped to compute the directional estimation of the dependence structure. This did not demonstrate a clear anisotropic pattern and, therefore, graphics are not presented here.

Figure 5(b) presents the estimations of $\theta^{\mathrm{tim}}(k)$ for pairs separated by a time lag k. In this case $(Y_j^{(1)}, Y_j^{(2)})$ in (2.1) represents $(\max_{s \in \mathcal{M}^*}(Y(t_j, s)), \max_{s \in \mathcal{M}^*}(Y(t_j + k, s)))$, with $\mathcal{M}^*$ the observation locations situated in S^* the

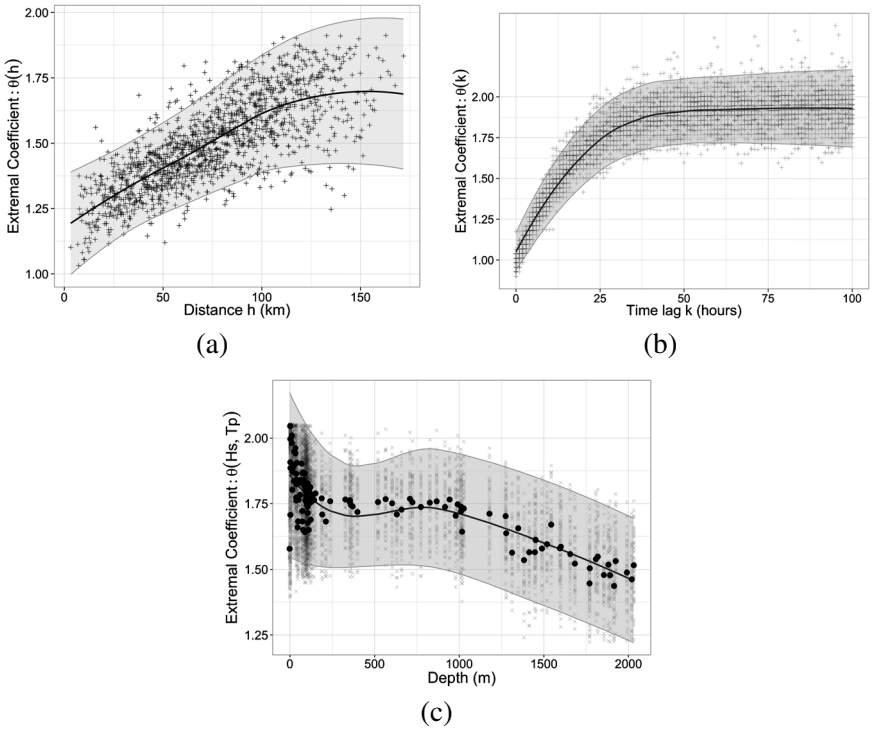


FIG. 5. *Estimations of the three extremal coefficients* (*see text for details*). *For each pair, the coefficients are estimated for the full period* (1961–2012) *of the hindcast on yearly block of data*. (a) *The extremal coefficients* $\theta^{\mathrm{spa}}(h)$ *estimated on* χ *from pairs of* Hs *values separated by a distance* h *given in kilometers*. *Estimations are binned to* 1,500 *distinct distances* h. (b) *Extremal coefficients* $\theta^{\mathrm{tim}}(k)$ *are estimated from pairs of* Hs *values separated by a lag* k *in hour.* (c) *Extremal coefficients* $\theta(\mathrm{Hs}, \mathrm{Tp})$ *estimated from the significant wave height* Hs *and the peak wave period* Tp *at locations* $s \in \chi$ *are ordered by their corresponding bathymetry. The dots are the median values from estimated pairwise coefficients*. *In each sub-panel*, *the straight line and its shadow envelope are respectively a fitted polynomial regression model and its* 95% *prediction interval*.

very littoral zone presented above. The arbitrary choice of S^* is still related to the final goal of the document: quantifying coastal hazards. With such littoral areas, only storms impacting the shoreline area are considered in the measure. We can observe from Figure 5(b) that $\widehat{\theta^{\mathrm{tim}}}(k)$ narrows 1.9 and becomes almost steady at $k = 50$. Hence, we can state that the dependence within a storm impacting the littoral will be considered as persistent only up to 50 hours.

The proposed uplifting procedure relies on a crucial hypothesis which is max-stable context. Indeed, we assume that the space–time dependence structures are constant in the extreme. Figures 5(a) and 5(b) show that this hypothesis is reasonable with our data, when considering a time lag smaller than 50 hours, corresponding to an extremal coefficient strictly inferior to 2.

Finally, to assess the dependencies between the two wave variables Hs and Tp observed at the same time and at the same location, we consider a third extremal coefficient $\theta(\mathrm{Hs}, \mathrm{Tp})$. Let $\mathrm{Hs}(t_j, s)$ and $\mathrm{Tp}(t_j, s)$ denote the significant wave height and the peak wave period at time t_j and location s, respectively. In this case, $(Y_j^{(1)}, Y_j^{(2)})$ in (2.1) represents $(\mathrm{Hs}(t_j, s), \mathrm{Tp}(t_j, s))$. Estimation $\widehat{\theta}(\mathrm{Hs}, \mathrm{Tp})$ is computed using the data from the subset χ. Figure 5(c) illustrates such estimation. By ordering the estimated bivariate extremal coefficients by the depth of the observation sites, we show that the deeper the sites, the more Hs and Tp remain dependent within their extreme realisations. In general, we can deduce that those two variables are fairly dependent, with an extremal coefficient inferior to 2, even if the waves mechanic may behave differently in very shallow waters.

3. Semiparametric storm uplifter.

3.1. *Extreme space–time processes.* In the sequel, $\{X(s,t), s \in S, t \in \mathcal{T}\}$ denotes a random space–time process with S a compact subset of $\mathbb{R}^d$ and $\mathcal{T}$ a compact subset of $\mathbb{R}^+$. Such a random process represents a random variables collection indexed by both space and time which is in the space of continuous real functions on $S \times \mathcal{T}$ denoted $C(S \times \mathcal{T})$. We suppose that $\{X(s,t), s \in S, t \in \mathcal{T}\}$ is in the domain of attraction of a max-stable process [de Haan and Lin (2001), de Haan and Ferreira (2006)]. In other words, we suppose that there exist continuous functions $a_n(s,t)$ positive and $b_n(s,t)$ such that the process

$$\left\{ \max_{1 \le i \le n} \frac{X_i(s,t) - b_n(s,t)}{a_n(s,t)} \right\}_{(s,t) \in S \times \mathcal{T}}$$

with $X_1, \ldots, X_n$ independent copies of X, converges in distribution to a max-stable process η in $C(S \times \mathcal{T})$. Since convergence of marginals and convergence of dependence structure can be split up, we consider, in the sequel, the standardised process $1/(1 - G_{X(s,t)}(X(s,t)))$ where $G_{X(s,t)}$ corresponds to the distribution of $X(s,t)$. Such a process has marginal standard Pareto distributions and belongs to the domain of attraction of the unit Fréchet distribution. Following Thibaud and Opitz (2015), it is convenient to fix a high threshold function $u(s,t)$ and to assume that the marginal distributions of this process satisfy

$$P\big(X(s,t) > x\big) = \big[1 + \xi(s,t)\big(x - \mu(s,t)\big)/\sigma(s,t)\big]_+^{-1/\xi(s,t)}, \tag{3.1}$$

for $x > u(s,t)$, with real parameters $\mu(s,t) < u(s,t)$, $\sigma(s,t) > 0$ and $\xi(s,t)$, such that the right-hand side of (3.1) is less than unity.

As a consequence, to result in a process with standard Pareto margins, we can define the standardised process X^* as follows:

$$X^*(s,t) = T\big(X(s,t)\big) = \big[1 + \xi(s,t)\big(X(s,t) - \mu(s,t)\big)/\sigma(s,t)\big]^{1/\xi(s,t)}. \tag{3.2}$$

3.2. *Method*. As presented in the Introduction, the outline of the methodology consists of four steps. First, data are marginally transformed. This enables us to manipulate the data on a standard scale. Here, we use a transformation to reach the standard Pareto scale. Then we need to extract storms from the dataset. Once storms are extracted, the data are uplifted to higher values, with a control on the marginal amplification coefficient. Finally, the data are transformed back to their original scale by inverting the transformation. Details of these four steps of the presented methodology are given in this subsection.

The first step consists in standardising $X(s,t)$ to a standard Pareto scale according to (3.2). In practice, the parameters are unknown and need to be estimated. In this first approach, we suppose the threshold and the parameters to be constant over time, depending only on space. One can alternatively use more sophisticated expressions of those quantities to deal with a potential nonstationarity of the process, for example, seasonality and directional effects might be better explained doing so [e.g., Jonathan, Ewans and Randell (2013)].

In each site, parameter estimations $\hat{\mu}(s)$, $\hat{\sigma}(s)$, $\hat{\xi}(s)$ are obtained by the maximum likelihood method using data above a high threshold $u(s)$ which can be chosen as a high quantile for a fixed s (here the 0.99-quantile). Since marginal data may have some short-term dependences, they are de-clustered before being used to estimate the parameters [Coles (2001), Section 5.3.2]. In this paper, the de-cluster procedure has been configured with an interval of 5 consecutive values below u_s to consider an exceedance as a new cluster, that is, 6 hours after the last exceedance. This step allows us to reach the independence condition assumed in the estimation procedure. Using such estimators in (3.2), let $\{\tilde{X}^*(s,t), s \in S, t \in \mathcal{T}\}$ denote the obtained standardised process. Note that this preprocessing step relies on different techniques from those used in Caires, de Haan and Smith (2011), Groeneweg, Caires and Roscoe (2012).

The second step consists in extracting storms on a standardised scale from the data. To extract the biggest storm, the maximum value of $\widetilde{X}^*(s,t)$ is searched over the subset of sites $\mathcal{M}^*$, which might be a single reference location, locations of the entire space S or locations of some area in between. This point leads to a distinct strategy of selection of storms from Caires, de Haan and Smith (2011), Groeneweg, Caires and Roscoe (2012). Let us assume this maximum occurs at time t_1. We fix the total storm duration as 2δ. Consequently, such a storm is a subset in the time dimension of $\{\tilde{X}^*(s,t), s \in S, t \in \mathcal{T}\}$, therefore, defined as $\tilde{Z}^* = \{\tilde{X}^*(s,t), s \in S, t \in \mathcal{T}_0 \subset \mathcal{T}\}$. For this first storm, $\mathcal{T}_0 = [t_1 - \delta, t_1 + \delta]$.

The period $\mathcal{T}_0$ is hidden from the selection of the second biggest storm. Furthermore, we introduce a time value which is a "precaution time-lag" ε to insure the independence of the storms. The selection of the second biggest storm will consist in identifying the maximum value of $\widetilde{X}^*(s,t)$ over the subset of sites $\mathcal{M}^*$ with $t \in \mathcal{T} \backslash [t_1 - \delta - \varepsilon, t_1 + \delta + \varepsilon]$. The two values δ and ε are generally defined according to expert advice or from preliminary analyses or both. In this study, the specific values of these parameters are given and explained in Section 4.1.

Algorithm 1: Storm selection

Input : $\{\tilde{X}^*(s,t),\ s \in S,\ t \in \mathcal{T}\}$, space–time observations on a standard scale.
p' the maximum number of storms to select.

Output: $\{\tilde{Z}_i^*, i \in \{1,\dots,p\}\}$ with $p \le p'$, a sorted collection of i.i.d. storms

begin
  $i = 1$, $\delta \leftarrow$ Cst, $\varepsilon \leftarrow$ Cst, $T \leftarrow \mathcal{T}$, $T' \leftarrow T$;
  while $(i \le p')$ and $(\max_{s\in\mathcal{M}^*, t\in T'} \tilde{X}^*(s,t) > 1)$ **do**
    $t_i \leftarrow \arg\max_t\{\tilde{X}^*(s,t)\}$; // $s \in \mathcal{M}^* \subseteq S$ and $t \in T'$.
    $\tilde{Z}_i^* \leftarrow \tilde{X}^*(\cdot,t)$ with $t \in T \cap [t_i - \delta,\ t_i + \delta]$;
    $T' \leftarrow T' \setminus [t_i - \delta - \varepsilon,\ t_i + \delta + \varepsilon]$;
    $i = i + 1$;
  return $\{\tilde{Z}_1^*, \tilde{Z}_2^*, \dots, \tilde{Z}_p^*\}$;

The general iterative scheme to select storms is presented in Algorithm 1. It is noticeable that the stop condition of the algorithm implies that there is at least one exceedance of the site marginal threshold in each selected storm. The algorithm would select storms until the required and arbitrary number of storms p' is reached or until the exceedance condition is no longer satisfied.

Finally, let $\{\tilde{Z}_i^*(s,t), i \in \{1,\dots,p\}\}$ denote a collection of such space–time processes and represent the p highest storms available in the transformed dataset.

It is relevant to compare them with each other in term of their extremeness. In the sequel, the definition of extremeness of a so-called storm $\{Z^*(s,t), s \in S, t \in \mathcal{T}_0 \subset \mathcal{T}\}$ relies on the level corresponding to the within-storm maxima $z_{\max} = \max_{s,t}\{Z^*(s,t), s \in \mathcal{M}^* \subset S, t \in \mathcal{T}_0 \subset \mathcal{T}\}$. Consequently, a storm $\{Z_1^*\}$ is considered more extreme than $\{Z_2^*\}$ if $z_{1,\max} > z_{2,\max}$.

In extreme value theory, a return period m is associated with a return level r_m. The return level r_m is reached once over the return period m in mean. By definition, this is no more than the $(1 - \frac{1}{m})$-quantile of the block maximum distribution. We define the return period of a storm $\{Z^*(s,t)\}$ as equal to the marginal return period associated with the within-storm maxima $z_{\max}$ observed at location $s_{\max}$. The location $s_{\max}$ is either fixed as a reference site or defined as equal to $\operatorname{argmax}_{s\in\mathcal{M}^*}\{\tilde{Z}^*(s,t)\}$.

The third step consists of an uplifting technique. To obtain more severe storms (with a longer return period), processes $\tilde{Z}_i^*, i \in \{1,\dots,p\}$ are multiplied by a coefficient factor superior to unity and denoted ζ_i. The coefficient ζ_i is applied to the entire duration of the storm i. Hence, $\zeta_i \tilde{Z}_i^*(s,t), \zeta_i > 1, i \in \{1,\dots,p\}$, is the collection of the uplifted storms at the standardised scale.

For the final step, each uplifted storm is transformed back to its original scale by

$$\tilde{Z}_i(s,t) = T^{\leftarrow}\big(\zeta_i \tilde{Z}_i^*(s,t)\big), \qquad i \in \{1,\dots,p\}, \tag{3.3}$$

where $T^{\leftarrow}(Y(s,t)) = \hat{\mu}(s) + \hat{\sigma}(s)\frac{[Y(s,t)]^{\hat{\xi}(s)}-1}{\hat{\xi}(s)}$.

We obtain here a collection of heavier extreme storms from a set of observed extreme storms.

It is important to highlight that an observed extreme storm $Z_i^*(s,t)$ is defined if and only if

$$\max_{s\in\mathcal{M}^*} Z_i^*(s,t) > 1, \tag{3.4}$$

meaning that there is at least one exceedance of the site marginal threshold. This uplifting proposition relies on a mathematical justification given in the Appendix. In this detailed proof, it has been shown that there is actually no limitation in uplifting bivariate processes $\{Z_{1,i}^*, Z_{2,i}^*\}$ conditioned to (3.4) is satisfied for one of the margin.

What is further remarkable is that such a uplift method of a space–time process appears as naturally linked to the GPD process framework. This framework was initially introduced by Ferreira and de Haan (2014). Dombry and Ribatet (2015) generalise this result by considering conditional events characterised through a continuous and homogeneous risk function $\ell(\cdot)$. The case from Ferreira and de Haan (2014) corresponds to $\ell(f) = \sup_{s\in S} f(s)$ and the ℓ function we are considering here corresponds to $\ell(f) = \max_j f(s_j,t)$. As a consequence, the limit of the conditional distribution we consider corresponds to the distribution of a GPD process.

Other remarks can be made with regard to the construction of the processes. First, note that in (3.3), the coefficient ζ_i relative to the uplifted storm i can be chosen in several ways as long as it is superior to 1.

We can consider, and this is in fact the choice we made, the special case $\zeta_i = \frac{T(z_m)}{T(z_{\max})}$, where $z_{\max}$ is still the within-storm maxima and z_m is the return level corresponding to the m-year return period at location $s_{\max}$. Implemented in Groeneweg, Caires and Roscoe (2012), Smith in Caires, de Haan and Smith (2011) interprets such a transformation as an uplift from a storm with a given return period to a storm with a return period equal to m. In that case, ζ_i is obviously storm-dependent and this choice enables us to uplift different storms to a comparable level. However, other choices for ζ_i could be proposed, for example, in Caires, de Haan and Smith (2011), de Haan proposes another approach which can be interpreted as an uplifting of the threshold of the peaks-over-threshold process Z_i. As another example, $\zeta_i, i = 1,\dots,p$ could be obtained as independent realisations of a standard Pareto distribution. In that case, our approach should be very similar to the constructive representation of the Pareto process proposed by Dombry and Ribatet (2015). To the best of our knowledge, there are few results about simulations of GPD processes and consequently our results may also be of interest.

4. Results.

4.1. *Uplifted storms*. The presented method is applied to the 52-year sea-states condition dataset. To cope with the computational demand of dealing with nearly 4000 locations, algorithms are implemented in a dedicated R code and parallelised via the Message Passing Interface (MPI) protocol. All computations are performed on a cluster composed of 96 cores, which reduces the overall computation duration to nearly 5 hours.

From this point on and for the sake of simplicity, the definition of storm embraces the multivariate space–time processes composed of Hs, Tp and directions ψ.

We worked on the 10 highest storms observed to uplift both Hs and Tp variables, resting on the proposed bivariate approach. In our case study, Hs is the variable that conditions the bivariate space–time processes selection. It avoids selecting events with high Tp but low Hs, a phenomenon that can be observed in nature. Consequently, only highly energetic wave processes are considered because at least one component in $\mathcal{M}^*$ exceeds its threshold. In this application, marginal thresholds correspond to marginal 0.99-quantiles.

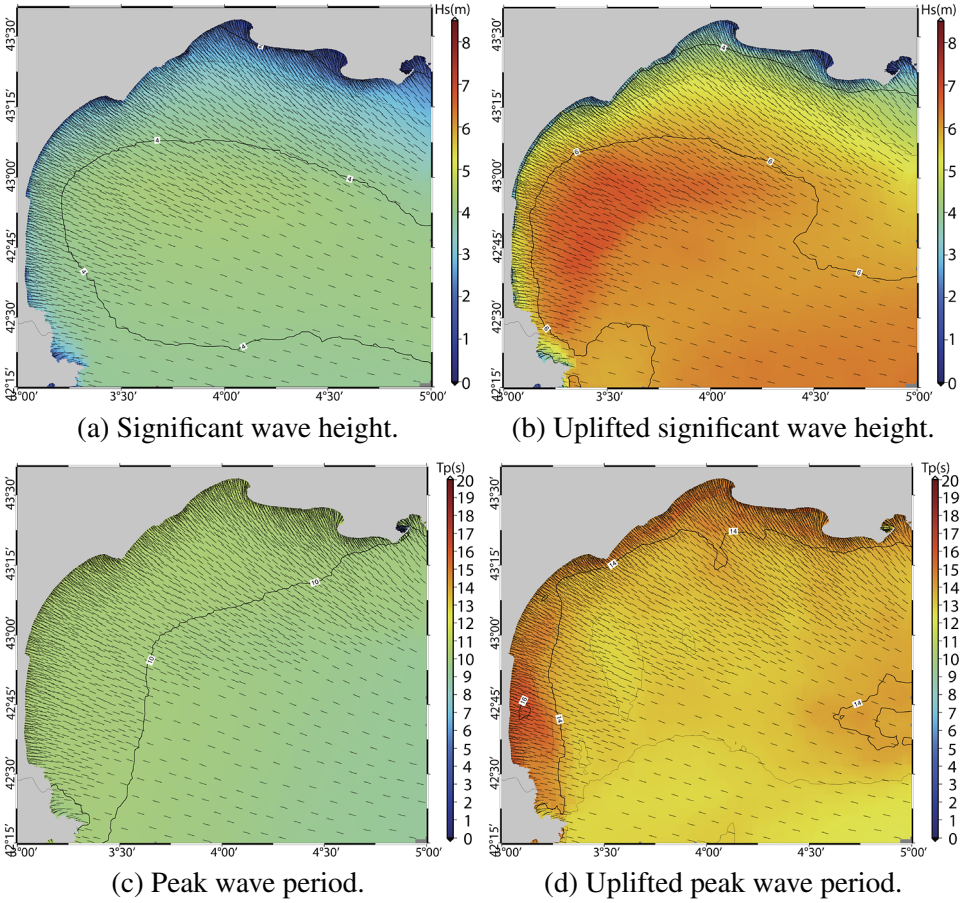
We are concerned with modelling storms that impact the coastline only. Hence, we chose to set S^* equal to the coastline-band area illustrated in Figure 4. This restriction in the storm detection area prevents the selection of offshore storms that do not propagate to the coast in the execution of Algorithm 1.

From the preliminary analysis in Section 2.2, we determine that storms last about 50 hours: the duration for which the extremal coefficient appears to be steady, revealing a persistence of the dependence structure within a storm up to that time. Thus, the selected value of δ is equal to 24 (hours). To select only i.i.d. storms, the value of ε is also equal to 24 (hours). This parameter is set to avoid the selection of overlapping storms. In this application, it would have been set to 0 without any consequence since no overlaps had been detected in this configuration.

Both $\zeta_{i,\mathrm{Hs}}$ and $\zeta_{i,Ts}$ are chosen to uplift original storms to m-year return period storms following the implementation of Groeneweg, Caires and Roscoe (2012). It is remarkable that any uplifted storms in the same return period might be compared to realisations of the distribution of the storms at this return period. Hence, having the control on the return period of storms is the easiest way to interpret and compare the impacts of storms from a coastal engineering point of view. In this application, $s_{\max}$—the within storm maxima—is chosen among the entire set $\mathcal{M}^*$ of locations available in the littoral area. The location $s_{\max}$ might be different for the two variables. Figure 6 illustrates one of the uplifted storms.

Note that mean wave directions are conserved during the uplift procedure.

Among the set of 10 scenarios, the variability of the fields observed are quite large, but are unsurprisingly dominated by fluxes from the south, southeast or east. This is a direct consequence of choosing $\mathcal{M}^*$ as a very littoral area.



(a) Significant wave height. (b) Uplifted significant wave height.

(c) Peak wave period. (d) Uplifted peak wave period.

FIG. 6. *Comparison of a storm uplifted to its* 100*-year return period, at its peak. The left panels illustrate the original storm; the right panels illustrate the uplifted storm. The arrows indicate the mean wave directions.*

4.2. *Uplifted storms at work: A risk analysis.* Coastal hazards such as submersion, erosion or beach contamination are usually quantified from formulae that require the computation of mass flux of energy towards the shoreline, given off the shoaling zone where waves do not interact significantly with the sea bottom. We usually distinguish between cross-shore and long-shore contributions, depending upon the goal of the application. For instance, the calculation of the alongshore-sand transport [Bagnold (1966), CERC (1984)] requires the long shore mass flux of energy. In the following, we strictly consider the long- shore impact ϕ of the deep water mass flux of energy $\mathcal{Q}$ to the shoreline, which is a relevant expression to tackle any analysis of shoreline dynamics. We model evolution of such a quantity during extreme wave storms.

For a given storm event S, we compute the impact $\phi_{i,t}^{(S)}$ at a location $c_i \in \mathcal{C}$ and at a time t of the mass flux of energy $\mathcal{Q}_{i,t}$ coming from waves at a location $l_i \in \mathcal{L}$ (see Figure 7). The long-shore impact is calculated by

$$\phi_{i,t}^{(S)} = \mathcal{Q}_{i,t} \sin(\omega_{i,t}) \cos(\omega_{i,t}), \tag{4.1}$$

where $\omega_{i,t}$ represents the angle of the wave propagation at l_i at a time t and is function of the wave direction $\psi_{i,t}$.

Practically, $\mathcal{Q}$ is derived from the variables Hs, Tp characterising the sea-state conditions at various points along an iso-bathymetric baseline. Such a mass flux of

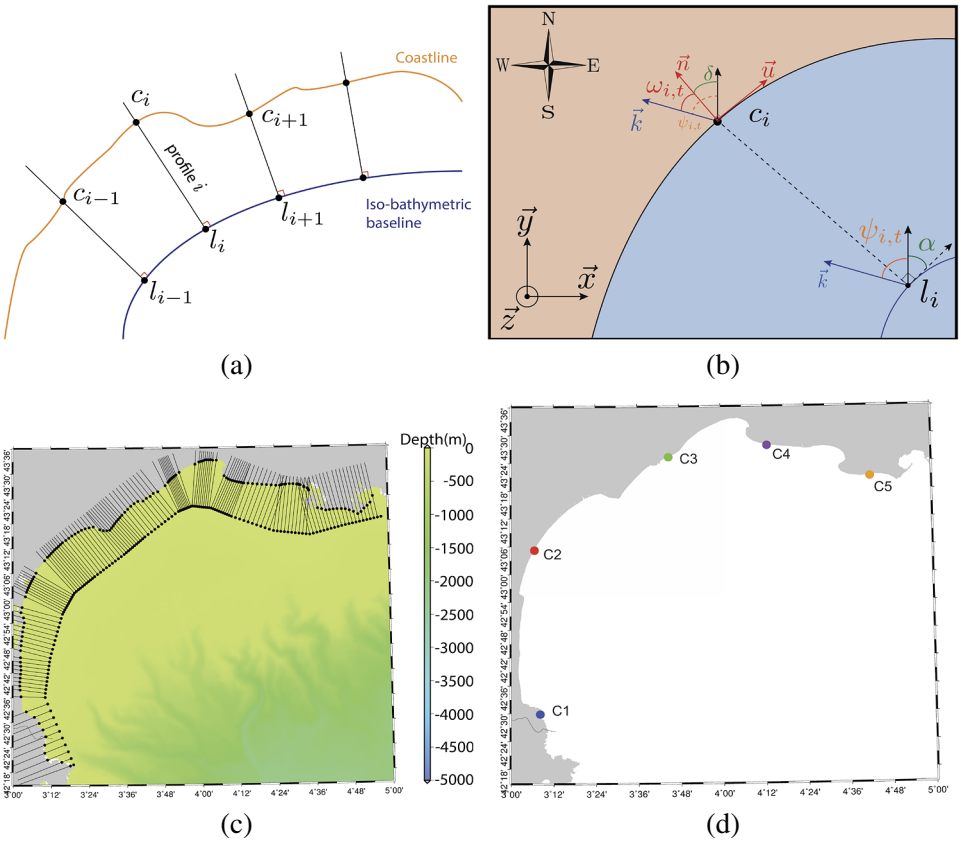


FIG. 7. (a) *A schematic representation of the baseline and the creation of the n profiles*. (b) *Illustration of angles used to compute the impact of the wave energy flux at point* l_i *to its coupled coast point* c_i. ωi_t *denotes the angle of interest*: *the angle between the observed direction of the waves* $\vec{k}$ *at location* l_i*—at a time t—and the cross-shore direction at location* c_i *denoted* $\vec{n}_i$. (c) *The actual profile construction over the GOL. Sea-states conditions are picked-up from a set* $\mathcal{L} = \{l_1, \ldots, l_n\}$ *of n points lying on an iso-bathymetric baseline. From those locations, n profiles normal to the baseline are created. The intersections of those profiles with the coastline derived form a set* $\mathcal{C} = \{c_1, \ldots, c_n\}$ *denoting the reference locations where mass flux energy are derived to. The number n is chosen to fit the resolution required along the shore*. (d) *The selected five locations analysed in the risk analysis*.

energy is classically given by

$$\mathcal{Q}_{i,t} = \frac{1}{8}\rho g \mathrm{Hs}_{i,t}^2 \mathrm{Tp}_{i,t}, \tag{4.2}$$

where ρ denotes the water volumetric mass density and g the gravity constant.

This procedure can be performed both with the storms extracted from the hindcast dataset to monitor the impact of the past events, or with the uplifted storms to assess the impact on the coast of more severe storms.

We compute the long-shore impact at any location c_i for some of the simulated (very) extreme storms. A set of 5 locations from the available c_i [see Figure 7(d)] has been picked as a reference to discuss the assessment of the long-shore impact at the coastline of the GOL under extreme conditions. These locations are manually selected to provide a good covering of the coastline with only few locations for the sake of clarity.

Regarding the angles presented in Figure 7, a positive value of ψ is interpreted as a long-shore contribution in the direction of $\vec{u}$—the tangent at the coast. A negative value is interpreted as a long-shore contribution in the opposite direction, that is, $-\vec{u}$.

Figure 8 gives an overview of the various possibilities offered by the simulation of storms in the assessment of long-shore impact.

First, Figure 8(a) shows the response of the impact model at the 5 reference locations to an uplifted storm at a 100-year return period. Regarding this figure, it is very clear that in this configuration c_2, c_3 and c_4 are impacted towards the west and southwest directions, revealing the presence of an eastern wave forcing. By contrast and since $\psi > 0$ at c_1, this site is impacted towards $\vec{u}$, that is, to the north or northwest at c_1. From such a figure, the time evolution of the long-shore impact regarding the simulated extreme process can be explored.

We may also look at the variability of the long-shore impact when storms vary in extremeness, as defined above. Figure 8(b) represents what could be expected in terms of long-shore impact, at one location and for a given storm uplifted to various return periods.

Another interesting information in the assessment of long-shore impact is to look at the response ψ for several storms uplifted to the 100-year return period. This is illustrated in Figure 8(c) for the point c_5, which is situated at the very east of the GOL. From this figure, we can state that the long-shore impact is likely to be towards the west, catching a consequent amount of energy from the storm coming from the open sea boundary of the GOL (i.e., from the east/southeast). This remark is in accordance with a physical observation that is identified when looking at the shoreline: the formation of sandy spits.

However, and still in Figure 8(c), some of the selected storms have a positive impact during their realisation. This is not really surprising, because as it is located at the edge of the GOL, this shoreline location is also subject to be hit by south

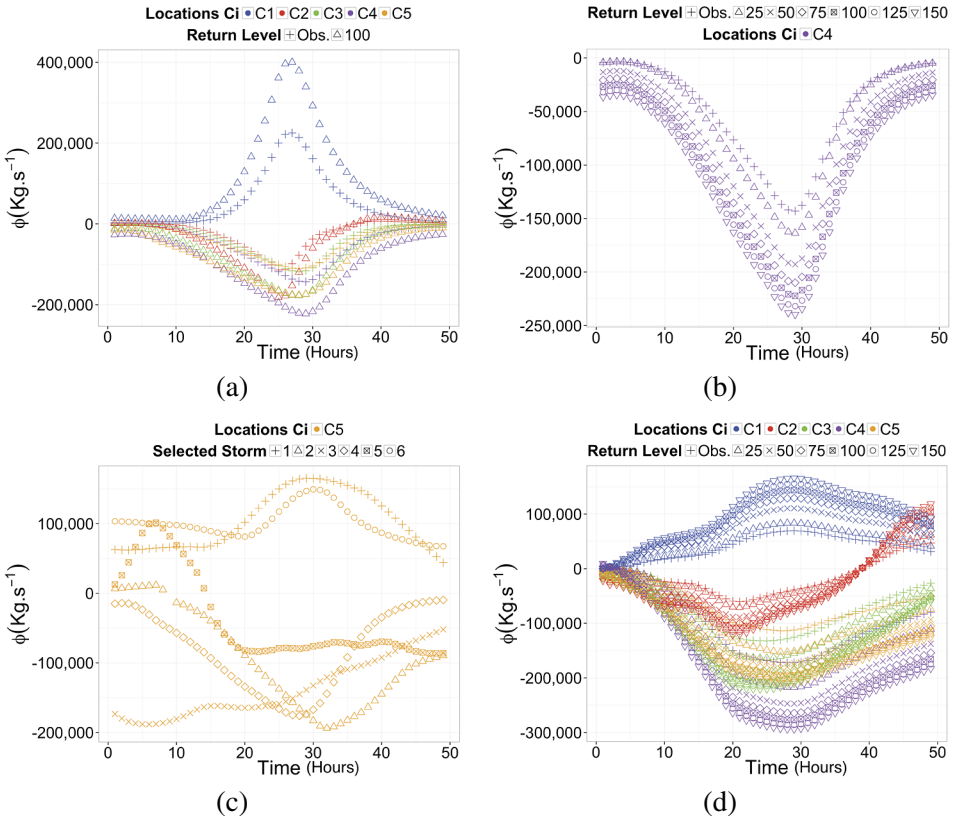


FIG. 8. *Evaluation of the long-shore impact* ϕ; (a) *at the* 5 *locations* c_i *for an observed storm uplifted to the* 100*-year return period*; (b) *at the location* c_4 *for an uplifted storm to the* 25, 50, 75, 100, 125 *and* 150*-year return periods. The impact computed from values of the observed storm are given as well for reference*; (c) *at the location* c_5 *for a sample of observed storms, uplifted to the* 100*-year return period*; (d) *at the* 5 *locations* c_i *for an observed storm uplifted to the* 25, 50, 75, 100, 125 *and* 150*-year return periods. The impact computed from values of the observed storm are also given for reference.*

and southwest storms, which are less frequent but even more damaging than the eastern ones.

Finally, Figure 8(d) is a mix of the possible combinations. It provides a simultaneous preview for various return periods of the storm and at the 5 locations of interest. Spatial patterns of long-shore impact regarding the intensity of a storm might be determined from such a figure.

5. Discussion. We introduced a semiparametric approach to simulate bivariate extreme space–time wave processes. Our motivation was to simulate more extreme storms than those already observed in order to assess event-scale coastal

hazards in such situations. In practice, these storms would feed physical littoral models, which depend strongly on the time evolution of the forcing extreme event.

We applied the methodology presented on a reanalysis dataset covering the GOL area in the northwestern Mediterranean sea.

To demonstrate the benefits of such a method, some simulated storms were used in a risk analysis. Thanks to the simulated processes based on which a control of the extremeness is provided, we showed that the variability of the littoral longshore impact can be assessed, both spatially and through the time evolution. Such results are of the utmost interest in coastal engineering applications, such as the construction of seawalls along the coastline.

This method is especially suitable for its relatively low-cost computational requirement. Indeed, the highest demand concerns the marginal fits, which is an easily parallelisable code. Simulating a set of extreme storms with a physical model would take days where our proposed method will take hours. The proposed method can therefore be applied on massive space–time dataset, as described in this application. Mathematically justified, this method reaches its goal to seamlessly simulate reliable space–time extreme events at a more extreme scale than the ones observed.

However, some limits of the method itself and its implementation should be highlighted. As often when dealing with EVT approaches, we suppose that the underlying dependence structure through time and space is preserved from extreme but observable events to more extreme events. However, it is difficult to physically validate this assumption. As emphasized by Bortot, Coles and Tawn (2000), asymptotic dependence is a limiting property which cannot be verified with certainty from data alone. Usually, the check of the extremal dependence structure relies on modelling properties, arguing from reasonable agreements between empirical and model-based estimates of particular extremal probabilities. Unfortunately, such checking procedures are not possible under our approach, since no particular form of extremal dependence is assumed. As a consequence, if our assumption of a constant space–time extremal dependence for small lags is not satisfied, our approach would lead to an overestimation of the extremal dependence.

Because we are dealing with bivariate space–time processes only, we assume that the third variable defining sea-states conditions [namely the direction $\psi(t)$] remains unchanged in distribution for highest storms. There are good physical reasons to make this hypothesis, such as the GOL orientation, which will never change. Indeed, the open boundary of the GOL, which is southeast oriented, will naturally prohibit the observation of high waves being southeast oriented near the coastline [see Chailan et al. (2014) for further details on the GOL orientation and the implied fetch constraint]. Hence, it seems appropriate to conserve the wave directions from observed storms for heavier storms to keep them physically valid. Consequently, this restriction on the wave directions of the simulated storms to those that have already been observed can be seen both as a strength and a limitation.

From a more practical point of view, it could be argued that the storm size in the Algorithm 1 is fixed and symmetric around the peak value of the storm. This may not reflect the reality for all storms. Therefore, replacing the current fixed size by an adaptive one might be of interest to better represent those storms.

Note that there is no limitation in the methodology to select smaller or longer storms, conditionally to the fact that at least one component excesses its marginal threshold. Regarding coastal risk assessment analysis, selecting smaller storms would result in an underestimation of the length of a storm, and consequently of the overall quantity of interest (e.g., wave energy). By taking storms lasting too long, the opposite may occur. In such cases, selecting the rightful duration of a storm is a true challenge. The use of the extremal coefficient expressing the temporal dependence within storms is by definition a good indicator to determine the storm duration.

In the Algorithm 1, ε is set to avoid the selection of dependent storms and therefore respect statistical assumptions. To avoid dependent storms, it is convenient to always set $\varepsilon = \delta$.

Other parameters of the algorithm can be debated, such as the littoral area S^*. Because its definition is paramount to assess littoral hazards, it could be interesting to evaluate the sensitivity of the storm detection regarding this area.

In this first approach and even if seasonality is found in the data, fixed marginal thresholds are used for the margins transformation. It would be valuable to use more sophisticated expressions of the thresholds to handle the nonstationarity of the data. The use of directional covariates in the thresholds rather than omnidirectional ones might also significantly improve the marginal fits.

In this paper, we have not addressed the estimation uncertainties on marginal fits and their propagation. Block bootstrapping is usually used for assessing such uncertainties. Nevertheless, one practical difficulty is the choice of the blocks to consider, especially in a space–time context. In a similar vein, the validation of the uplifted storms is hard to afford, if not impossible. We recommend using techniques inspired by cross-validation, but practical limitations arise. Uplifted storms are multivariate space–time processes and the first constraint is to find a measure to compare them. Assuming a reduced-dimension measure, the second limit is any uplifted storm that has to be seen as a realisation from the multivariate space–time distribution of storms. Unfortunately, this distribution might only be estimated empirically and many realisations must be used to estimate it correctly. Yet we do not possess enough extreme realisations, by definition. One way to avoid the lack of realisations is to lower the threshold to detect storms in our dataset. Unfortunately, doing so would violate the hypothesis of the method: the need for an exceedance over a "high" threshold, to approximate the asymptotic results.

Beyond those limitations, this method appears promising and opens many perspectives. It would be interesting to extend this approach to the multivariate context because that would allow us to integrate additional variables describing the

environmental phenomenon at a very extreme scale. However, the underlying dependence structures of the considered variables must be thoroughly investigated before being able to justify this extension with the presented assumptions.

Another perspective of work is to apply the method on larger regions. For such an application, the choice of letting $s_{\max}$ be located respectively to each variable should be reviewed. With a wider area, various and independent physical processes might be caught at the peak of storm. Consequently, $\zeta_{i,\mathrm{Hs}}$ and $\zeta_{i,\mathrm{Tp}}$ can be determined from two different processes. On this basis, the uplifted storms might be physically unrealistic.

Other datasets and applications could also be considered. Most notably, we are interested in applying this method in the context of rain-storms. Such an application would allow us to explore the space–time variability of extreme rain-storms scenarios with a plenty set of derived applications. For instance, a simulated scenario can then feed a rainfall-runoff model to study their consequences in terms of floods.

A future work would be the comparison between simulated storms issued by the presented semiparametric approach and those issued by other parametric approaches, and in particular the generalized Pareto processes. Such a comparison would be valuable since both approaches present similarities.

At the same time and after having performed a small risk analysis using some of the simulated extreme space–time waves events, one challenge is to use those storms to feed heavy computational physical models assessing other coastal hazards, such as a flood overland model. In our opinion, this challenge may represent the foundation of the next generation of coastal flood early warning systems such as Delaware's Coastal Flood Monitoring System (CFMS).[2]

APPENDIX SECTION

A mathematical justification of the storm uplift can be obtained through the following asymptotic equivalence for conditional distributions.

Indeed, following Caires, de Haan and Smith (2011),

$$P\left(\frac{T^{\leftarrow}(\zeta_i Z_i^*(s,t)) - b_{n\zeta_i}}{a_{n\zeta_i}} \in A \Big| \max_{s\in\mathcal{M}^*} Z_i^*(s,t) > 1\right)$$

has the same limit (as $n \to \infty$) as

$$P\left(\frac{Z_i(s,t) - b_n}{a_n} \in A \Big| \max_{s\in\mathcal{M}^*} Z_i^*(s,t) > 1\right),$$

where $Z_i^*(s,t) = [1 + \xi(s,t)(Z_i(s,t) - b_n(s,t))/a_n(s,t)]^{1/\xi(s,t)}$ and $T^{\leftarrow}(y) = b_n + a_n \frac{y^{\xi}-1}{\xi}$.

[2] http://coastal-flood.udel.edu/

Let us drop both i and (s,t) indexes for the sake of simplicity in the left part of the conditional probability. The former limit equivalence is valid since following Ferreira and de Haan (2014)-Section 4.2,

$$
\begin{aligned}
&P\left(\frac{T^{\leftarrow}(\zeta Z^*)-b_{n\zeta}}{a_{n\zeta}}\in A\Big|\max_{s\in\mathcal{M}^*}Z_i^*(s,t)>1\right)\\
&\quad=P\left(\frac{a_n}{a_{n\zeta}}\frac{[\zeta^{\xi}(1+\xi\frac{Z-b_n}{a_n})]-1}{\xi}-\frac{b_{n\zeta}-b_n}{a_{n\zeta}}\in A\Big|\max_{s\in\mathcal{M}^*}Z_i^*(s,t)>1\right)\\
&\quad=P\left(\frac{a_n\zeta^{\xi}}{a_{n\zeta}}\frac{1+\xi\frac{Z-b_n}{a_n}-\zeta^{-\xi}}{\xi}-\frac{b_{n\zeta}-b_n}{a_{n\zeta}}\in A\Big|\max_{s\in\mathcal{M}^*}Z_i^*(s,t)>1\right)\\
&\quad=P\left(\frac{a_n\zeta^{\xi}}{a_{n\zeta}}\left(\frac{Z-b_n}{a_n}-\zeta^{-\xi}\left[\frac{b_{n\zeta}-b_n}{a_n}-\frac{\zeta^{\xi}-1}{\xi}\right]\right)\in A\Big|\max_{s\in\mathcal{M}^*}Z_i^*(s,t)>1\right)\\
&\quad=P\left(\frac{Z-b_n}{a_n}\in\frac{a_{n\zeta}\zeta^{-\xi}}{a_n}A+\zeta^{-\xi}\left(\frac{b_{n\zeta}-b_n}{a_n}-\frac{\zeta^{\xi}-1}{\xi}\right)\Big|\max_{s\in\mathcal{M}^*}Z_i^*(s,t)>1\right).
\end{aligned}
$$

From de Haan and Ferreira (2006), it can be deduced that:

1. $\lim_{n\to\infty}\frac{a_{n\zeta}}{a_n}=\zeta^{\xi}$ (see proof of Lemma 1.2.9, p. 24);
2. $\lim_{n\to\infty}\frac{b_{n\zeta}-b_n}{a_n}=\frac{\zeta^{\xi}-1}{\xi}$ [consider $U(n)$ as in Theorem 1.1.2 for decomposing $\frac{b_{n\zeta}-b_n}{a_n}$ as $\frac{U(n)-b_n}{a_n}-\frac{U_{n\zeta}-b_{n\zeta}}{a_{n\zeta}}\frac{a_{n\zeta}}{a_n}+\frac{U(n\zeta)-U(n)}{a_n}$ and use 1.1.20, p. 10].

Then

$$
\frac{a_{n\zeta}\zeta^{-\xi}}{a_n}\to 1\quad\text{and}\quad\zeta^{-\xi}\left(\frac{b_{n\zeta}-b_n}{a_n}-\frac{\zeta^{\xi}-1}{\xi}\right)\to 0
$$

uniformly for $(s,t)\in S\times\mathcal{T}$ as $n\to\infty$ and the result follows using the convergence to types theorem [see Embrechts, Klüppelberg and Mikosch (1997), Theorem A1.5].

There is no limitation to extend this reasoning to a bivariate context. In accordance with our chosen approach, only one of the two considered processes is concerned with the conditional event. Indeed, the conditional event has no impact on the aforementioned probability developments. Hence, we can similarly show that

$$
\begin{aligned}
P\Bigg(&\frac{T_1^{\leftarrow}(\zeta_{1,i}Z_{1,i}^*(s,t))-b_{1,n\zeta_i}}{a_{1,n\zeta_i}}\in A_1,\\
&\frac{T_2^{\leftarrow}(\zeta_{2,i}Z_{2,i}^*(s,t))-b_{2,n\zeta_i}}{a_{2,n\zeta_i}}\in A_2\Big|\max_{s\in\mathcal{M}^*}Z_{1,i}^*(s,t)>1\Bigg),
\end{aligned}
$$

where $T_1^{\leftarrow}(y) = b_{1,n} + a_{1,n}\frac{y^{\xi_1}-1}{\xi_1}$ and $T_2^{\leftarrow}(y) = b_{2,n} + a_{2,n}\frac{y^{\xi_2}-1}{\xi_2}$, has the same limit (as $n \to \infty$) as

$$P\left(\frac{Z_{1,i}(s,t) - b_{1,n}}{a_{1,n}} \in A_1, \frac{Z_{2,i}(s,t) - b_{2,n}}{a_{2,n}} \in A_2 \Big| \max_{s\in\mathcal{M}^*} Z_{1,i}^*(s,t) > 1\right).$$

Acknowledgments. We would like to thank A. Laurent and F. Bouchette who have both contributed significantly to this document.

This work was performed within the context of both the LITTOCMS and MISTRAL-LITTORAL projects. It also forms part of the GLADYS initiative; see www.gladys-littoral.org.

REFERENCES

BAGNOLD, R. A. (1966). An approach to the sediment transport problem. General Physics Geological Survey, Prof. Paper.

BECHLER, A., BEL, L. and VRAC, M. (2015). Conditional simulations of the extremal t process: Application to fields of extreme precipitation. *Spat. Stat.* **12** 109–127. MR3346645

BEIRLANT, J., GOEGEBEUR, Y., TEUGELS, J. and SEGERS, J. (2004). *Statistics of Extremes. Theory and Applications*. Wiley, Chichester. MR2108013

BORTOT, P., COLES, S. and TAWN, J. (2000). The multivariate Gaussian tail model: An application to oceanographic data. *J. Roy. Statist. Soc. Ser. C* **49** 31–49. MR1817873

BOUCHETTE, F., SABATIER, F., SYLAIOS, G., MEULÉ, S., LIOU, J. L., HEURTEFEUX, H., DENAMIEL, C. and HWUNG, W. (2012). SUBDUNE tool: Quasiexplicit formulation of the water level along the shoreline. *Rev. Paralia* **12** 223–232.

BRUNEL, C., CERTAIN, R., SABATIER, F., ROBIN, N., BARUSSEAU, J. P., ALEMAN, N. and RAYNAL, O. (2014). 20th century sediment budget trends on the Western Gulf of Lions shoreface (France): An application of an integrated method for the study of sediment coastal reservoirs. *Geomorphology* **204** 625–637.

CAIRES, S., DE HAAN, L. and SMITH, R. L. (2011). On the determination of the temporal and spatial evolution of extreme events Technical Report, Deltares. Report 1202120-001-HYE-004 (for Rijkswaterstaat, Centre for Water Management).

CAMPMAS, L., BOUCHETTE, F., MEULE, S., PETITJEAN, L., SOUS, D., LIOU, J.-Y., LEROUX-MALLOUF, R., SABATIER, F. and HWUNG, H.-H. (2014). Typhoons driven morphodynamics of the Wan Tzu Liao sand barrier (Taiwan). *Coastal Eng. Proc.* **1** sediment–50.

CASTRUCCIO, S., HUSER, R. and GENTON, M. G. (2016). High-order composite likelihood inference for max-stable distributions and processes. *J. Comput. Graph. Statist.* **25** 1212–1229. MR3572037

CERC (1984). Shore Protection Manual.

CHAILAN, R. (2015). Application of Scientific Computing and Statistical Analysis to Address Coastal Hazards Ph.D. thesis University of Montpellier.

CHAILAN, R., TOULEMONDE, G., BOUCHETTE, F., LAURENT, A., SEVAULT, F. and MICHAUD, H. (2014). Spatial assessment of extreme significant waves heights in the Gulf of Lions. *Coastal Eng. Proc.* **1** management–17.

COLES, S. (2001). *An Introduction to Statistical Modeling of Extreme Values*. Springer, London. MR1932132

DAVIS, R. A., KLÜPPELBERG, C. and STEINKOHL, C. (2013a). Max-stable processes for modeling extremes observed in space and time. *J. Korean Statist. Soc.* **42** 399–414. MR3255398

DAVIS, R. A., KLÜPPELBERG, C. and STEINKOHL, C. (2013b). Statistical inference for max-stable processes in space and time. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **75** 791–819. MR3124792

DAVISON, A. C. and HUSER, R. (2015). Statistics of extremes. *Annu. Rev. Stat. Appl.* **2** 203–235.

DAVISON, A. C., PADOAN, S. A. and RIBATET, M. (2012). Statistical modeling of spatial extremes. *Statist. Sci.* **27** 161–186. MR2963980

DE HAAN, L. (1984). A spectral representation for max-stable processes. *Ann. Probab.* **12** 1194–1204. MR0757776

DE HAAN, L. and DE RONDE, J. (1998). Sea and wind: Multivariate extremes at work. *Extremes* **1** 7–45. MR1652944

DE HAAN, L. and FERREIRA, A. (2006). *Extreme Value Theory*. Springer, New York. MR2234156

DE HAAN, L. and LIN, T. (2001). On convergence toward an extreme value distribution in $C[0, 1]$. *Ann. Probab.* **29** 467–483. MR1825160

DIEKER, A. B. and MIKOSCH, T. (2015). Exact simulation of Brown-Resnick random fields at a finite number of locations. *Extremes* **18** 301–314. MR3351818

DOMBRY, C., ENGELKE, S. and OESTING, M. (2016). Exact simulation of max-stable processes. *Biometrika* **103** 303–317. MR3509888

DOMBRY, C. and EYI-MINKO, F. (2013). Regular conditional distributions of continuous max-infinitely divisible random fields. *Electron. J. Probab.* **18** 1–21. MR3024101

DOMBRY, C., ÉYI-MINKO, F. and RIBATET, M. (2013). Conditional simulation of max-stable processes. *Biometrika* **100** 111–124. MR3034327

DOMBRY, C. and RIBATET, M. (2015). Functional regular variations, Pareto processes and peaks over threshold. *Stat. Interface* **8** 9–17. MR3320385

EASTOE, E. F. and TAWN, J. A. (2009). Modelling non-stationary extremes with application to surface level ozone. *J. R. Stat. Soc. Ser. C. Appl. Stat.* **58** 25–45. MR2662232

EMBRECHTS, P., KLÜPPELBERG, C. and MIKOSCH, T. (1997). *Modelling Extremal Events. Applications of Mathematics* (*New York*) **33**. Springer, Berlin. MR1458613

EMBRECHTS, P., KOCH, E. and ROBERT, C. (2016). Space-time max-stable models with spectral separability. *Adv. in Appl. Probab.* **48** 77–97. MR3539298

ENGELKE, S., MALINOWSKI, A., KABLUCHKO, Z. and SCHLATHER, M. (2015). Estimation of Hüsler-Reiss distributions and Brown-Resnick processes. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **77** 239–265. MR3299407

FERREIRA, A. and DE HAAN, L. (2014). The generalized Pareto process; with a view towards application and simulation. *Bernoulli* **20** 1717–1737. MR3263087

GOULDBY, B., MÉNDEZ, F. J., GUANCHE, Y., RUEDA, A. and MÍNGUEZ, R. (2014). A methodology for deriving extreme nearshore sea conditions for structural design and flood risk analysis. *Coastal Eng.* **88** 15–26.

GROENEWEG, J., CAIRES, S. and ROSCOE, K. (2012). Temporal and spatial evolution of extreme events. *Coastal Eng. Proc.* **1** management–9.

GUTIERREZ, B. T., PLANT, N. G., THIELER, E. R. and TURECEK, A. (2015). Using a Bayesian network to predict barrier island geomorphologic characteristics. *J. Geophys. Res., Earth Surf.* **120** 2452–2475.

HERRMANN, M. and SOMOT, S. (2008). Relevance of ERA40 dynamical downscaling for modeling deep convection in the Mediterranean Sea. *Geophys. Res. Lett.* **35**.

HERRMANN, M., SEVAULT, F., BEUVIER, J. and SOMOT, S. (2010). What induced the exceptional 2005 convection event in the northwestern Mediterranean basin? Answers from a modeling study. *J. Geophys. Res., Oceans* (1978–2012) **115**.

HUSER, R. and DAVISON, A. C. (2013). Composite likelihood estimation for the Brown-Resnick process. *Biometrika* **100** 511–518. MR3068451

HUSER, R. and DAVISON, A. C. (2014). Space-time modelling of extreme events. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **76** 439–461. MR3164873

JONATHAN, P., EWANS, K. and RANDELL, D. (2013). Joint modelling of extreme ocean environments incorporating covariate effects. *Coastal Eng*. **79** 22–31.

LANTUÉJOUL, C. and BEL, L. (2014). Simulation conditionnelle du processus de Schlather. In 46*èmes Journées de Statistique de la SFdS*.

MICHAUD, H. (2011). Impacts des vagues sur les courants marins: Modélisation multi-échelle de la plage au plateau continental Ph.D. thesis Université Montpellier II-Sciences et Techniques du Languedoc.

MICHAUD, H., ROBIN, N., ESTOURNEL, C., MARSALEIX, P., LEREDDE, Y., CERTAIN, R. and BOUCHETTE, F. (2013). 3D hydrodynamic modeling of a microtidal barred beach (Sète, NW Mediterranean Sea) during storm conditions. In *Proc*. 7*th Int*. *Conf. on Coastal Dynamics*, *Arcachon France* **139** 1183–1194.

PADOAN, S. A., RIBATET, M. and SISSON, S. A. (2010). Likelihood-based inference for max-stable processes. *J. Amer. Statist. Assoc*. **105** 263–277. MR2757202

RAILLARD, N., AILLIOT, P. and YAO, J. (2014). Modeling extreme values of processes observed at irregular time steps: Application to significant wave height. *Ann*. *Appl*. *Stat*. **8** 622–647. MR3192005

RASCLE, N. and ARDHUIN, F. (2013). A global wave parameter database for geophysical applications. Part 2: Model validation with improved source term parameterization. *Ocean Model*. **70** 174–188.

RIBATET, M., COOLEY, D. and DAVISON, A. C. (2012). Bayesian inference from composite likelihoods, with an application to spatial extremes. *Statist*. *Sinica* **22** 813–845. MR2954363

SCHLATHER, M. and TAWN, J. A. (2003). A dependence measure for multivariate and spatial extreme values: Properties and inference. *Biometrika* **90** 139–156. MR1966556

SHABY, B. A. (2014). The open-faced sandwich adjustment for MCMC using estimating functions. *J. Comput*. *Graph*. *Statist*. **23** 853–876. MR3224659

SHABY, B. A. and REICH, B. J. (2012). Bayesian spatial extreme value analysis to assess the changing risk of concurrent high temperatures across large portions of European cropland. *Environmetrics* **23** 638–648. MR3019056

SMITH, R. L. (1990). Max-stable processes and spatial extremes. Preprint. Univ. Surrey.

THIBAUD, E. and OPITZ, T. (2015). Efficient inference and simulation for elliptical Pareto processes. *Biometrika* **102** 855–870. MR3431558

TOLMAN, H. L. (2014). User Manual and System Documentation of WAVEWATCH III® version 4.18. Technical Report 316.

WADSWORTH, J. L. and TAWN, J. A. (2014). Efficient inference for spatial extreme value processes associated to log-Gaussian random functions. *Biometrika* **101** 1–15. MR3180654

WANG, Y. and STOEV, S. A. (2011). Conditional sampling for spectrally discrete max-stable random fields. *Adv. in Appl. Probab*. **43** 461–483. MR2848386

UNIVERSITY OF MONTPELLIER
2 PLACE EUGÈNE BATAILLON
34095 MONTPELLIER CEDEX 5
OCCITANIE
FRANCE
E-MAIL: romain.chailan@umontpellier.fr
gwladys.toulemonde@umontpellier.fr
jean-noel.bacro@umontpellier.fr