Received: 20 May 2018 | Revised: 5 September 2018 | Accepted: 7 September 2018

DOI: 10.1002/env.2541

RESEARCH ARTICLE

WILEY

A model for the directional evolution of severe ocean storms

S. Tendijck[1] | E. Ross[2] | D. Randell[2] | P. Jonathan[3,4]

[1]TU Delft Electrical Engineering, Mathematics and Computer Science, 2600 AA Delft, The Netherlands

[2]Shell Global Solutions International B.V., 1031 HW Amsterdam, The Netherlands

[3]Shell Research Ltd., London SE1 7NA, U.K.

[4]Department of Mathematics and Statistics, Lancaster University, Lancaster LA1 4YW, U.K.

Correspondence
P. Jonathan, Shell Research Ltd., London SE1 7NA, U.K.; or Department of Mathematics and Statistics, Lancaster University, Lancaster LA1 4YW, U.K.
Email: philip.jonathan@shell.com

Abstract

Motivated by recent work on Markov extremal models, we develop a nonstationary extension and use it to characterize the time evolution of extreme sea state significant wave height (H_S) and storm direction in the vicinity of the storm peak sea state. The approach first requires transformation of H_S from a physical to a standard Laplace scale achieved using a nonstationary directional marginal extreme value model. The evolution of Laplace-scale H_S is subsequently characterized using a Markov extremal model and that of the rate of change of storm direction described by an autoregressive model, the evolution variance of which is H_S-dependent. Simulations on the physical scale under the estimated model give realistic realizations of storm trajectories consistent with historical data for storm trajectories at a northern North Sea location.

KEYWORDS

extreme, Markov extremal model, nonstationary, significant wave height, storm trajectory

1 | INTRODUCTION

We are often interested in understanding the evolution of time series of extreme values, potentially nonstationary with respect to covariates. For example, in an oceanographic setting, there is interest in understanding the evolution of a severe ocean storm, consisting of a set of consecutive sea states, in time from sea state to sea state. This involves joint modeling of multivariate time series of a combination of variables, some of which (e.g., sea state significant wave height, henceforth H_S for brevity) are extreme and others (e.g., covariates such as sea state storm direction) are not. Such a model is important for reliable design of marine structures, enabling estimation of distributions of variables such as crest elevation, individual wave height, and total water level for whole storm events consisting of multiple dependent sea states (as opposed to estimates for isolated sea states) given storm peak characteristics. These can then be combined with distributions for the statistics of storm peak magnitudes and storm rates of occurrence to estimate design conditions for return periods of arbitrary length, as explained, for example, in the work of Feld, Randell, Ross, and Jonathan (2018). Such a model also allows the "directional dissipation" of an ocean storm to be characterized, facilitating consistent estimation of return values across directional sectors (e.g., Ross, Randell, Ewans, Feld, & Jonathan, 2017).

In the context of characterizing heat wave evolution, Winter and Tawn (2016, 2017) introduce a Markov extremal model (MEM) for an interval of extreme values of time series (transformed to a standard Laplace marginal scale) motivated by the conditional extremes model of Heffernan and Tawn (2004). In this work, we present a nonstationary extension of the work of Winter and Tawn incorporating evolution of sea state H_S and direction. In outline, the approach consists of the following steps: (a) a common directional marginal model based on the work of Ross et al. (2017) is established for H_S of all sea states of all storms, given direction; (b) the marginal model is used to transform the values of H_S given direction to the standard Laplace scale using the probability integral transform; (c) intervals of time series of threshold exceedances of Laplace-scale H_S (of different lengths, but including occurrence of the storm peak, H_S^{sp}) and corresponding storm direction (together referred to as "storm trajectories") are isolated, for each of a large number of storm events;

(d) a MEM of appropriate order is estimated to describe the evolution of Laplace-scale H_S relative to the storm peak; (e) an autoregressive time-series model of appropriate order with heterogeneous evolution variance dependent on Laplace-scale H_S is established for the rate of change of storm direction on its trajectory; (f) models from steps (a), (d), and (e) are coupled to facilitate simulation of realistic storm trajectories on a physical scale. The method developed here is motivated by ocean engineering requirements regarding characterization of storm evolution but is generally applicable to nonstationary time series of extreme events such as precipitation, wind, ocean current (including features such as solitons), and storm surge in an environmental context.

The key idea behind the proposed approach is the conditional extremes model of Heffernan and Tawn (2004). This model is advantageous for two main reasons. (a) It allows the characterization of more general forms of extremal temporal and spatial dependence (including both asymptotic dependence and asymptotic independence) compared with its main competitors based on max-stable process modeling (e.g., Chavez-Demoulin & Davison, 2012; Davison, Padoan, & Ribatet, 2012; Huser & Davison, 2014; Padoan, Ribatet, & Sisson, 2010), which typically admit only asymptotic dependence, although extensions to asymptotic independence and mixtures of asymptotic dependence and asymptotic independence have been proposed (e.g., Rootzen, Segers, & Wadsworth, 2018; Wadsworth, Tawn, Davison, & Elton, 2017). (b) It allows the partitioning of large spatio-temporal problems into small pairwise ones. This facilitates computationally efficient inference, including incorporation of the effects of covariates compared with competitors based on max-stable processes (but see also computationally elegant hierarchical max-stable models of Reich & Shaby, 2012; Stephenson, 2009).

The layout of the article is as follows. A motivating application is given in Section 2, involving estimation of trajectories of extreme sea state H_S and corresponding storm direction for a location in the northern North Sea. Section 3 provides a description of the model. Application of the model to the northern North Sea example is then outlined in Section 4. The discussion of results and conclusions is provided in Section 5. The method by which distributions of MEM residuals are estimated is outlined in the Appendix.

2 | MOTIVATING APPLICATION

Significant wave height H_S can be defined as four times the standard deviation of the ocean surface elevation at a spatial location for a specified period of observation. The application sample is taken from the hindcast of Reistad et al. (2011), which provides time series of significant wave height and (dominant) wave direction for three-hour sea states for the period September 1957 to December 2012 at a northern North Sea location. Extreme sea states in the North Sea are dominated by winter storms originating in the Atlantic Ocean and propagating eastwards across the northern part of the North Sea. Sea states at northern North Sea locations are usually more intense than those at the southern North Sea locations. Directions of propagation of extreme seas vary considerably with location, depending on land shadows of the British Isles, Scandinavia, and the coast of mainland Europe, and fetches associated with the Atlantic Ocean, the Norwegian Sea, and the North Sea itself. In the northern North Sea, the main fetches are the Norwegian Sea to the north, the Atlantic Ocean to the west, and the North Sea to the south. Extreme sea states from the directions of Scandinavia to the east and the British Isles to the south-west are not likely. The shielding by these land masses is more effective for southern North Sea locations, resulting in a similar directional distribution but reduced wave heights by comparison with northern North Sea locations.

With θ indicating the direction from which waves propagate measured clockwise from the north, Figure 1a shows a directional plot of the complete sample of H_S data in gray. Storm peak occurrences H_S^{sp} shown in black are local maxima of storm trajectories estimated, as outlined in Section 3. The effect of the land shadow of Norway in particular is visible at $\approx 220°$. Figure 1b gives a polar plot of storm trajectories for 15 typical storms. The lengths and directional characteristics of storm trajectories vary considerably. Figure 2 shows the evolution of H_S (panel (a)) and θ (panel (b)) for the same 15 storms. The objective of the current work is to estimate a realistic model for storm trajectories such as those illustrated in Figure 1b and Figure 2.

3 | MODEL

We estimate a model for the evolution of an interval I of time series of consecutive values for threshold exceedances of sea state significant wave height, and corresponding direction $\{Y_t, \Theta_t\}_{t \in I}$, including the storm peak event (i.e., the occurrence of the largest value of H_S for the storm), sampled at some constant rate. For convenience, we assume time labels are

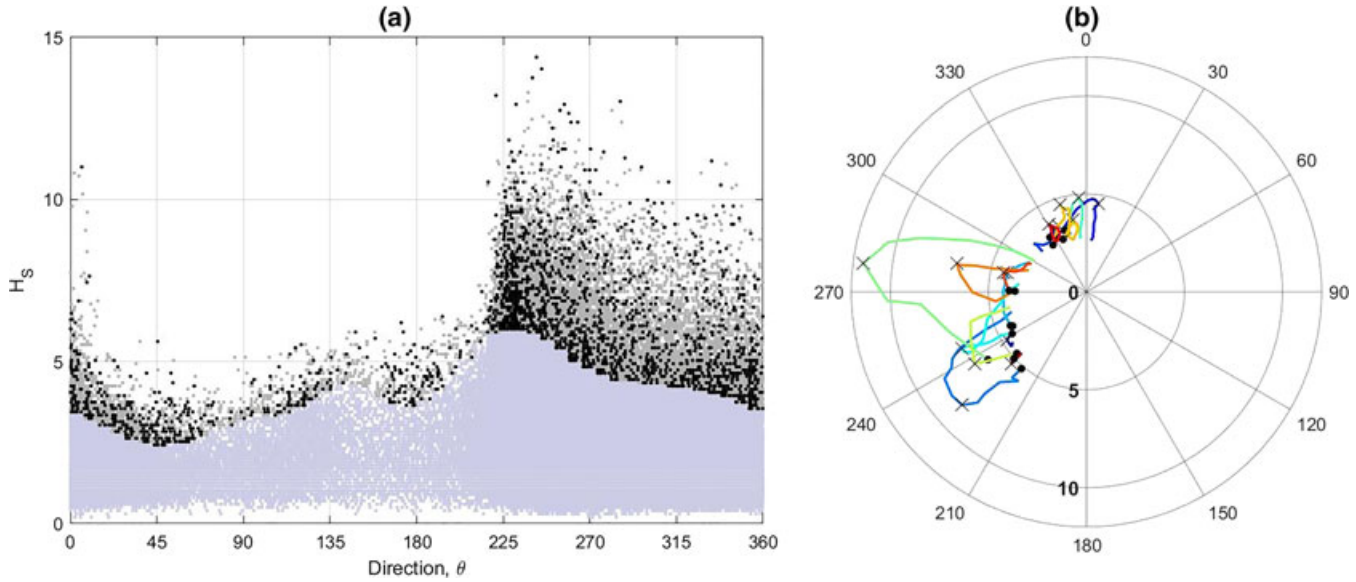

FIGURE 1 Motivating application. (a) Directional plot of storm peak H_S (black) and (sea state) H_S (gray) for all storms. (b) Polar plot of H_S trajectories with direction for 15 typical storms

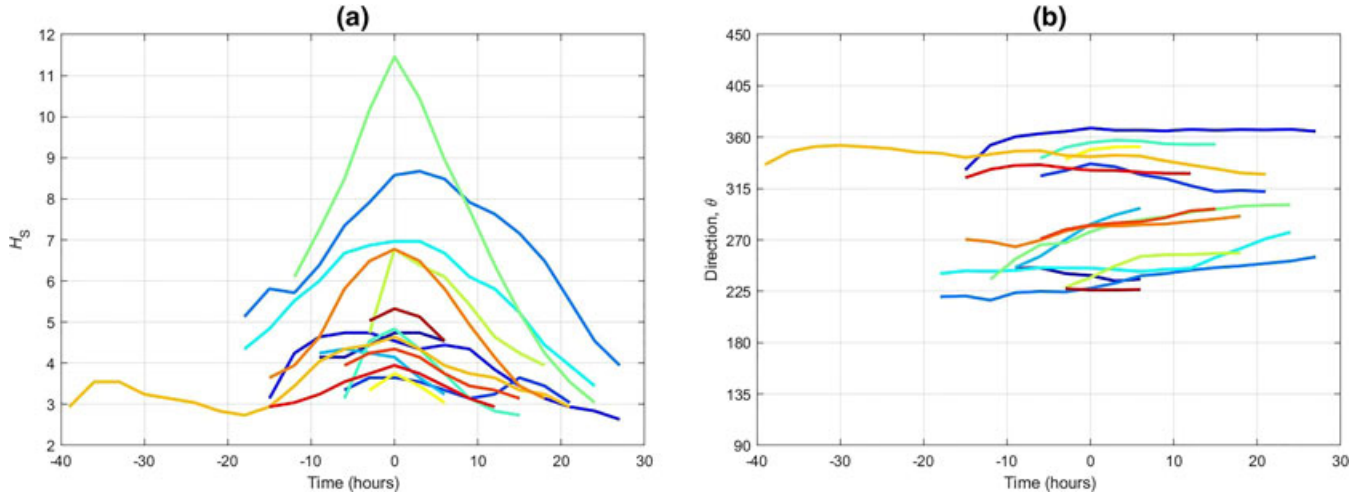

FIGURE 2 Time evolution for the 15 typical storms (shown also in Figure 1). (a) H_S in time. (b) θ in time

centered such that, for interval I, $t = 0 \in I$ corresponds to the time of the storm peak event. We estimate the model using a sample of m intervals $\{I_k\}$ of different lengths and data $\{\{y_{tk}, \theta_{tk}\}_{t \in I_k}\}_{k=1}^{m}$. The model is an extension of MEM of Winter and Tawn (2016, 2017), expressed for time series on a marginal standard Laplace scale. We therefore start the model description by outlining a nonstationary marginal model assumed applicable for all occurrences of H_S used to transform the sample of H_S values to the standard Laplace scale.

3.1 | Marginal modeling, transformation to Laplace scale, and isolation of storm trajectories

We assume marginally that sea state significant wave height Y (for all intervals and all times regardless of magnitude) follows a two-part truncated gamma-generalized Pareto distribution (see Ross et al., 2017) with density

$$f_Y(y|\gamma, \zeta, \xi, \nu, \psi) = \begin{cases} \tau \times f_{TG}(y|\gamma, \zeta, \psi) \text{ for } y \leq \psi \\ (1-\tau) \times f_{GP}(y|\xi, \nu, \psi) \text{ for } y > \psi, \end{cases}$$

where gamma shape γ and (inverse) scale ζ, generalized Pareto shape ξ and scale ν, and threshold ψ are all functions of direction. In this work, ψ is estimated using nonstationary quantile regression prior to inference for gamma and generalized Pareto parameters, although in general ψ can be estimated as part of a single whole-sample inference, as in the work of Randell, Turnbull, Ewans, and Jonathan (2016). Furthermore,

$$f_{TG}(y|\gamma,\zeta,\psi) = \frac{f_G(y|\gamma,\zeta)}{F_G(\psi|\gamma,\zeta)} \text{ for } y \in [0,\psi],$$

with gamma density

$$f_G(y|\gamma,\zeta) = \frac{\zeta^\gamma}{\Gamma(\gamma)} y^{\gamma-1} \exp(-\zeta x)$$

and generalized Pareto density

$$f_{GP}(y|\xi,\nu,\psi) = \frac{1}{\nu}\left(1 + \frac{\xi}{\nu}(x-\psi)\right)^{-1/\xi-1}.$$

Using F_Y to represent the estimated cumulative distribution function corresponding to density f_Y, we transform the sample to a standard Laplace "X" scale using

$$X = \begin{cases} \log[2F_Y(Y)], & \text{if } F_Y(Y) < 1/2 \\ -\log[2\{1-F_Y(Y)\}], & \text{otherwise.} \end{cases}$$

On the Laplace scale, we identify storm intervals (and hence storm trajectories) as corresponding to contiguous sequences of values of $X > \eta$ for some constant threshold η. Typical examples of the trajectories of the form $\{X_t, \Theta_t\}_{t\in I}$ isolated in this way are illustrated on the physical scale (as $\{Y_t, \Theta_t\}_{t\in I}$) in Figure 1b and Figure 2, for η corresponding to the quantile of the standard Laplace distribution with nonexceedance probability 0.75. Note that data for storm trajectories (with $X > \eta$) only are used in all subsequent analyses.

Storm events are assumed independent for sufficiently large η. In practice, there is some dependence between the value of threshold η and the validity of the independence assumption. For any actual location and time, the extent of between-storm dependence for a given threshold is further conditional on the atmospheric conditions generating the storm events. It is of course important to examine the sensitivity of inferences to the choice of threshold.

In order to simulate storm events, we would typically model the rate of occurrence of storm peak X_0 as a function of storm peak direction Θ_0 using a nonstationary Poisson model (following section 3.2 of Randell et al., 2016, itself motivated by Chavez-Demoulin & Davison, 2005). However, in the current work, for simplicity, because our focus is within-storm evolution, we simply sample Θ_0 from the set of historical storm peak directions.

3.2 | MEM for H_S

Winter (2015) and Winter and Tawn (2017) describe a general kth-order Markov extremal model `MEM(k)` for the evolution of time series $\{X_t\}$ of extreme events on the standard Laplace scale exceeding threshold η. Here, we apply the model in turn to the "prepeak" and "postpeak" intervals of storm trajectories to describe respectively the evolution of a storm to and from its peak in time.

In the current work, as will be justified in due course in Section 4, a second-order MEM (`MEM(2)`) is appropriate. We therefore present this specific model here, as it is simpler to explain than a general kth-order model, yet captures all the important features of `MEM(k)`. We further restrict the presentation of the model to the "postpeak" interval, noting that the corresponding analysis has also been performed for the "prepeak" interval. Finally, we note that `MEM` has been implemented by us as described in the works of Winter (2015) and Winter and Tawn (2017). We therefore give an outline of the model and its estimation here, referring interested readers to the said articles for details.

For the "postpeak" portion $\{X_t\}_{t\geq 0}$ of the time series following the storm peak, we fit a model of the form

$$[X_{t+1}, X_{t+2}] = [\alpha_1, \alpha_2] X_t + X_t^{[\beta_1,\beta_2]} [\mu_1 + \sigma_1 Z_1, \mu_2 + \sigma_2 Z_2] \text{ for } X_t > \eta$$

with parameters $\alpha_1, \alpha_2 \in [-1, 1]$ and $\beta_1, \beta_2 \in (-\infty, 1]$, where $[Z_1, Z_2]$ is a dependent random variable, independent of X_t, with unknown distribution function $G_{1:2}$, where element-wise multiplication is assumed. Threshold η and parameters $\{\alpha_j\}$, $\{\beta_j\}$, $\{\mu_j\}$ and $\{\sigma_j\}$ are taken to be constant. We need to estimate parameters $\{\alpha_j\}$, $\{\beta_j\}$, $\{\mu_j\}$, and $\{\sigma_j\}$, and the distribution $G_{1:2}$.

Following the approach of Heffernan and Tawn (2004), parameters $\{\alpha_j\}$, $\{\beta_j\}$, $\{\mu_j\}$, and $\{\sigma_j\}$ are jointly estimated by assuming *temporarily* that Z_1 and Z_2 are independently distributed with standard Gaussian distributions, and

consequently that

$$\left(X_{t+j}|X_t = x\right) \sim N\left(\alpha_j x + \mu_j x^{\beta_j}, \sigma_j^2 x^{2\beta_j}\right), \quad x > \eta, \quad j = 1, 2.$$

As explained in detail in the Appendix, we then use the fitted parameters to calculate the set of residuals from the fit. This set of residuals is then used in kernel density estimation to estimate the joint distribution $G_{1:2}$ of $[Z_1, Z_2]$ (and the temporary assumption that this distribution is Gaussian is relaxed). As shown in the work of Winter and Tawn (2017), to be able to simulate under the fitted model, we also require the conditional distribution $G_{2|1}$, which is also estimated from the same kernel density estimation.

Now that parameters $\{\alpha_j\}$, $\{\beta_j\}$, $\{\mu_j\}$, and $\{\sigma_j\}$, and the conditional distribution $G_{2|1}$ have been estimated, the basic simulation for $\{X_t\}_{t\geq 0}$ relative to the storm peak event X_0 proceeds as follows. First, a value for X_0 is sampled from its marginal distribution. Then, a realization of the time series of X_t is simulated using

$$X_1 = \alpha_1 X_0 + X_0^{\beta_1}\left(\mu_1 + \sigma_1 Z_1\right), \text{ for } X_0, X_1 > \eta, \text{and}$$
$$X_t = \alpha_2 X_{t-2} + X_{t-2}^{\beta_2}\left(\mu_2 + \sigma_2 Z_{2|1}\right), \text{ for } X_t > \eta, t = 2, 3, \ldots.$$

The simulation continues iteratively until a value of $X_t \leq \eta$ is generated, at which time the simulation is terminated. Moreover, if a value of $X_t > X_0$ is generated during the simulation, the corresponding trajectory is also rejected because the value of the trajectory peak is no longer the desired value X_0.

A similar procedure is used to estimate the corresponding model for the "prepeak" portion $\{X_t\}_{t\leq 0}$, and simulate under it. Diagnostics for the choice of `MEM` order are discussed further in Section 4. We note that Keef, Papastathopoulos, and Tawn (2013) propose extra constraints on each (α, β) pair, which, although not used in this work, could be easily introduced. Further, physical intuition suggests that values of α should be positive.

3.3 | Directional model

Here, we seek a time-series model for the evolution of storm direction Θ_t at time t (relative to storm peak at $t = 0$). We find that this is best achieved in terms of the rate of change of storm direction on a transformed marginal scale. Given the sample $\{\theta_{tk}\}$ of directions, we calculate the corresponding sample $\{\dot{\theta}_{tk}\}$ of rate of change of direction $\dot{\Theta}_t$ by simple differencing. We then transform this sample marginally using its empirical cumulative distribution function and the probability integral transform to the corresponding Gaussian scale sample $\{\delta_{tk}\}$ of transformed rate random variable Δ_t such that

$$\delta_{tk} = \Phi^{-1}\left(\tilde{F}_{\dot{\Theta}}(\dot{\theta}_{tk})\right),$$

where $\tilde{F}_{\dot{\Theta}}$ is the empirical marginal distribution function of rate of change of storm direction estimated from sorting the full sample $\{\dot{\theta}_{tk}\}$ and Φ is the cumulative distribution function of the standard Gaussian distribution. Exploratory analysis of the sample $\{\delta_{tk}\}$ in Section 4 suggests that Δ_t can be modeled using an order k autoregressive model `AR`(k) (with $k = 1, 2, \ldots$) according to

$$(\Delta_t|X_t = x) \sim N\left(\sum_{j=1}^{k} \phi_j \Delta_{t-j}, \varsigma^2(x)\right)$$

with autoregressive parameters $\{\phi_j\}$ and variance $\varsigma^2(x)$, where

$$\varsigma^2(x) = \lambda_1 \exp(-\lambda_2 x) + \lambda_3$$

with $\lambda_1, \lambda_2, \lambda_3 > 0$. For $k = 1$, the parameter ϕ_1 of the `AR`(1) model is constrained by stationarity conditions to the interval $(-1, 1)$. For $k > 1$, stationarity requires that the roots of the corresponding lag polynomial for the `AR`(k) process lie outside the unit circle, thereby constraining the values of the parameter set $\{\phi_j\}$.

The choice of functional form for ς^2 is motivated by the observation that there is little change in storm direction when the value of storm severity is high. A number of similar functional forms for ς^2 were examined in the application described in Section 4, with the form above providing good performance. Diagnostics for the choice of autoregressive model order are discussed further in Section 4.

3.4 | Simulation under the model

To generate a single "postpeak" storm trajectory using the fitted model, the simulation procedure is shown below. An analogous approach is used to simulate the "prepeak" trajectory.

Algorithm 1 Simulation scheme for "post-peak" storm trajectory

simulate storm peak X_0, Θ_0;
for $t = 1, 2, \ldots$ **do**
  simulate X_t using `MEM`;
  if $X_t \geq X_0$ **then**
    reject trajectory and restart;
  end
  if $X_t < \eta$ **then**
    stop and save trajectory;
  end
  simulate Θ_t using directional model;
  simulate Y_t on physical scale using marginal transformation;
end

In the algorithm, in general, Θ_0 would be sampled from a Poisson model for the rate of occurrence of storm peak events noted in Section 3.1; in the current work, we simply sample a value of Θ_0 from the set of historical storm peak directions. The corresponding value of Δ_0 is then sampled at random from the zero-mean Gaussian distribution with appropriate variance, following Section 3.3.

4 | APPLICATION

4.1 | Marginal modeling and transformation to the standard marginal scale

Using Bayesian inference, we estimate the nonstationary gamma-generalized Pareto marginal model (Section 3.1) for the full sample of H_S, using a directional quantile threshold ψ corresponding to nonexceedance probability $\tau = 0.95$ illustrated effectively in Figure 1a. Posterior median parameter estimates are illustrated in Figure 3 as thick solid lines.

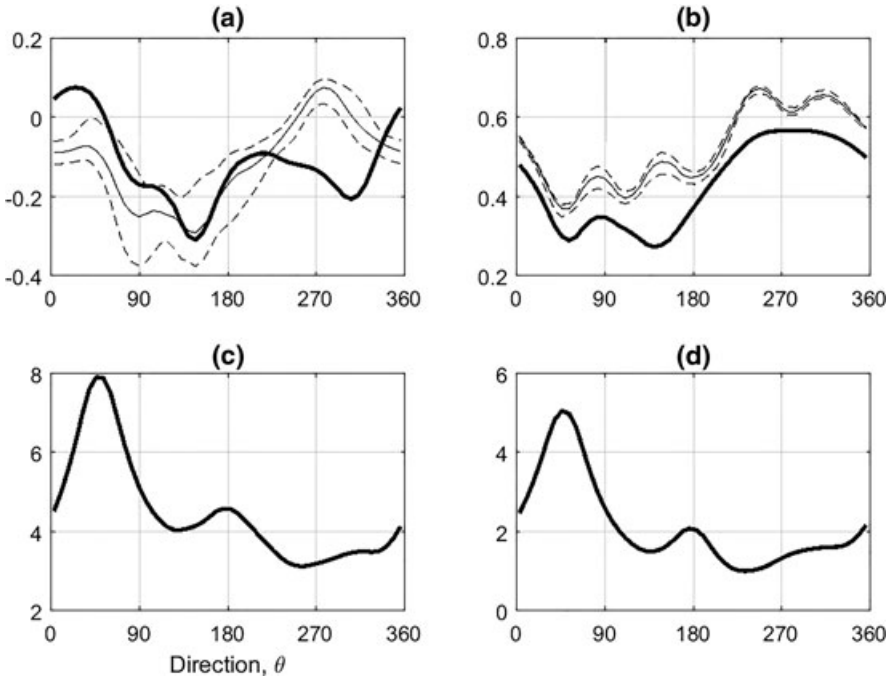

FIGURE 3 Parameter estimates (a) ξ, (b) ν, (c) γ, and (d) ζ for the directional marginal gamma-generalized Pareto model for sea state significant wave height H_S and storm peak significant wave height H_S^{sp}. Thick solid lines are median parameter estimates for H_S. Thin solid lines are median estimates for H_S^{sp}. Thin dashed lines are 95% uncertainty bands for H_S^{sp}. The directional threshold ψ used is illustrated in Figure 1

Figure 3 also shows parameter median estimates and 95% credible intervals for a generalized Pareto model of H_S^{sp} events exceeding the same directional quantile threshold ψ. We estimate the extreme value model for H_S assuming observations to be independent. In reality, because consecutive values of H_S are actually dependent, the credible intervals estimated from the H_S analysis are biased low. We therefore do not show credible intervals for parameters of the H_S model in the Figure. Nevertheless, the credible intervals for the H_S^{sp} model parameters give an indication as to the likely uncertainties in the H_S model parameters because storm peak events are assumed independent.

It is interesting to note in Figure 3 that the directional characteristics of storm peak events (H_S^{sp}) and those of sea state H_S are similar. From a statistical perspective, this might be expected because the generalized Pareto distribution with the same parameters applies equally to cluster maxima and arbitrary exceedances.

Then, we transform the full H_S sample from physical "Y" scale to the standard Laplace marginal "X" scale using the median parameter estimates from the H_S model and identify contiguous intervals above a second (constant) threshold η corresponding to a nonexceedance probability of 0.75, thereby isolating a sample of storm trajectories. We examined the sensitivity of subsequent inferences to the choice of η and found these to be relatively stable.

4.2 | MEM for H_S

For brevity in Sections 4.2 and 4.3, we discuss and illustrate model estimates for the "prepeak" portion of the storm trajectory only. Note that the corresponding analysis for the "postpeak" portion was also preformed, and exhibited similar characteristics to those discussed below. Both "prepeak" and "postpeak" models are used in the simulation of storm trajectories in Section 4.4.

We consider a number of diagnostics for the Laplace-scale data to judge an appropriate order k for the `MEM(k)` model to estimate. Figure 4a shows the partial autocorrelation function for X_t, indicating strong effects at lag 1 and 2, and potentially at larger lags also. Figure 4b shows estimates for the extremal dependence summary statistic χ (e.g., Eastoe, Koukoulas, & Jonathan, 2013) defined for time lag τ and threshold ω by

$$\chi(\tau, \omega) = \Pr\left(X_{t+\tau} > \omega | X_t > \omega\right).$$

The empirical estimate from the sample is shown in black. Other estimates are obtained from simulations under `MEM` models of different orders fitted to the sample. Inspection of Figure 4b suggests that `MEM(2)` provides a better description than `MEM(1)` and that higher order models provide no additional benefit over `MEM(2)`.

Figure 5 shows scatter plots for pairs of parameter estimates for the `MEM(2)` model from a bootstrap resampling study.

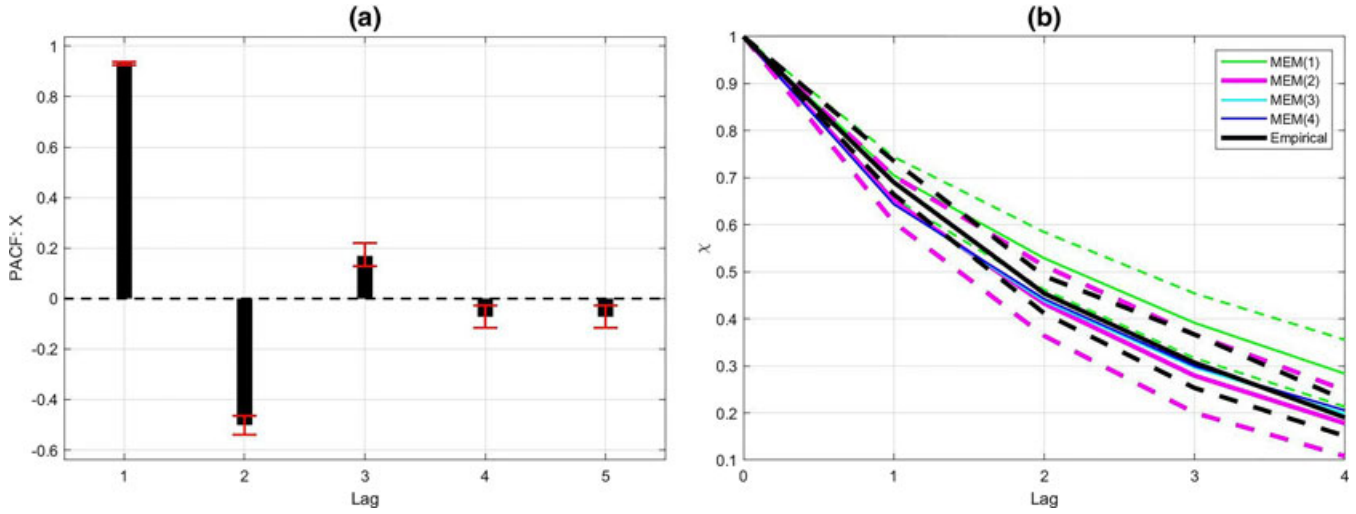

FIGURE 4 Choice of model order for the Markov extremal model (MEM) of Laplace-scale H_S, X. (a) Sample partial autocorrelation function with bootstrap 95% uncertainty bands in red. (b) Tail summary statistic χ corresponding to a nonexceedance probability of 0.995 for the sample and for realizations under the models of different orders. 95% bootstrap uncertainty bands are shown for the original sample (black), `MEM(1)` (green), and `MEM(2)` (pink) only. Uncertainty bands for `MEM(3)` and `MEM(4)` estimates have similar widths to those of `MEM(2)` and are omitted for clarity. PACF = partial autocorrelation function

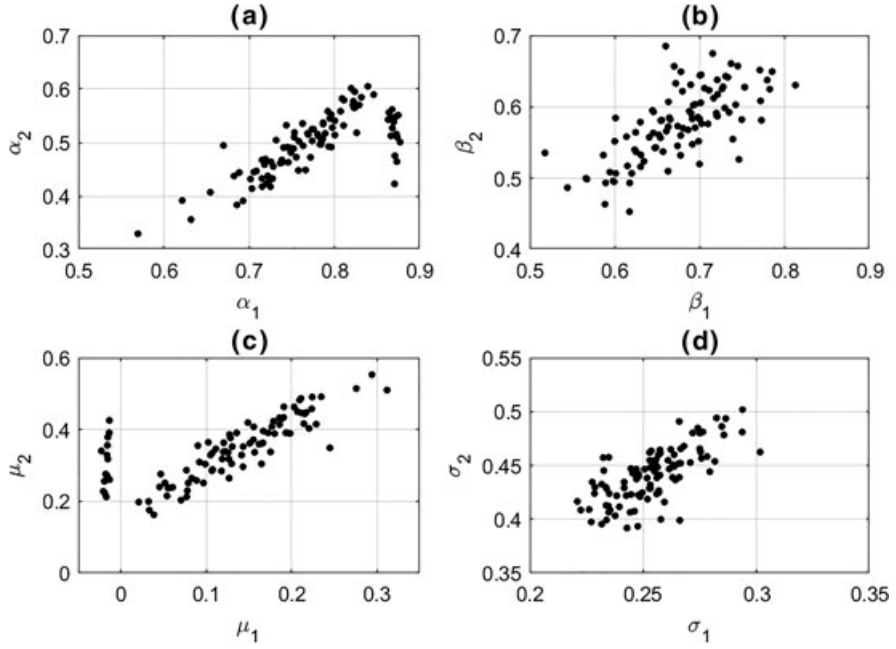

FIGURE 5 Scatter plots of α_2 on α_1, β_2 on β_1, μ_2 on μ_1, and σ_2^2 on σ_1^2 for the `MEM(2)` prepeak model for H_S from a bootstrap resampling analysis, using a threshold with nonexceedance probability 0.75. MEM = Markov extremal model

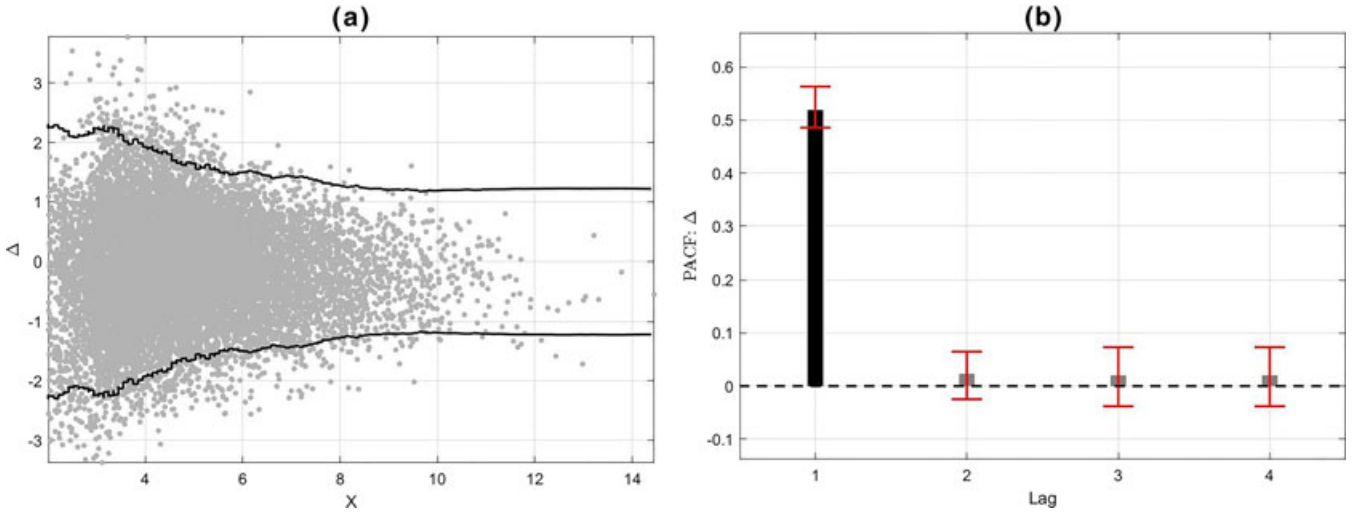

FIGURE 6 Directional diagnostics. (a) Scatter plot (gray dots) of Δ_t on X_t from the sample. Empirical estimates for 2.5% and 97.5% percentiles of distribution of $\Delta_t|X_t$ (estimated assuming normality; black lines). (b) Sample partial autocorrelation function for Δ_t with bootstrap 95% uncertainty bands in red. PACF = partial autocorrelation function

4.3 | Directional model for θ

Figure 6a shows the full sample of values for Δ_t on X_t, suggesting that the variance of Δ_t reduces with increasing X_t, and motivating the choice of directional autoregressive evolution variance in Section 3.3. Figure 6b shows the partial autocorrelation function for Δ_t, indicating that only lag 1 is significant. We therefore proceed to fit an `AR(1)` model for Δ_t with nonstationary evolution variance $\varsigma^2(X_t)$. Note that data for storm trajectories (with $X > \eta$) only were used to estimate the partial autocorrelation function for Δ_t.

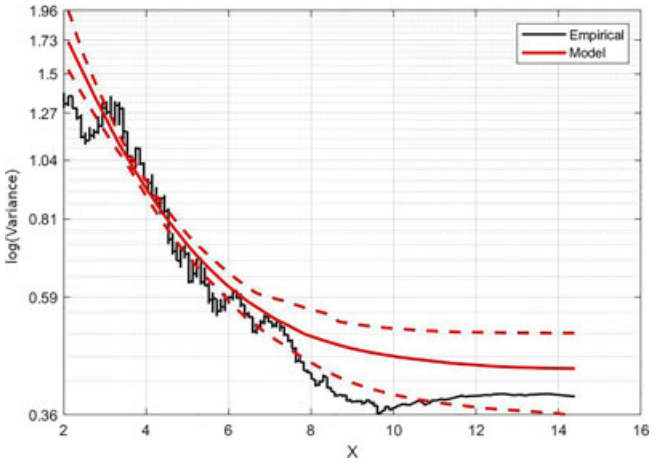

FIGURE 7 Directional model. Empirical (black) and model (red) estimates (with 95% bootstrap uncertainty band) for the decay of $\varsigma^2(X_t)$ with X_t

The estimated form of $\varsigma^2(X_t)$ with X_t is illustrated in Figure 7 with a 95% bootstrap uncertainty band, together with an empirical estimate direct from the sample.

4.4 | Simulation under the model and model validation

We use the simulation procedure from Section 3.4 to generate storm trajectories under the estimated nonstationary `MEM(2)` model. As indicated in Algorithm 1, the first step in the simulation would typically be to sample a realization storm peak significant wave height and direction. This requires estimation of a model for the rate of occurrence of storm peak events in a given direction and estimation of a model for the size of storm peak significant wave height given direction (the latter similar to that illustrated in Figure 3). In the current work, because focus is on estimation of storm evolution given storm peak, rather than estimating the storm peak itself, we choose for clarity and simplicity to sample storm peak events with replacement from the historical sample of storm peaks rather than from a statistical model for storm peak characteristics.

Figure 4b and Figure 5 illustrate that the extremal dependence χ of simulated trajectories with lag agrees well with that of the original sample and that nonstationary characteristics of the evolutionary variance of Δ_t are also in good agreement. Figure 8 shows 15 typical realizations in the same format as in Figure 2; the characteristics of the trajectories in the two figures appear similar.

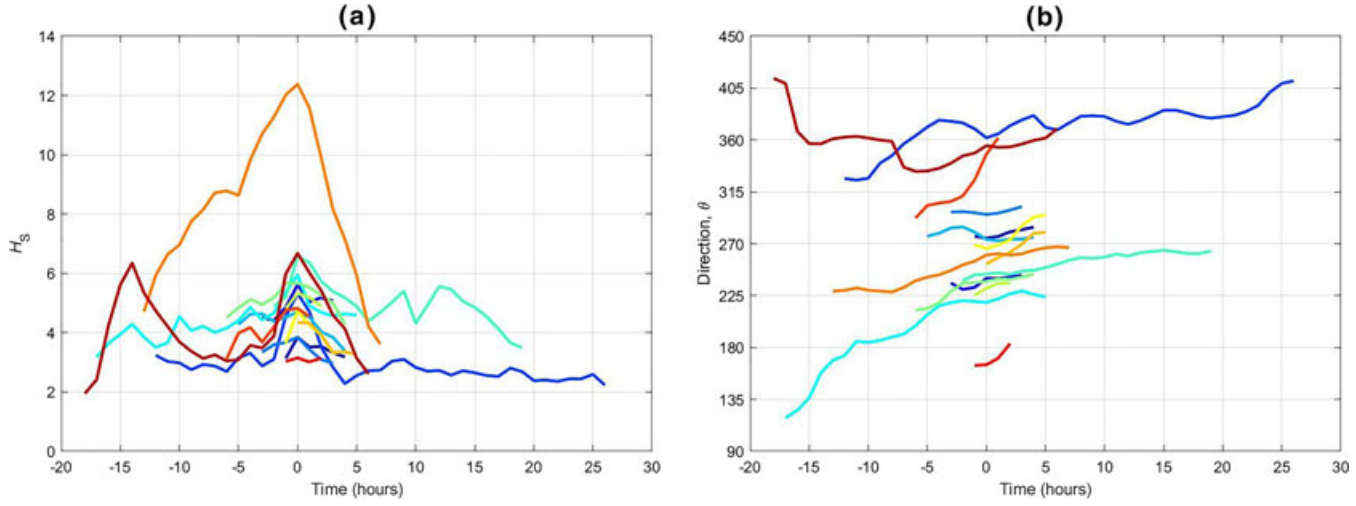

FIGURE 8 Time evolution of 15 typical simulated realizations of storm trajectories. (a) H_S in time. (b) θ in time

Furthermore, we choose to examine whether the distribution of lengths of storm trajectories generated under the model are consistent with that of the original sample; this is of particular interest because we do not explicitly seek to estimate storm length in the model from Section 3. Figure 9 illustrates this comparison for both `MEM(1)` and `MEM(2)` models. Agreement for the `MEM(2)` model is considerably better as might be expected.

Figure 10 compares the full distribution of sea state H_S from the original sample with that estimated from simulation under the `MEM(2)` model, in terms of the probability density function and the tail distribution. There is some suggestion that the mode of the body of simulated H_S is to the left to that of the original sample, but agreement is generally again good.

Figure 11 is a directionally resolved version of Figure 9b, comparing the distributions of storm lengths per directional octant. Figure 12 is a directionally resolved version of Figure 10, comparing the distributions of H_S per directional octant. As in Figures 9 and 10, agreement between simulations under the model and the original sample is good.

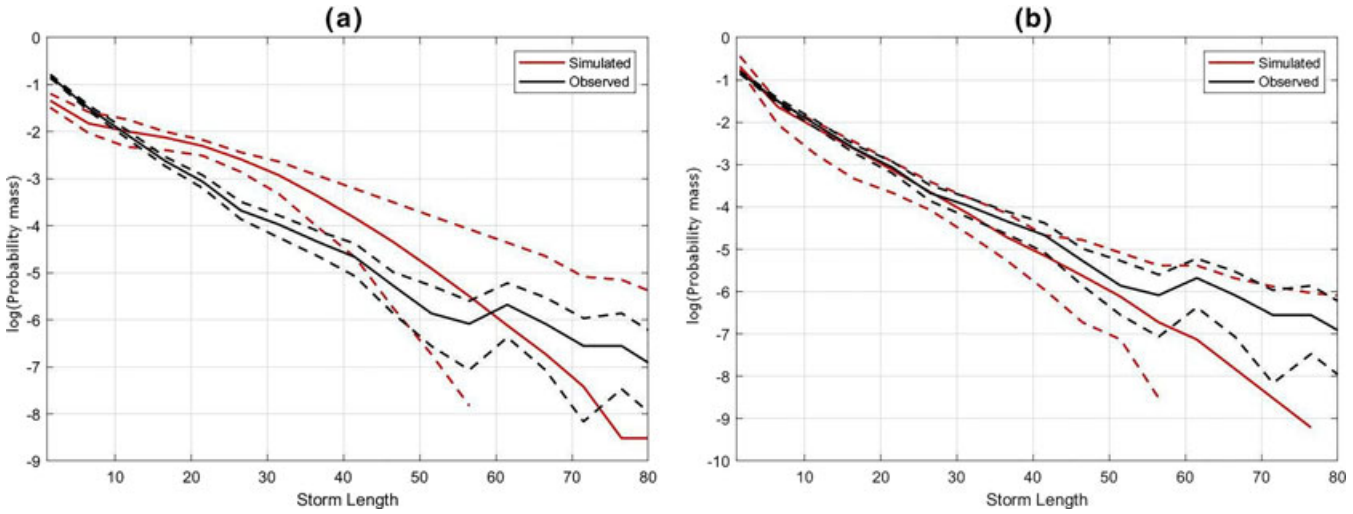

FIGURE 9 Logarithm of the probability mass for the number of sea states in a storm with the first-order directional model and (a) `MEM(1)` and (b) `MEM(2)`. Black lines give empirical sample estimates with a 95% bootstrap uncertainty band. Red lines are estimated by simulation under the model

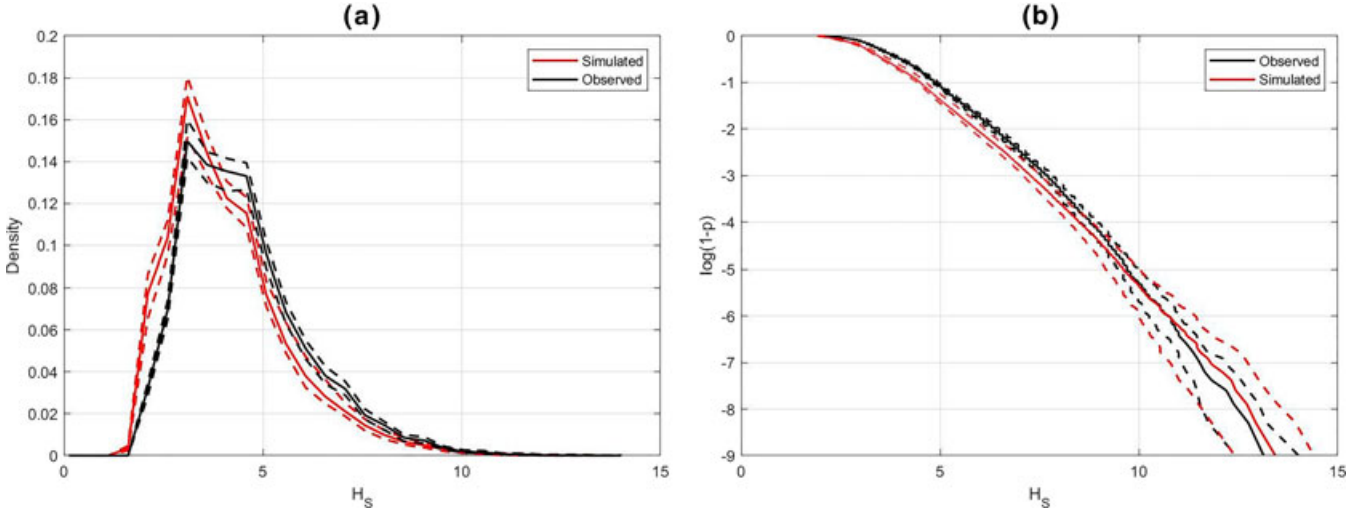

FIGURE 10 Comparison of distribution of H_S from the sample (black) and simulation under the fitted `MEM(2)` model (red) with 95% bootstrap uncertainty bands. (a) Density. (b) Tail of distribution

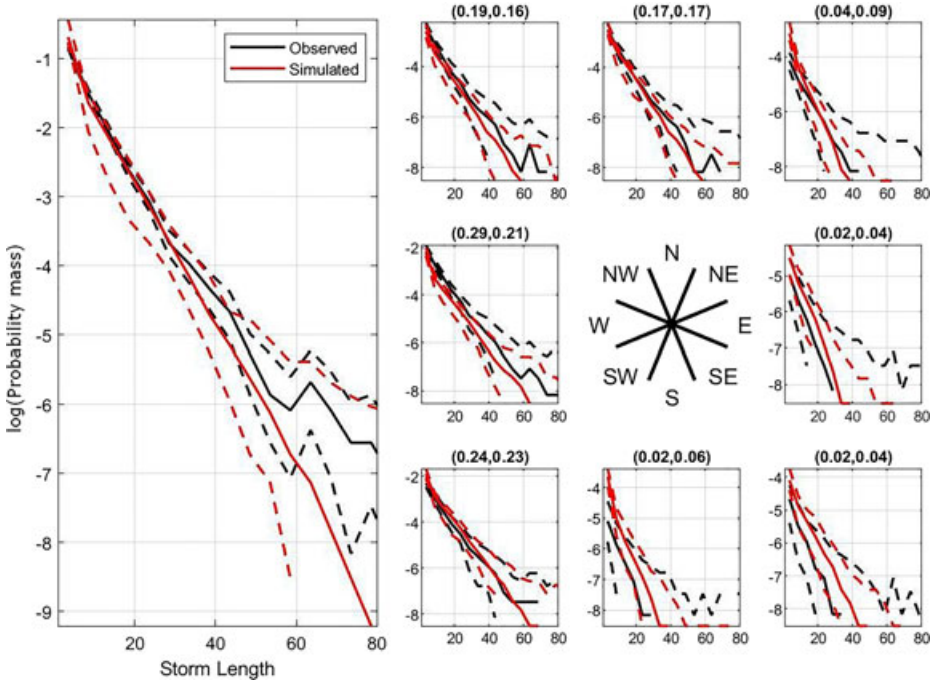

FIGURE 11 Directional comparison of logarithm of probability mass for storm lengths. The left-hand panel shows the omni-directional comparison, and the smaller plots show comparisons for eight directional octants centered on cardinal and intercardinal directions. Each panel shows the original sample tail (black) and the simulated tail (red) with 95% bootstrap uncertainty bands. Titles of smaller panels give the fraction of storm peak occurrences in the directional octant, first from the original sample and, then, from simulation

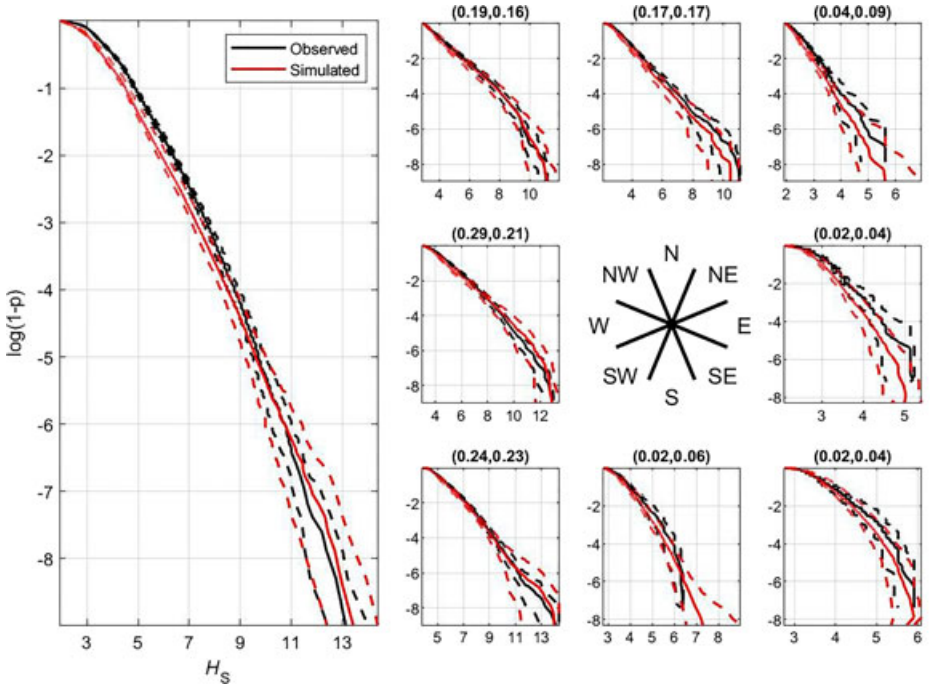

FIGURE 12 Directional comparison of tails of distributions of H_S. The left-hand panel shows omni-directional comparison, and the smaller plots show comparisons for eight directional sectors centered on cardinal and intercardinal directions. Each panel shows the original sample tail (black) and the simulated tail (red) with 95% bootstrap uncertainty bands. Titles of smaller panels give the fraction of storm peak occurrences in the directional octant, first from the original sample and, then, from simulation

5 | DISCUSSION AND CONCLUSIONS

We develop a nonstationary MEM as an extension of Winter and Tawn (2016, 2017) and use it to characterize the time evolution of extreme sea state significant wave height (H_S) in the vicinity of the storm peak sea state. The approach first requires transformation of H_S from physical to standard Laplace scale, achieved using a nonstationary directional marginal extreme value model. The evolution of Laplace-scale H_S is subsequently characterized using a second-order MEM, and the evolution of rate of change of storm direction, transformed to Gaussian scale, is described by a first-order autoregressive model, where the evolution variance of which is H_S-dependent. Simulations on the physical scale under the estimated model give realistic realizations of storm trajectories consistent with historical data, in terms of the distributions of trajectory lengths and total marginal distribution of H_S, at a northern North Sea location.

We consider the original MEM to be an important contribution to the environmental statistics and ocean engineering literature because simple descriptions of the evolution of time series of extreme values of variables from the physical environment are needed in many contexts. Given that covariate effects are present in the majority of ocean engineering settings, we consider the current work to be a useful extension of MEM, allowing more realistic application in such settings. We note as obvious examples the estimation of surge trajectories (e.g., Ross, Sam, Randell, Feld, & Jonathan, 2018) for ocean storms and the characterization of ocean current soliton events. The main motivation for the current work is the development of a statistical model to describe evolution of storm trajectories so that, given storm peak characteristics, realistic statistical simulators of storm evolution (as opposed to "storm matching" to events in a library of historical storms used in, for example, the work of Feld, Randell, Wu, Ewans, & Jonathan, 2015) can be used for met-ocean design, thereby facilitating a more formal model-based uncertainty quantification.

There are many ways in which the current nonstationary model can be extended. In the current work, we have assumed, given nonstationary marginal transformation to Laplace scale, that MEM parameters (on Laplace scale) are stationary with respect to covariates; we judge this to be a reasonable starting assumption supported by model diagnostics in the current application. However, in general, we suspect that this may not always be the case and that MEM parameters themselves may show covariate dependence. Keef et al. (2013) propose additional constraints on the conditional extremes parameters, which could be incorporated.

It would be interesting to consider the joint evolution of time series of variables such as sea state H_S and spectral peak period T_P during a storm, nonstationary with respect to a common set of covariates, or joint modeling of time series such as sea state H_S, wind speed, and current speed with a much larger set of potential covariates. We might also consider the temporal evolution of storm trajectories jointly for multiple locations in a neighborhood. The between-variable, between-time dependence structure of such a model would be interesting and probably challenging to identify.

ACKNOWLEDGEMENT

We acknowledge the useful conversations with Graham Feld and Matthew Jones at Shell, and Jonathan Tawn at Lancaster University.

ORCID

E. Ross http://orcid.org/0000-0002-0287-0611
D. Randell http://orcid.org/0000-0003-1127-5491
P. Jonathan http://orcid.org/0000-0001-7651-9181

REFERENCES

Chavez-Demoulin, V., & Davison, A. (2005). Generalized additive modelling of sample extremes. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, *54*, 207–222.

Chavez-Demoulin, V., & Davison, A. (2012). Modelling time series extremes. *REVSTAT*, *10*, 109–133.

Davison, A. C., Padoan, S. A., & Ribatet, M. (2012). Statistical modelling of spatial extremes. *Statistical Science*, *27*, 161–186.

Eastoe, E., Koukoulas, S., & Jonathan, P. (2013). Statistical measures of extremal dependence illustrated using measured sea surface elevations from a neighbourhood of coastal locations. *Ocean Engineering*, *62*, 68–77.

Feld, G., Randell, D., Ross, E., & Jonathan, P. (2018). Design conditions for waves and water levels using extreme value analysis with covariates. Submitted to Ocean Engineering, draft at http://www.lancs.ac.uk/~jonathan/

Feld, G., Randell, D., Wu, Y., Ewans, K., & Jonathan, P. (2015). Estimation of storm peak and intra-storm directional-seasonal design conditions in the North Sea. *Journal of Offshore Mechanics and Arctic Engineering*, *137*, 021102:1–021102:15.

Heffernan, J. E., & Tawn, J. A. (2004). A conditional approach for multivariate extreme values (with discussion). *Journal of the Royal Statistical Society: Series B*, *66*, 497–546.

Huser, R., & Davison, A. C. (2014). Space–time modelling of extreme events. *Journal of the Royal Statistical Society: Series B*, *76*, 439–461.

Keef, C., Papastathopoulos, I., & Tawn, J. A. (2013). Estimation of the conditional distribution of a vector variable given that one of its components is large: Additional constraints for the Heffernan and Tawn model. *Journal of Multivariate Analysis*, *115*, 396–404.

Padoan, S. A., Ribatet, M., & Sisson, S. A. (2010). Likelihood-based inference for max-stable processes. *Journal of the American Statistical Association*, *105*, 263–277.

Randell, D., Turnbull, K., Ewans, K., & Jonathan, P. (2016). Bayesian inference for non-stationary marginal extremes. *Environmetrics*, *27*, 439–450.

Reich, B. J., & Shaby, B. A. (2012). A hierarchical max-stable spatial model for extreme precipitation. *The Annals of Applied Statistics*, *6*, 1430–1451.

Reistad, M., Breivik, Ø., Haakenstad, H., Aarnes, O. J., Furevik, B. R., & Bidlot, J.-R. (2011). A high-resolution hindcast of wind and waves for the North Sea, the Norwegian Sea, and the Barents Sea. *Journal of Geophysical Research*, *116*, 1–18.

Rootzén, H., Segers, J., & Wadsworth, J. L. (2018). Multivariate peaks over thresholds models. *Extremes*, *21*, 115–145.

Ross, E., Randell, D., Ewans, K., Feld, G., & Jonathan, P. (2017). Efficient estimation of return value distributions from non-stationary marginal extreme value models using Bayesian inference. *Ocean Engineering*, *142*, 315–328.

Ross, E., Sam, S., Randell, D., Feld, G., & Jonathan, P. (2018). Estimating surge in extreme North Sea storms. *Ocean Engineering*, *154*, 430–444.

Stephenson, A. G. (2009). High-dimensional parametric modelling of multivariate extreme events. *Australian & New Zealand Journal of Statistics*, *51*, 77–88.

Wadsworth, J. L., Tawn, J. A., Davison, A. C., & Elton, D. M. (2017). Modelling across extremal dependence classes. *Journal of the Royal Statistical Society: Series B*, *79*, 149–175.

Winter, H. C. (2015). *Extreme value modelling of heatwaves* (PhD thesis). Lancaster University, Lancaster, UK

Winter, H. C., & Tawn, J. A. (2016). Modelling heatwaves in central France: A case-study in extremal dependence. *Journal of the Royal Statistical Society: Series C*, *65*, 345–365.

Winter, H. C., & Tawn, J. A. (2017). *k*th-order Markov extremal models for assessing heatwave risks. *Extremes*, *20*, 393–415.

How to cite this article: Tendijck S, Ross E, Randell D, Jonathan P. A model for the directional evolution of severe ocean storms. *Environmetrics*. 2018;e2541. https://doi.org/10.1002/env.2541

APPENDIX

ESTIMATION OF DISTRIBUTION OF MEM RESIDUALS

Following Section 3, for the "postpeak" portion $\{X_t\}_{t\geq 0}$ of the time series following the storm peak, the second-order Markov extremal model form is

$$[X_{t+1}, X_{t+2}] = [\alpha_1, \alpha_2] X_t + X_t^{[\beta_1,\beta_2]} [\mu_1 + \sigma_1 Z_1, \mu_2 + \sigma_2 Z_2] \text{ for } X_t > \eta,$$

where $[Z_1, Z_2]$ is a dependent random variable, independent of X_t, with unknown distribution function $G_{1:2}$, where element-wise multiplication is assumed. We write the full set of residuals available from the regression fit as $\{\hat{z}_{i1}, \hat{z}_{i2}\}_{i=1}^{n}$, where $\hat{z}_{ij} = (x_{t_i+j} - \hat{\alpha}_j x_{t_i} - \hat{\mu}_j x_{t_i}^{\beta_j})/(\hat{\sigma}_j x_{t_i}^{\beta_j})$, $j = 1, 2$ with $\{t_i\}$ representing the occurrence times of the corresponding observations of X, and $t = 0$ corresponding to the time of the storm peak. Then, following the work of Winter (2015), we define kernel density estimates for the density g_1 of Z_1 and $g_{1,2}$ of $[Z_1, Z_2]$ as

$$g_1(z_1) = \frac{1}{n} \sum_i \frac{1}{h_1} f\left(\frac{z_1 - \hat{z}_{i1}}{h_1}\right)$$

and

$$g_{1,2}(z_1, z_2) = \frac{1}{n} \sum_i \frac{1}{h_1 h_2} f\left(\frac{z_1 - \hat{z}_{i1}}{h_1}\right) f\left(\frac{z_2 - \hat{z}_{i2}}{h_2}\right),$$

where f is the kernel density, taken in this work to be Gaussian, and (h_1, h_2) are kernel widths estimated by inspection of resultant estimates for densities g_1 and $g_{1,2}$. Then, the cumulative distribution function $G_{2|1}$ of the conditional residual $Z_{2|1}$ can be written as

$$G_{2|1}(z_2|z_1) = \sum_i w_i(z_1) F\left(\frac{z_2 - \hat{z}_{i2}}{h_2}\right),$$

where F is the kernel cumulative distribution function corresponding to density f and weights $\{w_i(z_1)\}$ are given by

$$w_i(z_1) = \left(\sum_{i'} f\left(\frac{z_1 - \hat{z}_{i'1}}{h_1}\right)\right)^{-1} f\left(\frac{z_1 - \hat{z}_{i1}}{h_1}\right).$$