Second submission : April 2020

On the distribution of maximum crest and wave height at intermediate water depths

Matthias Schubert[a,*], Yanyun Wu[b], Jesper Tychsen[c], Michael Havbro Faber[a,d], John Dalsgaard Sørensen[d], Philip Jonathan[b,e]

[a] *Matrisk GmbH, Switzerland*
[b] *Shell Research Limited, London SE1 7NA, United Kingdom.*
[c] *TOTAL E&P Danmark A/S, Denmark*
[d] *Aalborg University, Department of Civil Engineering, Denmark*
[e] *Department of Mathematics and Statistics, Lancaster University LA1 4YF, United Kingdom.*

Abstract

We report new descriptions for the (probability) distributions of hourly maximum crest and wave height of water surface gravity waves for intermediate water depths. Estimated distributions are based on analysis of laboratory-scale measurements at the DHI wave basin. For a given sea state, the distribution of both hourly maximum crest and hourly maximum wave height, normalised by sea state significant wave height, is found to follow a generalised extreme value (GEV) distribution. Variation of the three parameters of the GEV distribution across sea states, is expressed in terms of a response surface model as a function of non-dimensional sea state Ursell number and wave steepness, and wave directional spreading angle. For inference, conventional Monte Carlo wave basin measurements are supplemented with measurements selected by means of a novel "pre-selection" sampling scheme using numerical simulations. This scheme effectively guarantees that extreme events from tails of distributions are produced, and reduces uncertainties associated with the estimated distributions. Estimation is performed using Bayesian inference, allowing uncertainties to be quantified, and providing estimates of posterior predictive tail distributions for sea states with arbitrary characteristics within the domain of sea state characteristics covered by the model.

Keywords: distribution, wave height, crest height, hourly maximum, extreme, generalised extreme value, metocean design

1. Introduction

In 2012, a photograph (taken in 2007) was discovered (Tychsen et al. 2016). It shows a plunging, breaking extreme wave, occurring in a sea state with significant wave-height H_{m0} of approximately 10m in the Danish Tyra field, offshore Jutland, with water depth d of approximately 45m. Further investigation revealed the existence of a video showing a second, even larger plunging breaker with crest height between 17m and 17.5m for the same storm and location. The storm was estimated to correspond to an event with return period of approximately 25 years, yet the crest observations exceeded the 10,000-year design criteria for both crest elevation and wave kinematics. This triggered an extensive study to quantify the characteristics of highly non-linear extreme wave events (Tychsen 2016). The current work forms part of that study.

Distributions for individual wave height and crest

Good estimation of extremes of individual wave height H and crest elevation E, and of maximum wave height $H_{\max}$ and crest elevation $E_{\max}$ on some spatio-temporal domain, are essential for the design and assessment of marine structures. There is a large literature on this subject stretching over many decades, addressing amongst other things estimation of (probability) distributions for H, E, $H_{\max}$ and $E_{\max}$ from theory and observation. The distribution of H in deep water has been modelled using the Rayleigh distribution (Longuet-Higgins 1952) and the Weibull distribution (Forristall 1978). Tayfun (1990) generalised the asymptotic model of Boccotti (1989) to include the effects of higher order non-linearities. Depth-induced wave breaking leading to restriction on the ratio of wave height to water depth required new parameterisations of these or other distributional forms for shallow water. To this end, Glukhovskiy (1966) proposed a Weibull parameterisation accommodating depth-limited breaking, modified by van Vledder (1991). Battjes and Groenendijk (2000) suggested a two-part Weibull-Weibull distribution and Wu et al. (2016) a Weibull-generalised Pareto distribution parameterised empirically in terms of sea state variables. Mendez et al. (2004), Forristall (2007), Katsardi and Swan (2011a,b) and Katsardi et al. (2013) reported studies of H in shallow water and also proposed Weibull-type distributions. Forristall (2000) provided a distribution for E motivated by observations and second-order theory, and summarised earlier work on crest height distributions starting with Haring et al. (1976). Katsardi et al. (2013) studied the tail of the distribution of E in shallow water using laboratory measurements. There is a large literature (including Tromans and Vanderschuren 1995, Prevosto et al. 2000 and Krogstad

*Corresponding author schubert@matrisk.com

and Barstow 2004) on the estimation of distributions for $H_{\max}$ and $E_{\max}$ for a defined time interval of observation, for example corresponding to the occurrence of one or more ocean storms, each consisting of multiple sea states with different characteristics; Mackay and Johanning (2018) provides a recent review.

The recent work of Karmpadakis et al. (2019) discusses the statistical distribution of the crest heights associated with surface waves in intermediate water depths, including the effects of varying sea-state steepness, effective water depth and directional spreading are investigated. They observe amplifications of crest heights of 5%-10% above second order for realistic directional spreads, and the increased influence of wave breaking as effective water depth reduces. The inferences of Karmpadakis et al. (2019) concerning the effects of sea-state steepness and directional spreading are in line with those of the current work.

In addition to measurement, and solution of equations of motion for surface gravity waves, asymptotic statistical theory can also be used to motivate useful parametric forms for the distribution of extremes of wave and crest height. Specifically for a random variable belonging to the maximum domain of attraction of a non-degenerate distribution (informally, a "max-stable" random variable), the distribution of peaks over threshold will converge to the generalised Pareto distribution, as threshold level increases. Further, the distribution of maxima (over some interval) of the random variable will converge to the generalised extreme value (GEV) distribution as the number of occurrences in the interval increases (e.g. Beirlant et al. 2004, Jonathan and Ewans 2013). A random variable M following the generalised extreme value distribution has cumulative distribution function F_M and probability density function f_M given by

$$\begin{aligned} F_M(m) &= \exp\{-t(m)\} \\ f_M(m) &= \frac{1}{\sigma}\, t(m)^{\xi+1} \exp\{-t(m)\} \end{aligned}$$

where

$$t(m) = \begin{cases} \left(1 + \frac{\xi}{\sigma}\,(m-\mu)\right)^{-1/\xi} & \text{if } \xi \neq 0 \\ \exp\{-\frac{1}{\sigma}\,(m-\mu)\} & \text{if } \xi = 0 \end{cases}$$

for shape parameter $\xi \in \mathbb{R}$, scale $\sigma \in \mathbb{R}$ and $\sigma > 0$, and location $\mu \in \mathbb{R}$. When $\xi < 0$, the distribution F_M has a finite upper end point $\mu - \sigma/\xi$.

Objective

The objective of the current work is to estimate statistical models for the distributions of hourly maximum crest height $E_{\max}$ and hourly maximum wave height $H_{\max}$, across a variety of different sea states. Models are estimated using observations (at laboratory scale) made at the DHI (formerly Danish Hydraulics Institute) wave basin at an intermediate water depth, incorporating highly non-linear effects such as wave breaking. The estimated statistical models are intended to be used to predict the distributions of $E_{\max}$ and $H_{\max}$ for other sea states with similar characteristics at similar intermediate water depths. The models therefore quantify the "short-term" distributions of $E_{\max}$ and $H_{\max}$, which can then be combined with long-term distributions of sea state variables to estimate "long-term" distributions of $E_{\max}$ and $H_{\max}$ for periods of up to 10^6 years (e.g. Hansen et al. 2019).

Outline

The article is arranged as follows. Section 2 introduces the data, and Section 3 provides a description of the different model elements considered, and the model form eventually chosen. Section 4 outlines application to the estimation of the distribution of hourly maximum crest height $E_{\max}$, and briefly summarises the corresponding analysis for hourly maximum wave height $H_{\max}$. Section 5 provides discussion and conclusions. Details of the modelling strategy, including a description of the so-called "numerical pre-selection" procedure for sampling from the extreme tail of distributions, Markov chain Monte Carlo sampling for Bayesian inference, response surface regression modelling, the hierarchy of statistical models considered, and MATLAB software for the estimated distributions, are given in appendices.

For brevity, unless stated otherwise, we use the term "distribution" to refer to probability distribution, "density" to refer to probability density function, and "tail" to refer to the upper tail of the probability distribution of a random variable.

2. Data

The data used in this study are measurements of hourly maximum wave and crest heights made at the DHI wave basin, more details of which are given in Bredmose et al. (2016). The set-up of the wave basin for the experiments described here is generally in line with Bredmose et al. (2016); a photograph taken during the experiments is given in Figure 1. Measurements were made at a lattice of $n_L = 100$ locations as illustrated in Figure 2, with dimensions given on actual laboratory (henceforth "lab") and equivalent full ("full") scales. The reference location is indicated. Laboratory water depth of 0.5 m corresponds

to 45m in full scale offshore, requiring a Froude number scaling of 1:90 and corresponding time scaling of $1{:}90^{0.5}$. In general terms, referring to Figure 2, waves propagate from left to right.

No walls where present in basin tests, and sloping beds made of crushed stones where placed around the sides in the basin. Furthermore, steel wave absorbers where installed in an M shape, to direct any reflections away from the test area. The closest wave absorber to any wave gauge was approximately 8m.

Figure 1: The DHI wave tank in operation, with acknowledgement to DHI and Total E&P Danmark A/S.

Just over 11 hours (lab-scale, corresponding to 105 hours at full scale) of continuous measurements were made at all locations for a total of $n_S = 50$ different sea states. From these measurements, $n_H = 105$ observations of hourly (full-scale) maxima $E_{\max}$ and $H_{\max}$, for each of n_L locations, for each of the n_S sea states were identified, with up-crossing definition for wave height. The corresponding sample can be written $\{m_{ijk}\}_{i=1,j=1,k=1}^{n_H,n_L,n_S}$, corresponding to n_H observations of each of the random variables $\{M_{jk}\}_{j=1,k=1}^{n_L,n_S}$, where M represents either $E_{\max}$ or $H_{\max}$. Individual sea states, indexed by k ($k = 1, 2, ..., n_S$), are characterised in terms of significant wave height H_{m0} (in metres), steepness ϵ, Ursell number U and directional spreading angle θ_{sp} (in degrees). The defining expressions for ϵ and U are

$$\epsilon = \frac{H_{m0}}{\lambda_p}, \quad U = \frac{H_{m0}}{k_p^2 d^3}$$

where λ_p is the wavelength at spectral peak frequency satisfying $\lambda_p = 2\pi/k_p$, $\omega_p^2 = g k_p \tanh(k_p d)$, $\omega_p = 2\pi/T_p$ and T_p is the spectral peak period. The set of n_S sea states was selected to cover a range of different values of H_{m0}, ϵ, U and θ_{sp} and combinations thereof appropriate for the location of interest, as illustrated in Figure 3. A listing of values of sea state characteristics is given in full in Table A.1 of Appendix A.

Figure 4 shows the mean quantile $\bar{q}_k(p)$ (per probability level p) of the distribution of $E_{\max}$ per sea state k, estimated using

$$\bar{q}_k(p) = \frac{1}{n_L} \sum_{j=1}^{n_L} q_{jk}(p) \text{ for } p \in (0, 1)$$

where quantile $q_{jk}(p)$ is estimated empirically by solving $\Pr(M_{jk} \leq q_{jk}(p_i)) = p_i$ for $p_i = (i - 0.5)/n_H$, $i = 1, 2, ..., n_H$ using the sample; each tail is constructed using n_H hourly maxima measurements. There is considerable variation in the characteristics of the tails between sea states which we seek to explain using statistical modelling, in two steps. First, extreme value analysis is used to characterise the tail for each sea state independently. Then, a regression model is established relating the parameters of extreme value models across sea states to the summary statistics of the sea state.

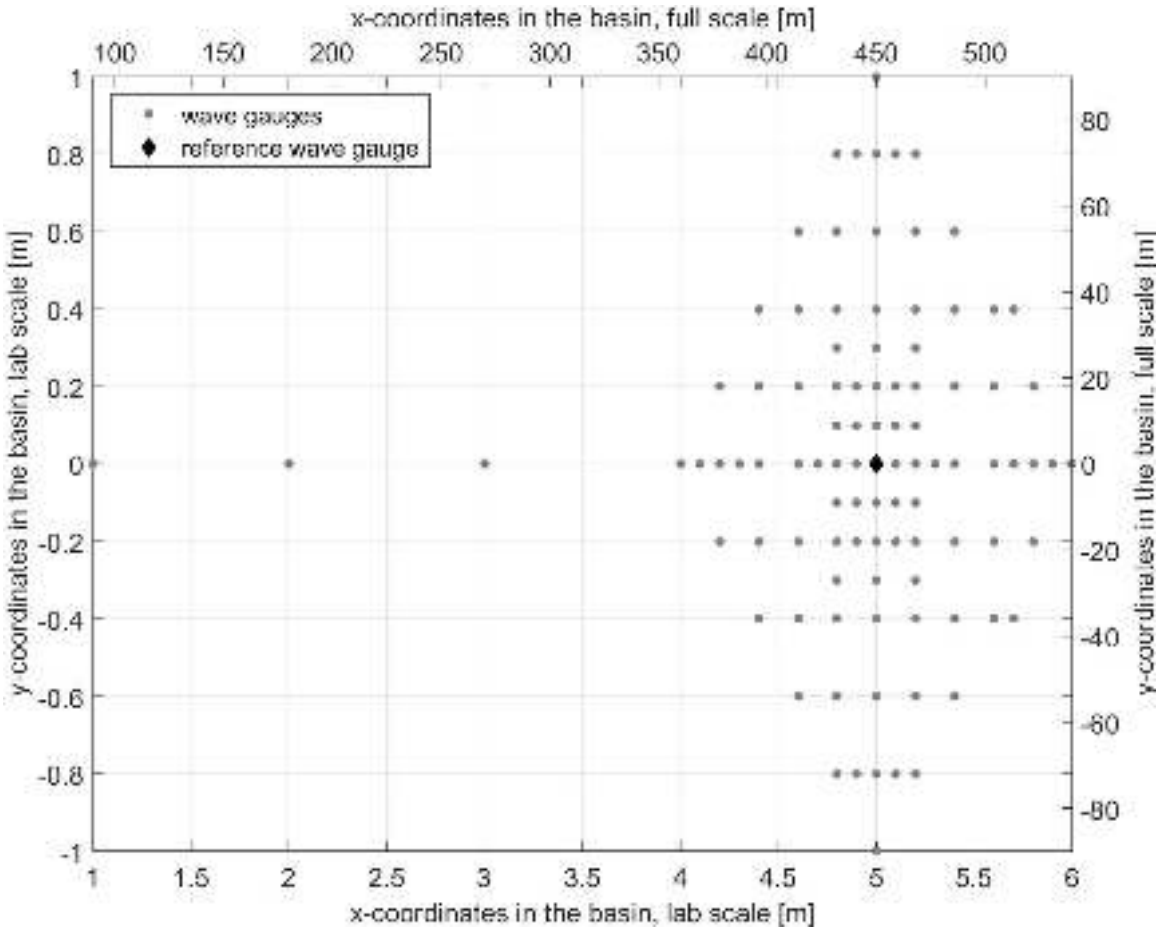


Figure 2: Locations of wave gauges in the wave basin at DHI, including square-shaped test area with dimensions on laboratory and full scales. The location of the reference gauge is marked with a diamond.

Figure 5 illustrates tails $q_{jk}(p)$ of M_{jk} at each of the n_L locations j for the two specific sea states k indicated in Figure 4, one sea state exhibiting (wave) breaking (Figure 5(a)) and the other no breaking (Figure 5(b)). In both figure panels, the upper solid (red) curves correspond to tails (full-scale) measured at each of the wave gauges, and the dashed curve to the tail of the areal maximum over all n_L gauges. It is clear that there is a small number of erroneous gauge measurements which are eliminated from subsequent analysis (and the estimation of areal maxima in the plot) following post-experiment assessment of gauges and measurements by experts at DHI. Nevertheless, there is also considerable variability between tails for different locations for a given sea state.

Upper tails estimated using numerical pre-selection

Tank experiments are costly and time-consuming. Observations of extreme events in a sea state are by their very nature rare. In the current work, observations of $n_H = 105$ values of hourly maxima M for each of $n_S = 50$ sea states are available. Estimation of return values corresponding to return periods $\gg 100$ hours using these observations therefore involves extrapolation. Supplementing these observations with those likely to correspond to long return periods is therefore of considerable advantage in characterising the tail of the distribution of M. Numerical modelling makes this feasible, provided that the numerical models adequately quantify the competing effects of non-linear wave amplification and wave breaking well present at intermediate water depths. In the current work, numerical simulation was performed using the compact formulation of Schäffer and Steenberg (2003) consistent with the work of Sharma and Dean (1981). Numerical simulations conducted by LIC Engineering and SchäfferWaves are designed so that they can be coupled to the wave paddle control system of the physical tank. In principle, therefore, up to the quality of the numerical modelling and wave tank test, the same waves can be generated in both the physical wave tank and its numerical simulator; of course, a numerical simulator provides only an approximation to the actual wave field. Nevertheless, we can use the numerical simulator to search for intervals of time-series of random paddle inputs which yield extreme crests at the reference location; see Appendix B for details. In particular, we might execute the numerical simulation of very long periods (corresponding to thousands of hours), and isolate sequences of paddle inputs which generate the very largest values for crest height at the reference location. We could then generate wave fields in the physical tank using the same sequences of paddle inputs, and compare observed and

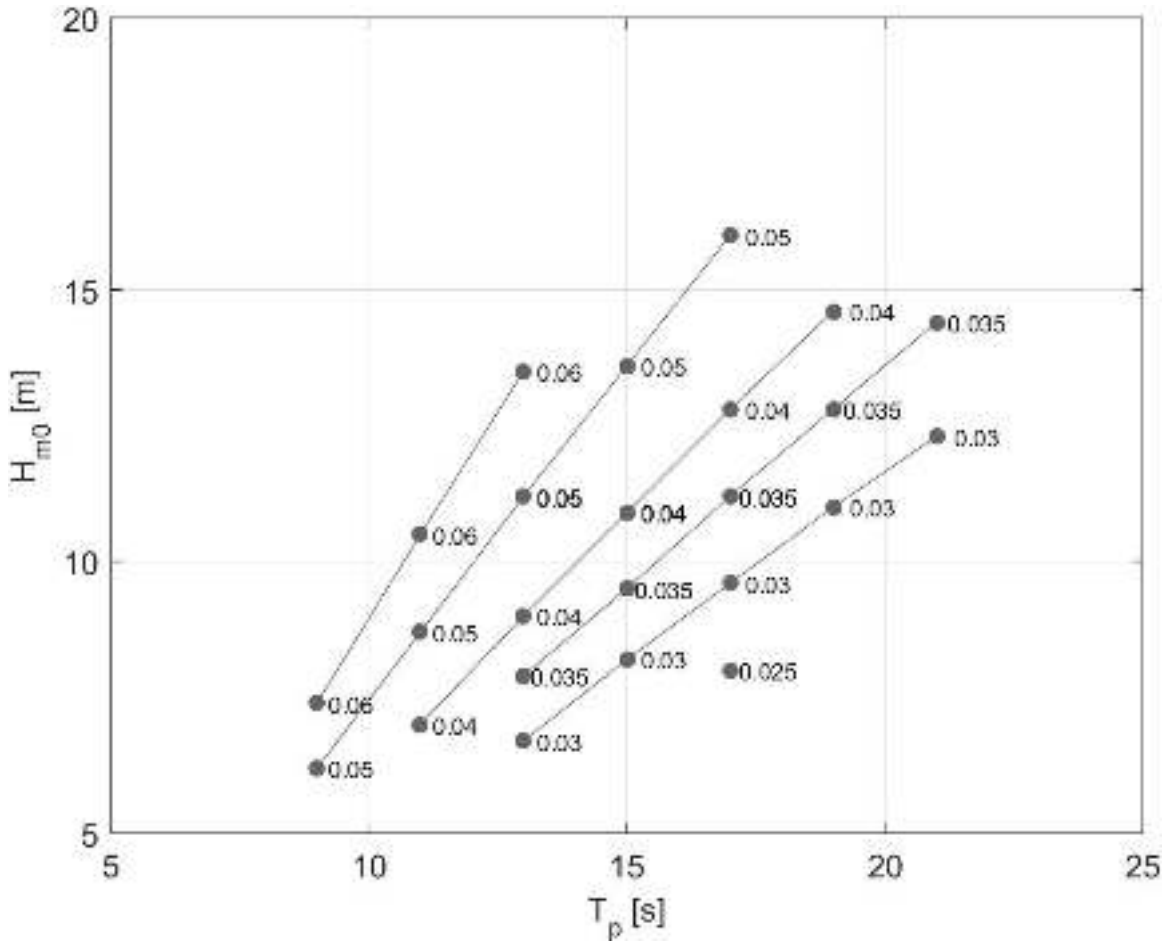


Figure 3: Locations of $n_S = 50$ sea states in H_s-T_p space. Lines indicate a constant steepness ϵ. Labels give values of ϵ. Note that different values of directional spreading θ_{sp} are also considered, not shown here.

numerically modelled crest heights. A useful numerical model will identify sequences of random inputs which actually provide extreme events in the physical tank. Suppose we execute the numerical tank for periods of T hours, with $T \approx 10^4$. The largest hourly maximum observed will not correspond to the T-hour event, because the numerical simulator is an imperfect model for the physical tank. However, it will correspond to a return value with different return period $\tilde{T}$, with $\tilde{T} \approx T$.

The numerical pre-selection and splicing procedure is described in Appendix B. Figure 6 provides a schematic representation of the process by which numerical pre-selected events for a given hour and sea state are used to provide an estimate for the basin hourly maximum for that hour. Using numerical simulation, paddle inputs corresponding to extreme numerical events for that hour are isolated. Paddle inputs corresponding to the single largest numerical event (only) are subsequently simulated in the wave basin. The dependence between the numerical maximum and the corresponding basin event, generated using the same paddle inputs, is therefore estimable. We cannot estimate the corresponding dependence for the k^{th} largest events, $k = 2, 3, ...$, since only the $k = 1$ event is simulated in the wave tank. We know from the discussion of Γ above that the basin hourly maximum does not always correspond to the numerical maximum for the same paddle inputs, yet we only have basin measurement of this event. We therefore use this basin measurement as an uncertain, biased estimate the actual basin hourly maximum, with uncertain, biased return value $\tilde{T}$ (estimated using Γ) in place of the nominal return value T. The statistical characteristics of Γ are estimated using the set of 105 hours of continuous simulations available for both numerical model and wave basin.

We assume that $\tilde{T}$ is a random variable dependent on T. In this section, we outline how the distribution $F_{\tilde{T}}$ of $\tilde{T}$ can be approximated. The analysis is performed independently per sea state at the reference location R. Suppose we have a sample of n_H hours of observations $\{m_i\}_{i=1}^{n_H}$ in the physical tank, and corresponding numerical simulations $\{\tilde{m}_i\}_{i=1}^{n_H}$ based on the same sequences of random paddle inputs, as follows. From the numerical simulations, we identify the realisation i^* which yields the largest value of maximum crest at the reference location

$$i^* = \underset{i \in [1,2,...,n_H]}{\text{argmax}} \{\tilde{m}_i\}.$$

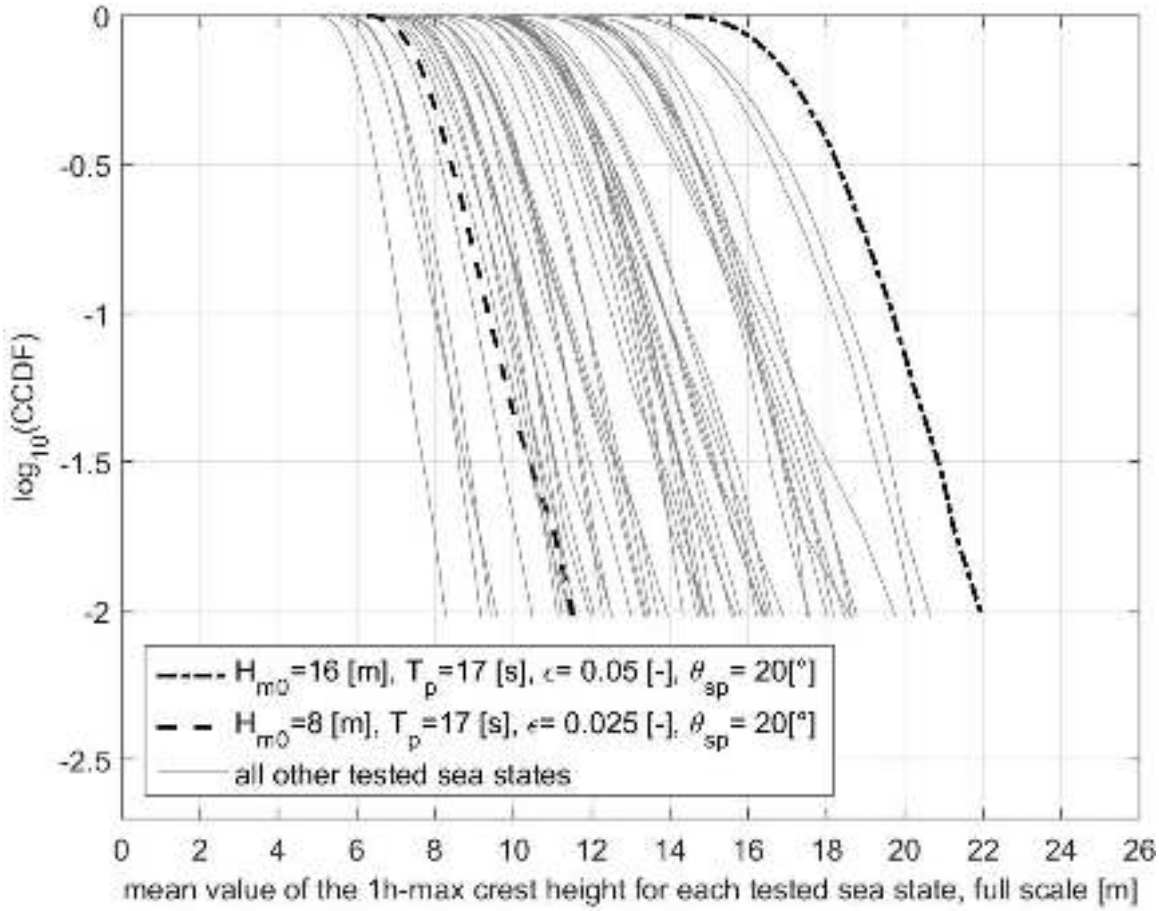


Figure 4: Tail of distribution of hourly maximum crest per sea state k, plotted as the mean quantile $\bar{q}_k(p)$ against $\log_{10}(1-p)$ for non-exceedance probability p. Tails for two sea states, used for illustrative purposes throughout this work, are highlighted. The (left-hand) dashed tail corresponds to a sea state exhibiting no wave breaking. The (right-hand) dot-dashed line corresponds to a sea state exhibiting wave breaking. "CCDF" refers to the complementary cumulative distribution function" or tail distribution, equal to one minus the cumulative distribution function.

Next we identify the rank γ of the corresponding physical crest m_{i^*} within the set of crests $\{m_i\}_{i=1}^{n_H}$ using

$$\gamma = \sum_{i=1}^{n_H} I\left(m_i \geq \tilde{m}_{i^*}\right),$$

where I is the indicator function with value unity when its argument is true, and zero otherwise. γ can be interpreted as follows: the set of random inputs yielding the largest value of hourly crest at the reference location in numerical experiments for a sea state, yields the Γ^{th} largest value in the actual physical tank experiment. In this sense, if the numerical experiment corresponds to a return period of T hours, the physical experiment corresponds to a return period of

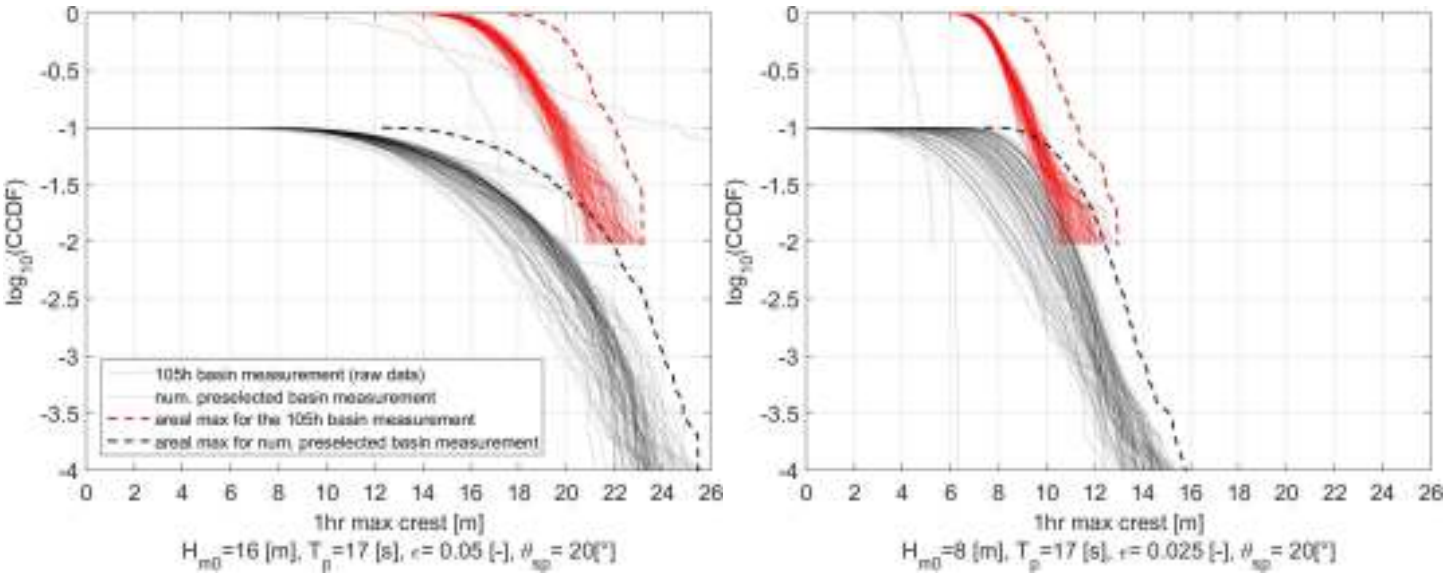
$$\tilde{T} = \frac{T}{\Gamma}$$

where γ is an estimate of random variable Γ. We then use the set of n_S values $\{\gamma_k\}_{k=1}^{n_S}$ across sea states to estimate the distribution F_Γ as

$$F_\Gamma(x) = \frac{1}{n_S} \sum_{k=1}^{n_S} I\left(\gamma_k \leq x\right).$$

Using F_Γ we are able to sample at random from the distribution of Γ and equivalently from the distribution $F_{\tilde{T}}$ of $\tilde{T}$ at each iteration of Bayesian inference, to incorporate our uncertainty about the value of $\tilde{T}$. By definition, the value of $\tilde{T}$ cannot exceed that of T.

In Figure 5, the lower (black) curves correspond to tails (each consisting of $n_M \approx 1000$ hourly maxima) measured using wave tests with n_M numerically pre-selected sequences of random paddle inputs. These sequences produced the largest n_M crests in the vicinity of the reference location for the sea state of interest in second-order numerical simulation for a period of 10^4 hours



(a) (b)

Figure 5: Tails for the two illustrative sea states exhibiting (a) wave breaking and (b) no wave breaking. Measured quantiles $q_{jk}(p)$ of hourly crest maxima for n_L locations j per sea state k, from direct measurement of n_H hours (top red) and from measurements using numerically pre-selected realisations (bottom black). Solid curves correspond to tails for specific locations. Dashed curves correspond to the tails of areal maxima. Black curves start at (0,-1) since they correspond to the 1000 numerically-preselected largest values based on a 10,000-year numerical simulation.

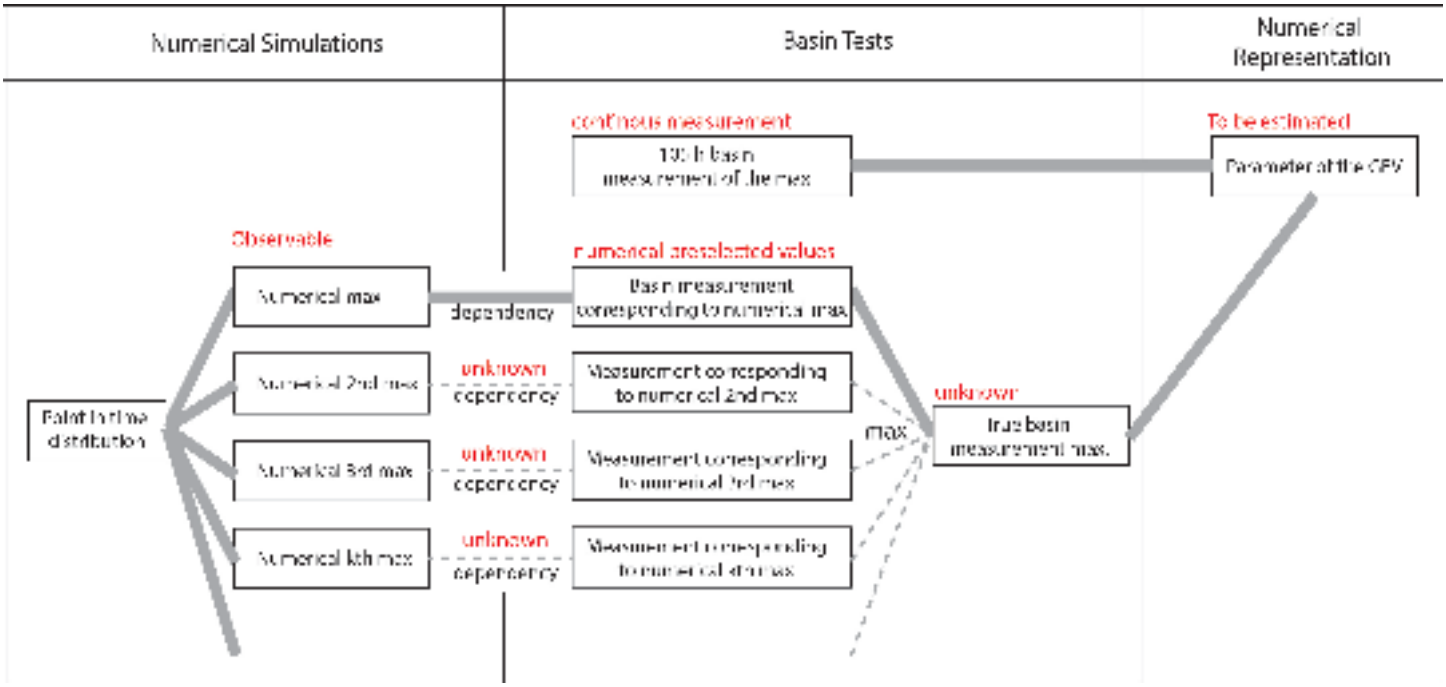
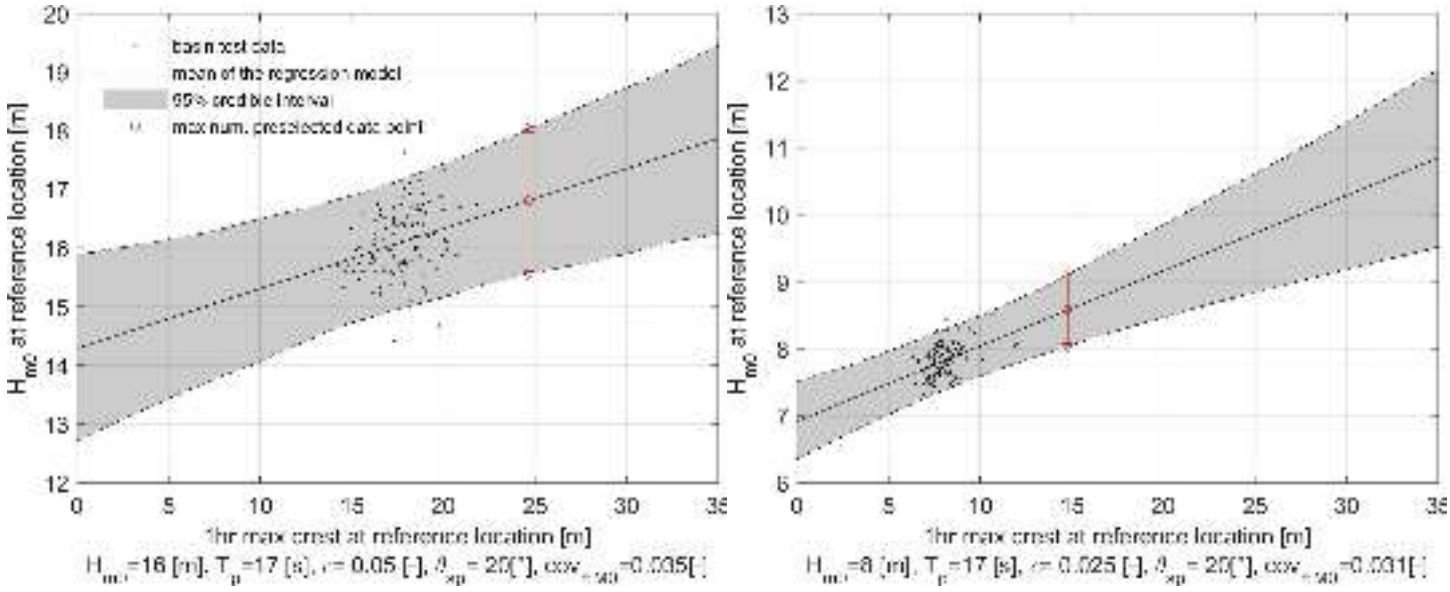


Figure 6: Schematic describing how numerical pre-selection is used to provide an uncertain, biased estimate of the hourly maximum event for crests and wave heights, for a given hour and sea state. Using numerical simulation, random paddle inputs corresponding to extreme numerical events for the hour are isolated. Paddle inputs corresponding to the single largest numerical event (only) are subsequently simulated in the wave basin, allowing the dependence between numerical maximum and the corresponding basin event, generated using the same paddle inputs, to be estimated. Estimation of dependence for the k^{th} largest events, $k = 2, 3, ...$ is not possible since only the largest numerical event is simulated in the wave tank. For statistical inference, the single basin measurement is adopted as an uncertain, biased estimate the actual basin hourly maximum, with uncertain, biased return value $\tilde{T}$ in place of the nominal return value T. The statistical characteristics of $\tilde{T}$ are estimated using the set of 105 hours of continuous simulations available for both numerical model and wave basin.

(full-scale). In this sense, the largest values of crest elevation at each location from the numerically pre-selected tails provide an estimate for the 10^4 hour maximum. Since the numerical wave simulation does not reflect the full physics of extreme waves, it is unlikely that maxima from numerical pre-selection actually correspond to the 10^4 hour maximum. However, observations from numerical pre-selection do provide evidence for the approximate location of the far tail of the hourly



(a) (b)

Figure 7: Illustration of estimation of hourly H_{m0} for numerically pre-selected measurement, required for local normalisation, for sea state exhibiting (a) wave breaking and (b) no wave breaking.

maximum crest distribution, and prove useful in estimating the upper tail, as discussed in Section 3. We use the notation $M_{\tilde{T}}$ to refer to the $\tilde{T}$-hour maximum of E or H obtained in this way. For definiteness, we use the terms "direct measurement" and "numerically pre-selected measurements" respectively to describe measurements made using conventional Monte Carlo wave test, and measurements using pre-selected sequences of random paddle inputs. Only wave basin measurements are used to estimate tail distributions; numerical simulation is used only to identify sequences of random inputs for numerical pre-selection.

Long-term and local normalisation

Extreme value models are estimated for laboratory data, normalised with respect to H_{m0} at the reference location (Figure 2), aiding the development of dimensionless models applicable for arbitrary H_{m0}. The target value of H_{m0} for a particular wave test is of course specified, but the actual value of H_{m0} achieved at the reference location is unknown and must be estimated; here we consider two different estimates achieved using different approaches to normalisation. The first "long-term" estimate uses all n_H hours of direct wave test to calculate H_{m0}; in practice, the long-term estimate is always in good agreement with the target value for the wave test. In the second "local" (or "short-term") approach, we normalise using the value of H_{m0} calculated for that specific hour of wave test. Normalisation using the two approaches leads to somewhat different extreme value models. Models estimated using the first long-term normalisation are likely to be more useful for practical design, and incorporate uncertainty due to hour-to-hour variability in H_{m0} for a given (fixed) long-term value. The current measurements suggest that the coefficient of variation of hourly H_{m0} over all sea states tested relative to the long-term value is ≈ 0.035. The second local normalisation assumes more certain temporally-local knowledge of H_{m0}.

For long-term normalisation, we assume that the same value of long-term H_{m0} (estimated over all n_H hours) is appropriate for both direct and numerically pre-selected data. For local normalisation using direct measurements, hourly estimates of H_{m0} are available. For numerically preselected extreme events however, only a relatively short interval of time series before and after the time of occurrence of the extreme event is available; this is insufficient to estimate hourly H_{m0} reliably. Therefore, hourly H_{m0} for numerically pre-selected events must be predicted empirically from the direct measurements. A simple linear regression model is used to estimate the hourly H_{m0} at the reference location as a function of the observed hourly maximum crest at the reference location, using the direct measurements. For our two illustrative sea states, estimated models for local normalisation are shown in Figure 7. There is considerable uncertainty in the predicted hourly H_{m0}, which is propagated into all subsequent inferences.

Henceforth, we adapt our notation such that M refers to hourly maxima $E_{\max}$ or $H_{\max}$ of E or H, suitably normalised.

3. Model components

We seek to establish statistical models for the distribution of hourly maximum crest height $E_{\max}$ and wave height $H_{\max}$, normalised with respect to H_{m0}, applicable to sea states similar to those listed in Table A.1 of Appendix A. To do this, we

make use of the sample of normalised direct measurements of $\{M_{jk}\}$ and normalised measurements of $\{M_{\tilde{T}jk}\}$ corresponding to numerically pre-selected realisations of the $\tilde{T}$-hour maximum. We perform inference in two parts: first, independently per sea state k, we estimate GEV models for hourly maxima using Bayesian inference; then, we develop response surface models to interpolate between the values of GEV model parameters for different sea states. Throughout the analysis, we are careful to propagate uncertainty so that predictions of distributions of hourly maxima for a previously unseen sea state reflect aleatory and epistemic uncertainty reasonably.

For a given sea state, we assume that the wave field generating $\{M_{jk}\}$ and $\{M_{\tilde{T}jk}\}$ is temporally and spatially homogeneous over the lattice of locations for each hourly observation. It follows that all measurements from all locations are useful in estimating a common marginal distribution F_{M_k} for sea state k. However, observations from different locations are spatially dependent, making joint modelling over all locations a challenging problem in spatial extremes (e.g. Ribatet et al. 2012); we seek to avoid addressing estimation of spatial (extremal) dependence (e.g. Wadsworth et al. 2017) directly in this work, adopting more pragmatic yet useful approximate approaches.

Here, we first outline the different approaches to approximate inference investigated by us, before describing the preferred approach eventually adopted. We subsequently report the analysis for normalised hourly maximum crest height in some detail in Section 4, providing a summary only of the analysis for normalised wave height. Since inference is performed independently per sea state, we suppress the subscript k in this section for clarity. Further, all probability density and cumulative distribution functions of normalised hourly maxima M are assumed dependent on a common set of GEV parameters ξ, σ, μ; all likelihoods are functions of these parameters.

3.1. Using observations at a single reference location only

As introduced in Section 1, the distribution of normalised maxima (of wave or crest height) can be described using the generalised extreme value distribution $\text{GEV}(\xi, \sigma, \mu)$. For a sample $\{m_{iR}\}_{i=1}^{n_H}$ at the reference location R, the sample likelihood is then

$$L_R = \prod_{i=1}^{n_H} f_M(m_{iR}).$$

In principle, any location might be used in place of the reference location, given the assumption of spatial marginal homogeneity. However, the fact that the wave basin is calibrated to produce the desired wave characteristics at the reference location would suggest that this location be used.

3.2. Using mean quantiles over all locations

We can also use the mean quantiles $\{\bar{q}(p_i)\}_{i=1}^{n_H}$ over all locations as a pseudo-sample of n_H values following the same marginal distribution, with likelihood

$$L_Q = \prod_{i=1}^{n_H} f_M(\bar{q}(p_i)).$$

In this pseudo-sample, averaging over locations reduces variability in the tail of the distribution, compared with using observations of normalised maxima from a single location. However, this approach does not reduce any bias present due to the experimental procedure.

3.3. Using areal maximum M_A

Since we assume that the wave field is temporally and spatially homogeneous over the lattice of locations for each hourly realisation of each sea state, it follows that the distribution F_{M_A} of normalised areal maximum M_A over all n_L locations must take the form

$$F_{M_A}(m) = F_M^{\theta_D}(m)$$

where $\theta_D \in [1, n_L]$ is sometimes known as the extremal coefficient, quantifying the extent of dependence between normalised maxima observed at different spatial locations. $\theta_D = 1$ corresponds to perfect dependence, so that all locations experience exactly the same value of M, and $\theta_D = n_L$ to perfect independence, so that M at different locations are independent; in the latter case, a 1-hour areal maximum corresponds exactly to a n_L-year maximum at any one location. In reality, of course, the value of θ_D for a particular wave field is unknown.

We could choose to estimate the distribution of normalised maximum M for any single location together with the areal maximum M_A jointly, assuming that both follow GEV distributions with different parameters, necessitating the estimation of 6 parameters in total. Since we are modelling wave and crest heights at intermediate water depths, it is reasonable to assume that the distributions of both M and M_A have a finite upper end-point expressible in terms of the corresponding GEV parameters. This upper end-point must therefore be common to both M and M_A, allowing elimination of one of the 6 parameters. Of course, strictly we would also need to model the extremal dependence between M and M_A, but might be able to justify assuming that this was small in practice.

3.4. Estimation of distribution of approximately independent maxima for a subset of locations

Extreme events at neighbouring locations exhibit spatial dependence. Nevertheless, we might select a subset I of locations with sufficiently large inter-locations distances, such that values of M can be regarded as effectively independent on the subset. The corresponding likelihood for sample $\{m_{ij}\}_{i=1,j\in I}^{n_H}$ is

$$L_I = \prod_{j\in I}\prod_{i=1}^{n_H} f_M(m_{ij}).$$

Independently per sea state, the procedure for subset selection of independent locations is to: (a) first include a single central location (typically R) in the subset; then (b) iterate, choosing the nearest location to any location in the existing subset, with absolute value of Pearson correlation (between values of normalised maxima M for the two locations) less than a pre-specified threshold δ; and finally (c) stop when no more locations can be added to the subset. This procedure, with $\delta = 0.25$, was found to yield subsets of locations with relatively low dependence between values of M. In practice, the number of independent locations per sea state varies from 1 to 9, in the $16,200\ m^2$ square-shaped basin test area (full scale), reflecting the extent of spatial dependence present.

3.5. Estimating the distribution of the $\tilde{T}$-year maximum $M_{\tilde{T}}$ for uncertain $\tilde{T}$, from numerical pre-selected realisations

For numerically pre-selected realisations corresponding to return period $\tilde{T}$, the distribution of $M_{\tilde{T}}$ at any location is $F_{M_{\tilde{T}}}(m) = F_M^{\tilde{T}}(m)$. If we are prepared to represent our uncertainty about $\tilde{T}$ using a "prior" distribution with density $f_{\tilde{T}}$ and cumulative distribution function $F_{\tilde{T}}$ on some domain $\mathcal{I}_{\tilde{T}}$, we can then estimate the distribution of $M_{\tilde{T}}$ using

$$F_{M_{\tilde{T}}}(m) = \int_{t\in\mathcal{I}_{\tilde{T}}} [F_M(m)]^t\, f_{\tilde{T}}(t)\ dt,$$

and use this to estimate a sample likelihood for an observation $m_{\tilde{T}}$ of $M_{\tilde{T}}$, given by

$$L_{\tilde{T}} = \int_{t\in\mathcal{I}_{\tilde{T}}} t f_M(m_{\tilde{T}})\, [F_M(m_{\tilde{T}})]^{t-1}\, f_{\tilde{T}}(t)\ dt.$$

Of course, the observation $m_{\tilde{T}}$ itself could correspond to a direct measurement at a reference location (as in Section 3.1, yielding likelihood $L_{\tilde{T}R}$), or to the mean maximum over all locations (as in Section 3.2, yielding likelihood $L_{\tilde{T}Q}$), or to observations over a subset of approximately independent locations (as in Section 3.4, yielding likelihood $L_{\tilde{T}I}$).

3.6. Selecting a reasonable estimation strategy

The purpose of tail modelling for individual sea states k is estimation of the joint distribution of GEV parameters ξ, σ, μ of the normalised hourly maximum value M at any single location. We could use any combination of modelling components outlined above to achieve this. In practice, we did not find modelling of areal maximum advantageous. For some sea states, observations at the reference location were missing. For these reasons, these modelling components were deemed less useful. As a result, an approximate likelihood L combining estimation using independent subsets of locations L_I, mean quantiles L_Q and numerical pre-selection (at the reference location) $L_{\tilde{T}R}$ was adopted

$$L = L_I + L_Q + L_{\tilde{T}R}$$

for marginal inference per sea state. The approximate likelihood is a heuristic attempt to exploit the complete data available (from multiple dependent locations, from direct and numerically pre-selected tests) in a broadly reasonable manner for inference; we do not claim optimality in any sense. Note that L_Q effectively contributes n_H observations (or "degrees of freedom") with reduced variability to the inference; the extent of reduction in variability increases with reduced spatial dependence. For a sea state with n_I approximately independent locations, L_I contributes approximately $n_H \times n_I$ degrees of freedom.

Bayesian inference is used independently per sea state, with approximate likelihood L, to generate samples from the joint posterior distribution of GEV shape ξ, scale σ and location μ for that sea state. The inference procedure is outlined in Appendix A.

3.7. Modelling GEV parameters across sea states

There is a large literature (Section 1) on the development of parametric models for wave and crest height distributions, the parameters of which are specified in terms of covariates such as significant wave height H_{m0}, spectral peak period T_P, steepness ϵ and directional spreading $\theta_{\rm sp}$ characterising the sea state from which the observations are made. In this section, we develop joint regression models for GEV parameters ξ, σ and μ for observations of normalised hourly maximum wave and crest heights at the reference location in terms of non-dimensional Ursell number U and steepness ϵ, and directional spreading angle $\theta_{\rm sp}$. Since the sea states selected for the study provide reasonable coverage of the domain of covariates at the location of interest, we adopt response surface models with intercept, linear, square and interaction terms in covariates. Therefore we estimate the model

$$\{\boldsymbol{y}|\boldsymbol{x}\} = \boldsymbol{x}'\boldsymbol{\beta} + \boldsymbol{\nu}$$

where $\boldsymbol{y}$ is the triad (ξ, σ, μ), $\boldsymbol{x}$ is a *design* vector consisting of known values of intercept, linear, square and interaction terms in the three covariates $(U, \epsilon, \theta_{\rm sp})$ and $\boldsymbol{\beta}$ is a matrix of regression parameters to be estimated. $\boldsymbol{\nu}$ is a vector of correlated random Gaussian errors, assumed to have zero mean. In developing the regression model, we are careful to propagate uncertainty from tail fits to individual sea states into estimates of the regression parameters $\boldsymbol{\beta}$ and the covariance matrix of $\boldsymbol{\nu}$. Details of the regression inference are given in Appendix A.

3.8. Applying crest and wave height models at different water depths

Models developed here for the distributions of ratios of hourly maxima $E_{\rm max}$ and $H_{\rm max}$ to H_{m0}, are parameterised in terms of ϵ_1, U and $\theta_{\rm sp}$ from a designed experiment on a pre-specified covariate domain (e.g. Table A.1). The models are applicable to situations where the competing influences of non-linear wave amplification and wave breaking as similar to those studied here. We believe that our models should therefore be applicable for sea states with values of ϵ_1, U and $\theta_{\rm sp}$ on the domain specified by Table A.1.

Nevertheless, all basin measurements were conducted at water depth of 0.5m corresponding to $d_0 = 45$m at full scale, for maxima over time period τ_0 corresponding to one hour (i.e. $\tau_0 = 1$). At any other depth d, using Froude number scaling, the distributions here correspond to maxima over a time period τ in hours given

$$\tau = \tau_0 \left(\frac{d}{d_0}\right)^{1/2}$$

with $\tau_0 = 1$ and $d_0 = 45$. Thus for $d = 60$m, the models developed here correspond to distributions of maxima for time periods of $\tau = 1.16$ hours. To obtain the distribution of hourly maxima at depth d, therefore, we need to "power down" the distribution at d_0. In the notation of Section 1, the cumulative distribution function $F_{M|d}$ of hourly maxima at depth d is related to the cumulative distribution function $F_{M|d_0}$ of hourly maxima at depth d_0 by

$$F_{M|d}(m) = \left(F_{M|d_0}\right)^{1/\tau}.$$

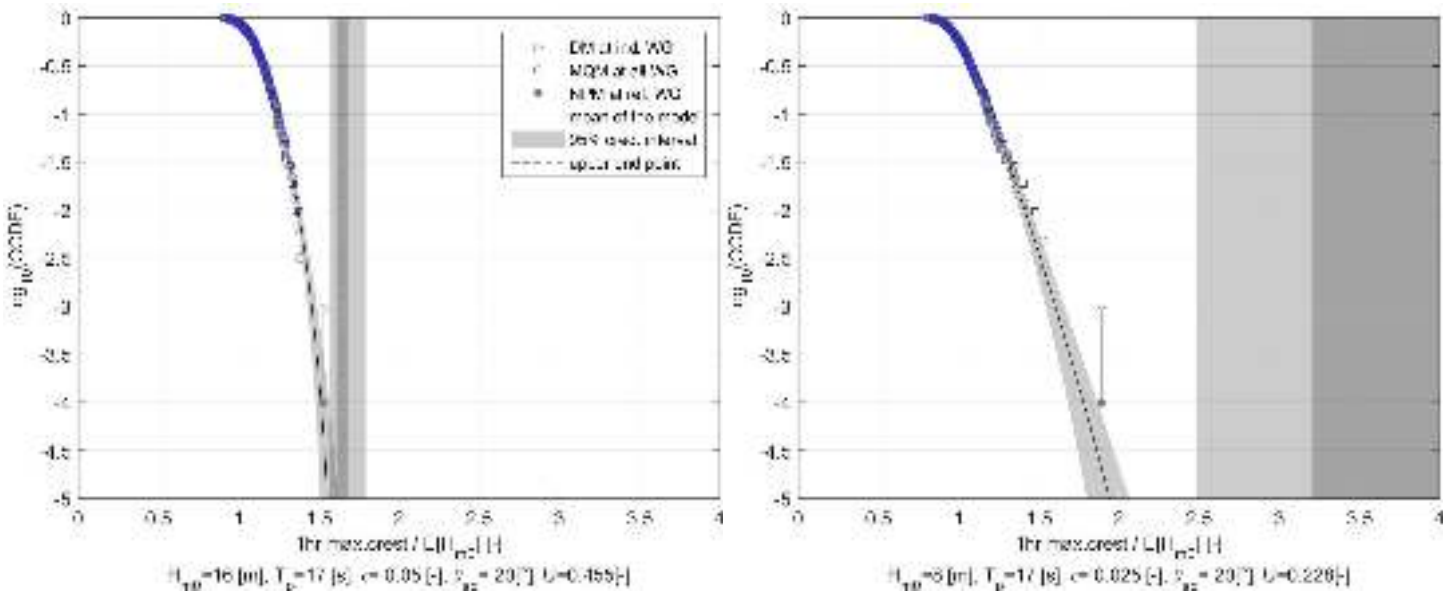
4. Models for the distribution of normalised $E_{\rm max}$ and $H_{\rm max}$

In this section we describe inferences made for distributions of long-term and locally normalised hourly maximum crest and wave height. Since the procedure for wave height is in essence the same as that for crest height, we provide a more complete description of crest modelling, and simply summarise inference for wave height. Freely-available algorithms accompanying this article provides descriptions of distributions of both hourly maximum crest and wave height, allowing the choice of local or long-term normalisation.

4.1. Marginal extreme value models for $E_{\rm max}$

Using the approximate likelihood described in Section 3.6, we estimate the triad of GEV shape ξ_k, scale σ_k and location μ_k for long-term normalised hourly maximum crest for each of $n_S = 50$ sea states k in turn, using Bayesian inference as outlined in Appendix A. In brief, independently per sea state, we first specify uniform prior distributions for ξ_k, σ_k and μ_k. Next we generate a chain of triads of GEV parameters which eventually converges to a correlated sample from the posterior joint distribution of parameters, using a Metropolis-Hastings with Gibbs Markov chain Monte Carlo scheme.

The resulting estimated tail distributions for the two illustrative sea states are shown in Figure 8. The estimated tail is represented by its posterior predictive mean and 95% credible interval. Wave basin data are represented by measurements at the independent set I of locations (shown as circles, see Section 3.4), by the mean quantile (as squares, see Section 3.2), and by the numerically pre-selected maximum (as discs, see Section 3.5). The magnitude of uncertainty in the numerically pre-selected maximum for specified return period $\tilde{T}$ is indicated by the vertical arrow. Also shown is the estimated upper end point of the normalised crest height distribution, represented in terms of its mean value (dashed line), 50% and 95%



(a) (b)

Figure 8: Tails for long-term normalised crest for the two illustrative sea states (a) exhibiting wave breaking, (b) not exhibiting wave breaking. Direct measurements summarised by sample values at approximately independent subset of locations (DM, circles), and mean quantiles measurements (MQM, $\bar{q}(p)$, squares). Numerically pre-selected measurement (NPM) shown as disk. Vertical arrow represents extent of uncertainty associated with estimation of $\tilde{T}$. Posterior predictive tail shown as mean and 95% credible interval. Distribution of upper end point illustrated in terms of mean, inter-quartile range and 95% credible interval. The x-axis of panel (b) is truncated at 4.

credible intervals. Agreement between observation and estimated model is good. For the sea state exhibiting wave breaking, we are considerably more confident about the shape of the tail distribution, and the location of the upper end point.

Figure 9 summarises the posterior distribution of model parameters for the two illustrative sea states. Unsurprisingly in light of Figure 8, the posterior density for ξ suggests considerably more negative values for the sea state exhibiting wave breaking, indicating a considerably shorter tail. Posterior densities generally have a Gaussian shape, with the exception of that for ξ in the non-breaking sea state.

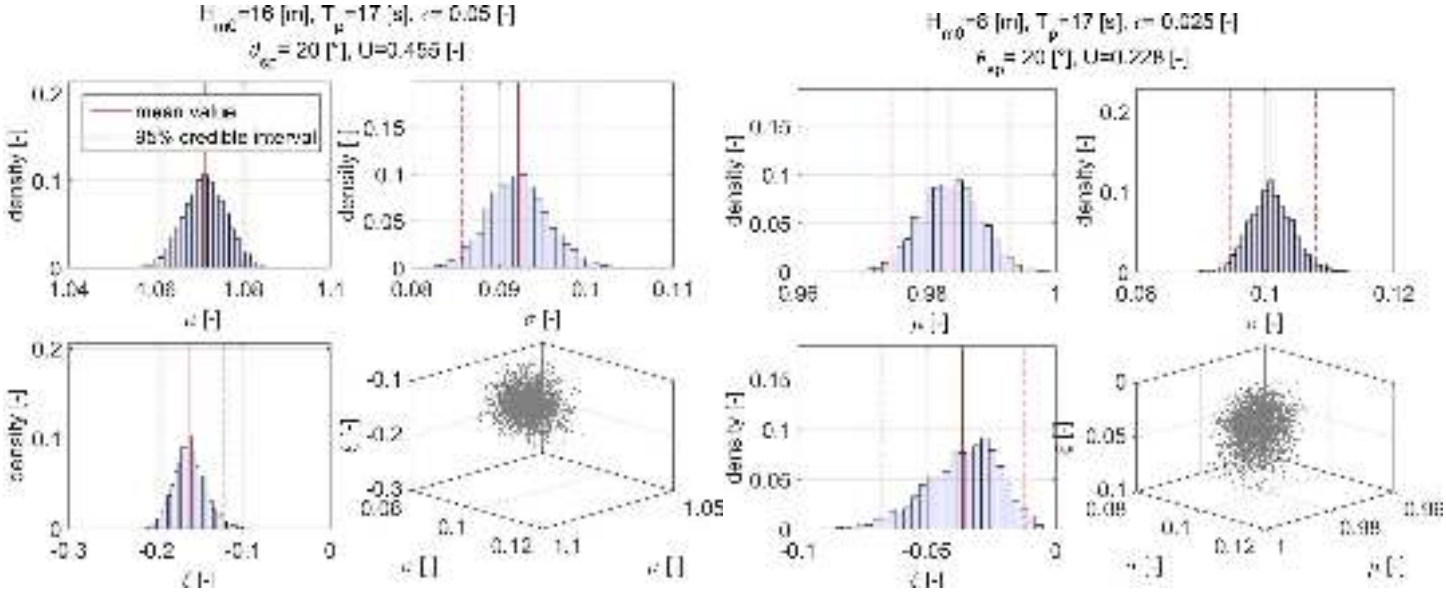
The corresponding estimated tail distributions for locally-normalised crests are shown in Figure 10. The characteristics of this figure are similar to those of Figure 8; an additional feature is the horizontal arrow representing "horizontal" uncertainty in the location of the numerically pre-selected measurement, due to our imprecise knowledge of H_{m0}. The goodness of fit of extreme value tails to both long-term normalised and locally-normalised data appears comparable.

4.2. Response surface regression for $E_{\max}$

Having estimated marginal extreme value models for sea states, and generated correlated samples from joint distributions of GEV parameters, we next estimate response surface models (Section 3.7) relating those GEV parameters to the summary characteristics of sea states, as explained in detail in Appendix A. In brief, we draw a triad of values of GEV parameters for each sea state at random from its posterior MCMC chain, and combine these to construct a new data set. We then repeat the process to generate a large number of data sets. For each data set in turn, we build a regression model to explain the variation in GEV parameters in terms of variation in sea-state Ursell number U, wave steepness ϵ and directional spreading angle θ_{sp}. We retain an ensemble of regression vectors and residuals (each corresponding to a particular data set) to represent the uncertainty in regression due to marginal fitting of GEV models, and (if required) the correlation between regression residuals for ξ, σ and μ.

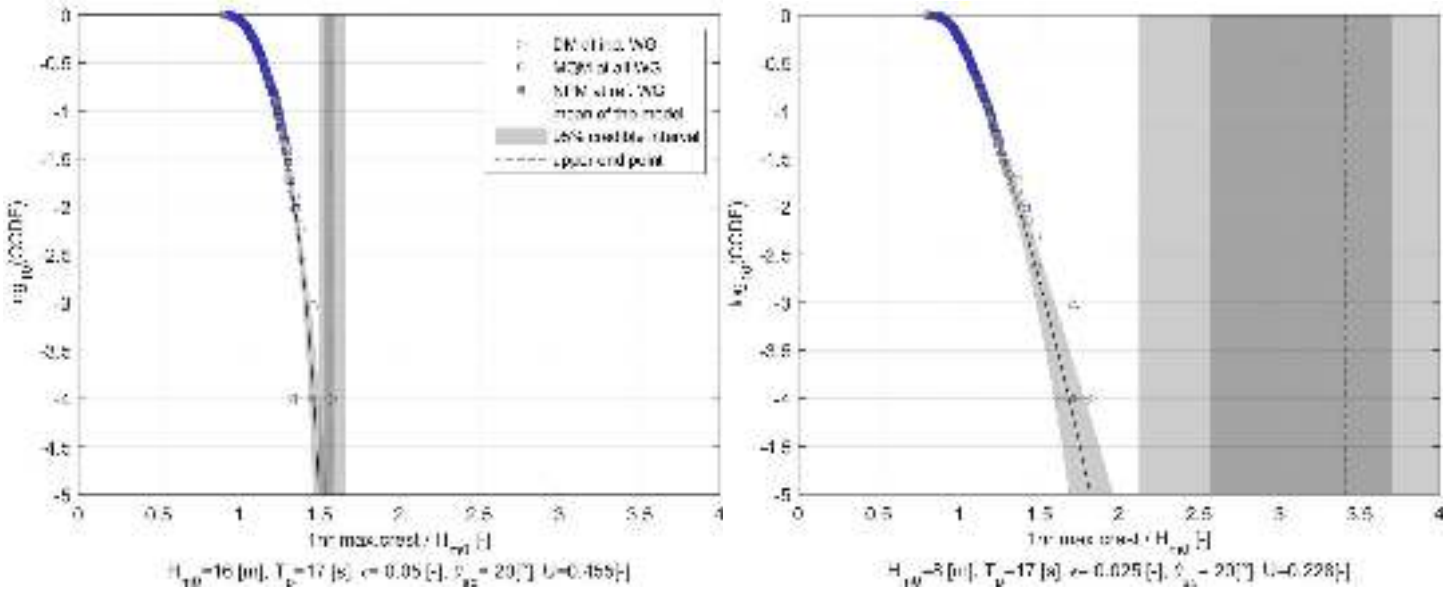
Figure 11 summarises the marginal posterior distributions for GEV parameters, obtained directly from marginal extreme value analysis (in red), or as estimated subsequently from the response surface regression (in grey). Distributions are represented by means (circles) and 95% credible intervals (bars). From the figure it is clear that we are able to describe variation in ξ, σ and μ across sea states reasonably well using the regression analysis; in general, the credible interval based on regression contains that based on extreme value analysis indicating relatively low bias from the regression. However, there is clearly increased uncertainty from the regression estimates, which is only to be expected. In particular, there are some cases for which the regression analysis does not provide much skill in estimating shape parameter ξ. It is interesting that some of the sea states for which estimation of ξ using regression is more problematic (e.g. sea states 6, 20, 21 and 36) have larger steepness, but also lower estimated values for both ξ and σ.

The systematic variation in GEV parameters and quantiles of estimated long-term normalised hourly maximum crest distributions is illustrated in Figure 12. Panel (a) shows contours of posterior predictive mean values for μ, σ and ξ as



(a) (b)

Figure 9: Estimated posterior densities of GEV shape ξ, scale σ and location μ, and scatter plot of posterior parameter samples from MCMC, for the two illustrative sea states (a) exhibiting wave breaking, (b) not exhibiting wave breaking. Long-term normalisation. Density plots also show mean and 95% credible intervals.



(a) (b)

Figure 10: Tails for normalised crest (using local normalisation with estimated H_{m0} for numerically pre-selected measurement) for the two illustrative sea states (a) exhibiting wave breaking, (b) not exhibiting wave breaking. Direct measurements summarised by sample values at approximately independent subset of locations (DM, circles), and mean quantiles (MQM, $\bar{q}(p)$, squares). Numerically pre-selected measurement (NPM) shown as disk. Vertical arrow represents extent of uncertainty associated with estimation of $\hat{T}$. Horizontal arrow represents extent of uncertainty associated with estimation of hourly H_{m0} for numerically pre-selected measurement. Posterior predictive tail shown as mean and 95% credible interval. Distribution of upper end point illustrated in terms of mean, inter-quartile range and 95% credible interval. The x-axis of panel (b) is truncated at 4. To be compared with Figure 8.

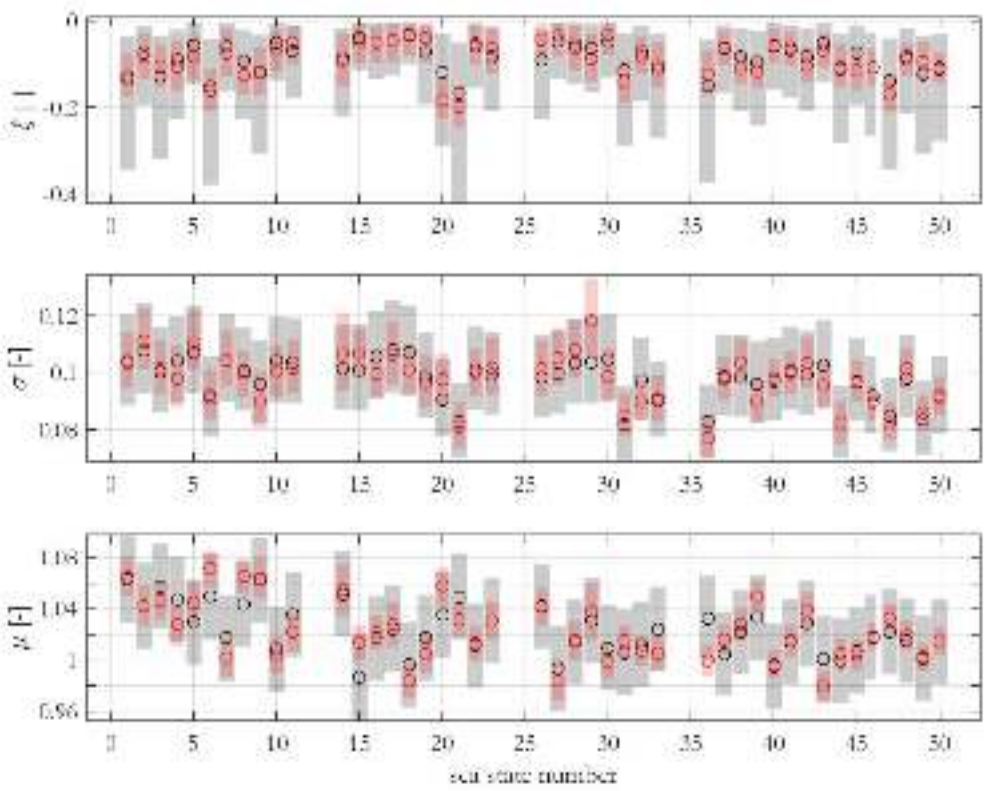


Figure 11: Posterior (predictive) distributions of sea state GEV shape ξ, scale σ and location μ of long-term normalised hourly maximum crest height. From initial extreme value analysis (in red, circle is mean, bar is 95% credible interval) and predicted from subsequent regression model (in grey).

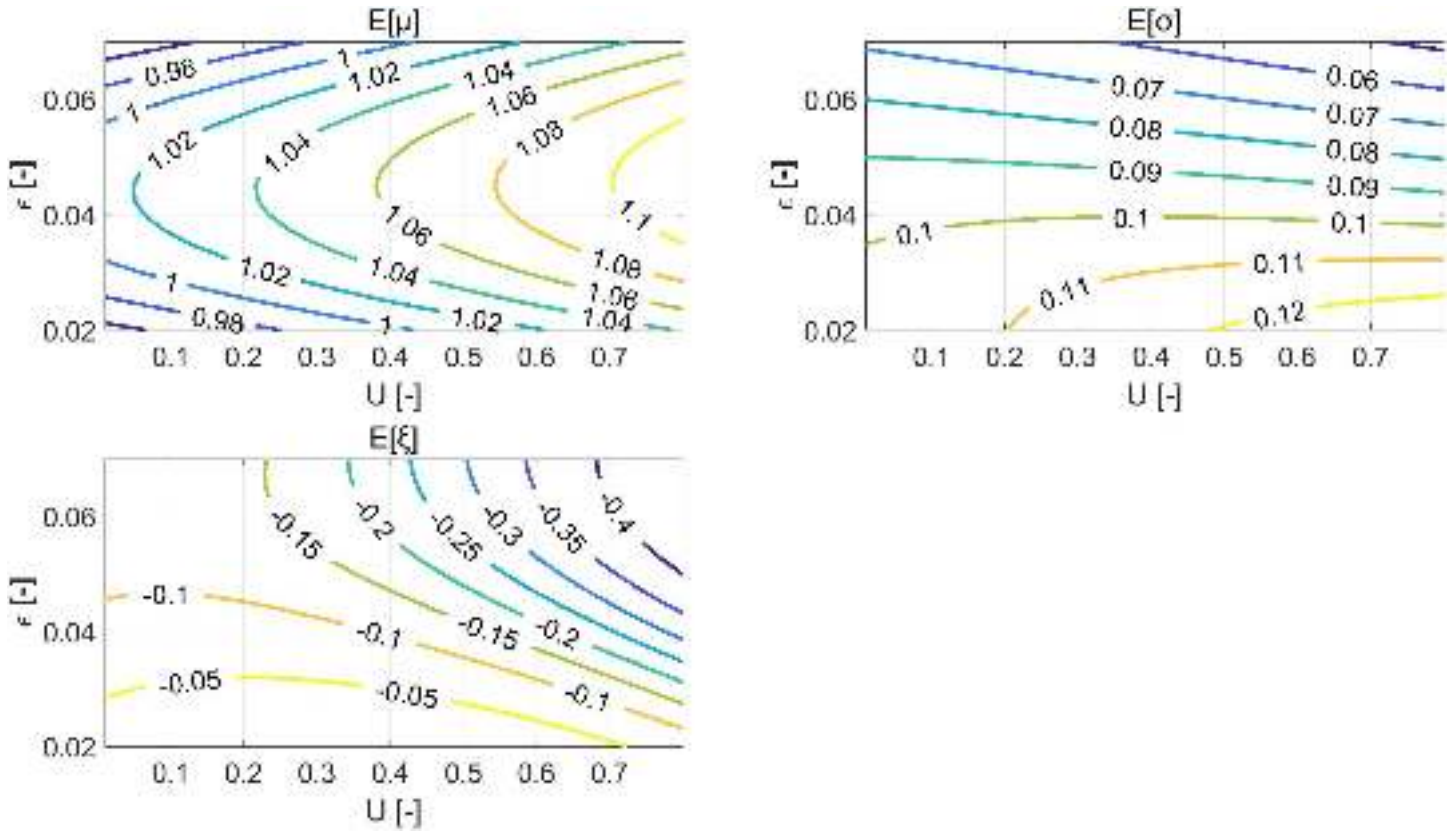
functions of sea state Ursell number U and wave steepness ϵ for the case of directional spreading angle $\theta_{sp} = 20°$. Panel (b) shows the corresponding contours for quantiles with non-exceedance probabilities of 0.5, 0.84, 0.95 and 0.99. As might be expected, the mean value increases with U; it has its maximum for intermediate values of ϵ; variation in μ and the 0.5 quantile are broadly consistent. However, the value of the 0.999 quantile of the distribution is determined primarily by ϵ; smaller wave steepness leads to longer tails of normalised hourly maximum crest. This effect is also clear from the contour plot for ξ.

Figure 13 summarises the relative significance of individual regression terms in the response surface model for parameters of GEV models for long-term normalised hourly maximum crest. Listed on the x-axis of the figure are each individual term considered in the regression for μ, σ and ξ in turn. For each term, the discs (and associated numerical values) give the empirical probability that the posterior density of the corresponding regression parameter does not include zero, and is therefore of greater importance in explaining variation in μ, σ and ξ between sea states. Over-interpretation of the figure can be misleading, because of the various interactions occurring, but nevertheless it is interesting to observe e.g. that the square of wave steepness is important in describing all responses, whereas the square of Ursell number appears to be important only for ξ, and (to a lesser extent) the square of directional spreading angle for σ. Interaction terms between U, ϵ and θ_{sp} appear to be generally important.

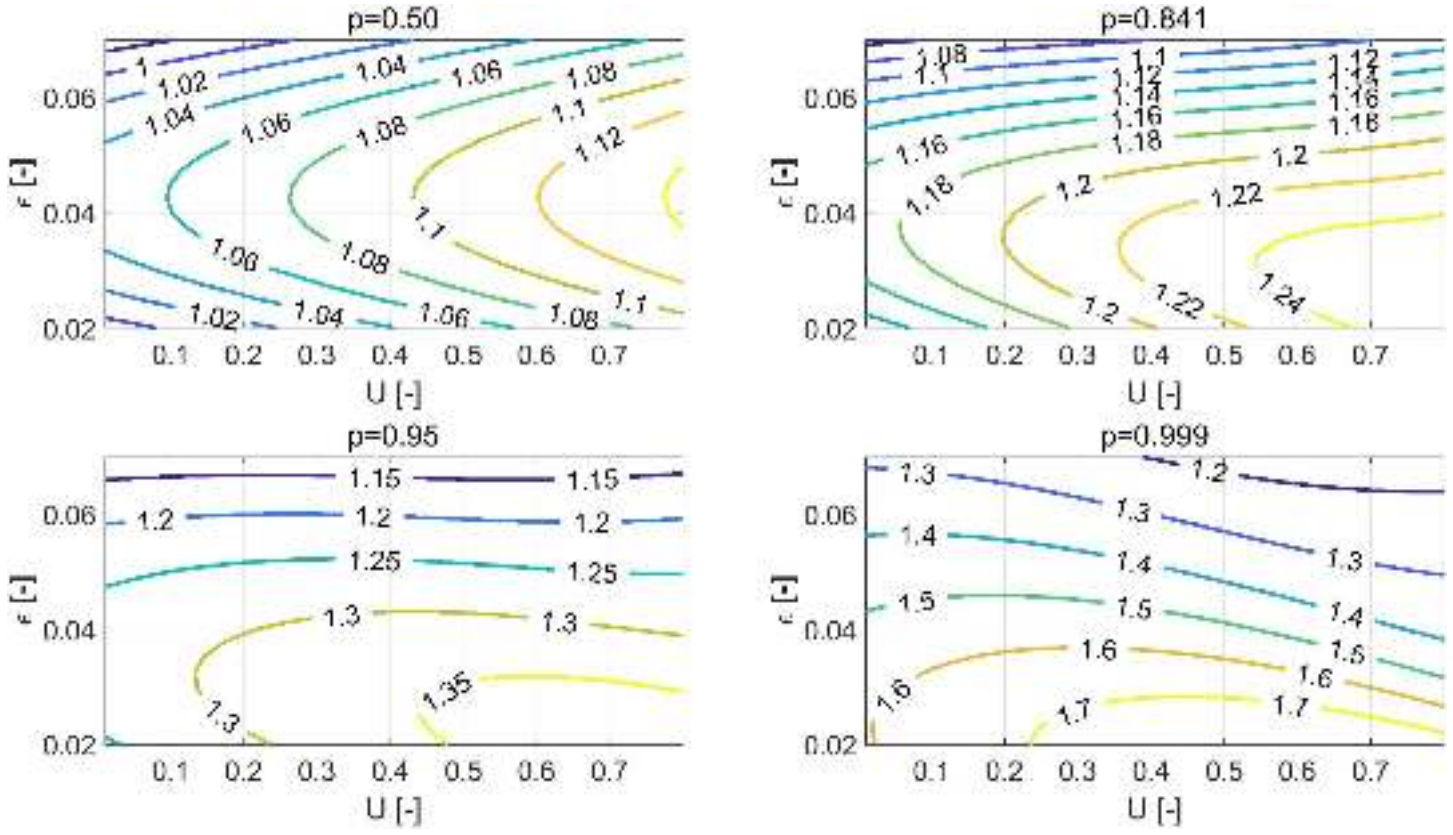
4.3. A model for the distribution of normalised $H_{\max}$

The procedure followed to establish parametric forms for the distribution of normalised $H_{\max}$ is analogous to that described above for normalised hourly maximum crest. For this reason we choose to omit details, and illustrate the inference for normalised hourly maximum wave height in terms of contour plots for GEV parameters ξ, σ and μ and corresponding quantiles of the GEV distribution, as a function on sea state Ursell number U and wave steepness ϵ, for the case of directional spreading angle $\theta_{sp} = 20°$.

Unsurprisingly, contour plots in Figure 14 have a number of features in common with the corresponding figure for crest height (Figure 12), e.g. the contours of σ and ξ. In constrast, extreme quantiles of normalised hourly maximum wave height reduce with U (the reverse of the effect observed in Figure 12 for crest height); this reflects well-established differences in non-linear influences on the distribution of wave height and crest distributions. Interestingly the behaviour of μ is somewhat



(a)



(b)

Figure 12: Tail model for long-term normalised hourly maximum crest. Contour plots of (a) posterior predictive means for GEV shape ξ, scale σ and location μ, and (b) posterior predictive quantiles of GEV tails, from regression model with directional spreading $\theta_{sp} = 20°$.

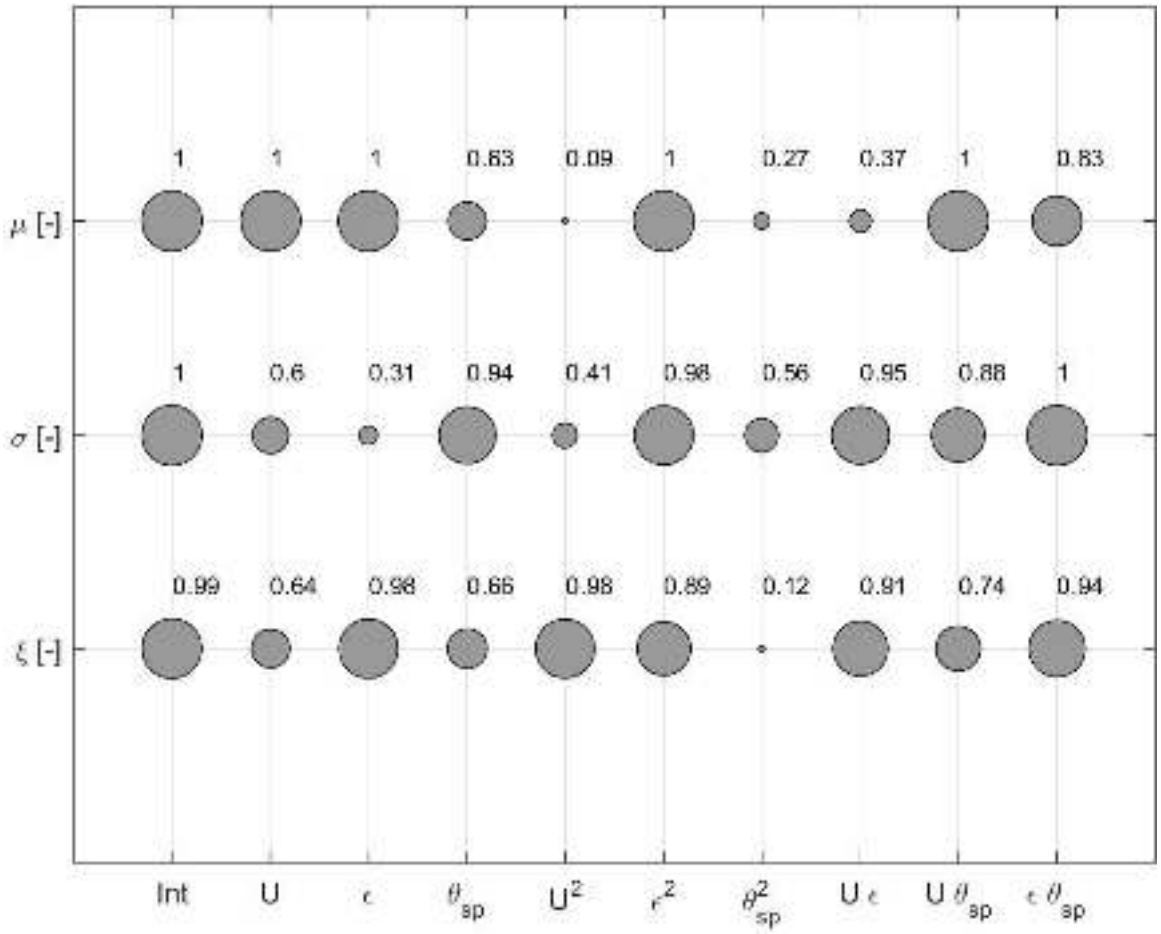


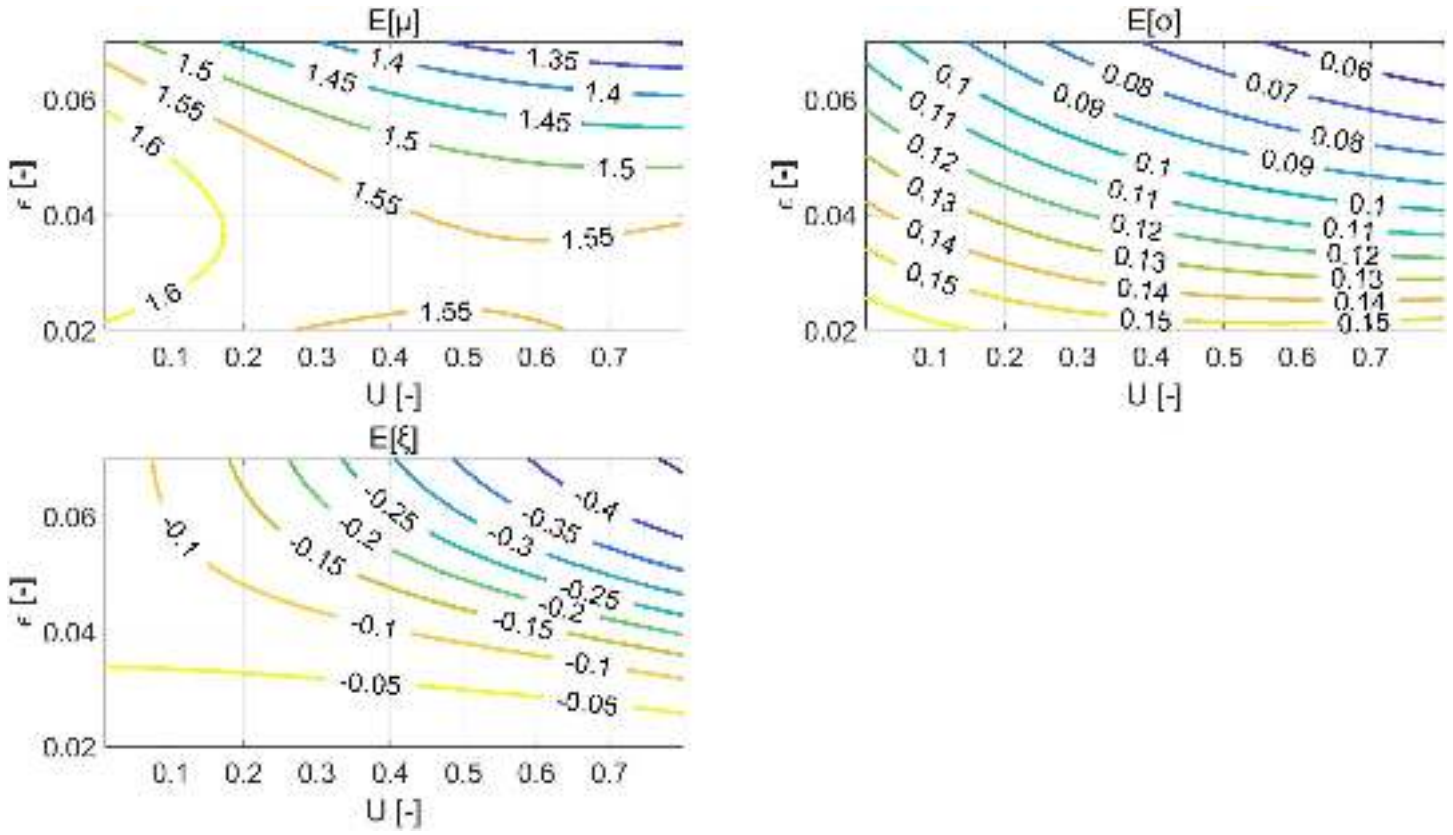
Figure 13: Estimated probabilities that regression terms are non-zero, in regression models for GEV shape ξ, scale σ and location μ for long-term normalised hourly maximum crest heights. Numbers give probabilities, proportional to disc radius in each case. Thus the probability that the coefficient of $U\epsilon$ (Ursell number $\times$ wave steepness) is non-zero for the regression model for GEV shape ξ is 0.91. `Int` corresponds to the regression intercept term.

different for normalised $H_{\max}$ and $E_{\max}$: Figure 14 exhibits a ridge of approximately constant μ (with U) at a steepness of approximately 0.03, which is different to features in the corresponding panel in Figure 12.

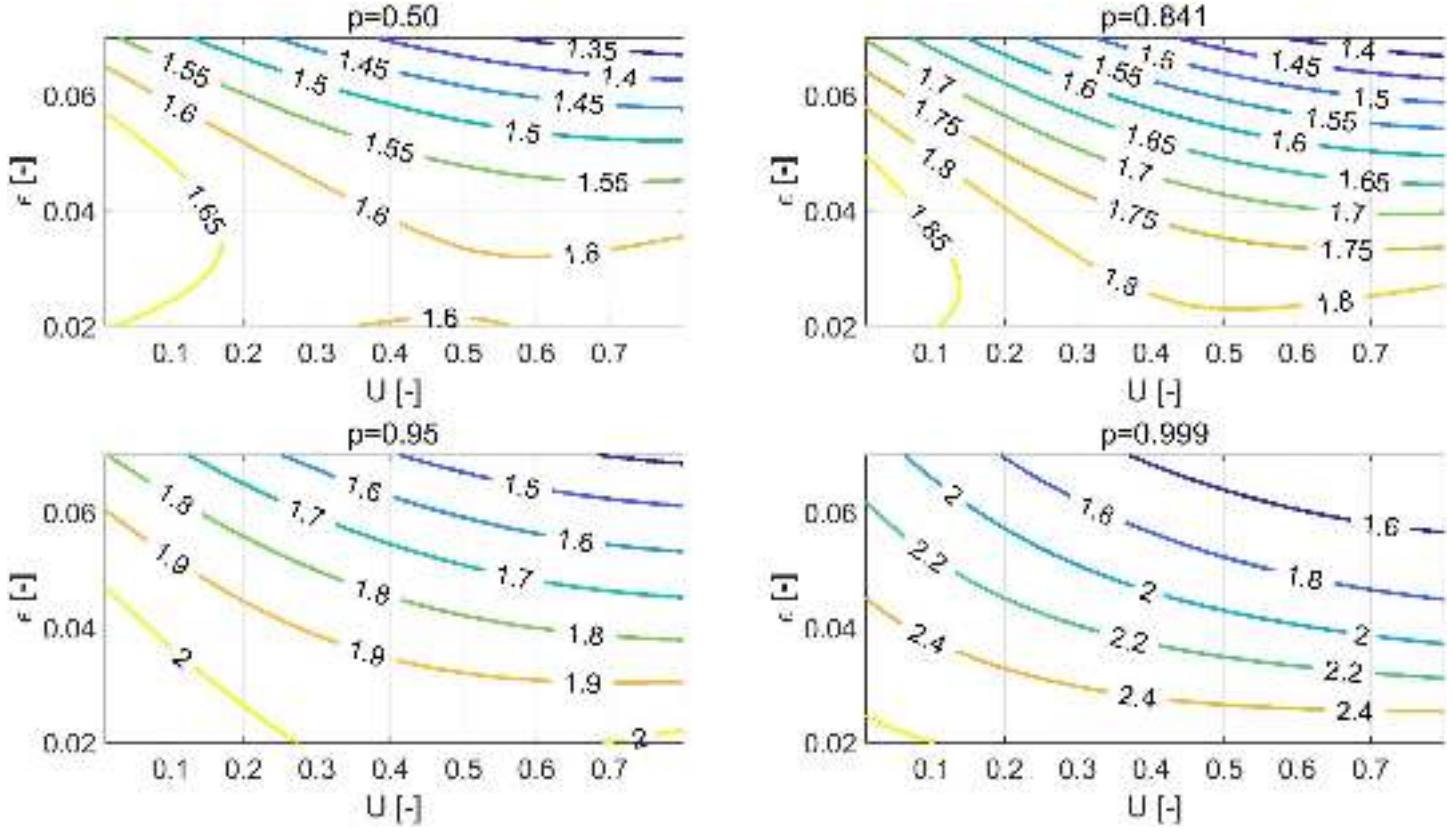
5. Discussion & conclusion

In this work, we report new models for the (probability) distributions of hourly maximum crest and wave height of ocean waves at intermediate water depths, incorporating non-linear effects such as spilling and breaking. Distributions for hourly maximum crests and wave heights, normalised by sea state significant wave height, follow the generalised extreme value (GEV) distribution, the parameters of which are themselves functions of the following sea-state characteristics: Ursell number U, wave steepness ϵ and directional spreading angle θ_{sp}. The distributions are established from measurements of 105 hourly maxima (full-scale) made at the DHI wave basin for 50 sea states, the characteristics of which are given in Table A.1 of Appendix A, and approximately cover the domain (full-scale) $H_{m0} \in [6, 16]$m, $T_P \in [9, 21]$s, $U \in [0.03, 0.65]$, $\epsilon \in [0.3, 0.6]$ and $\theta_{sp} \in [13, 30]^\circ$. The (full-scale) water depth for the wave basin experiments was 45m. Conventional Monte Carlo wave basin tests are supplemented with carefully numerically pre-selected tests, providing measurements at probability levels that without pre-selection would have required many years of basin testing. For numerically pre-selected tests, sequences of random inputs, known to produce large events in second-order numerical Monte Carlo simulations, are used to specify wave basin paddle motion. Numerical pre-selection therefore typically generates large crest and wave events, rarely observed in random Monte Carlo test, allowing better characterisation of the tails of their respective distributions. The numerically pre-selected maximum (NPM) is informative for the location of the shape of the tail of the crest and wave height distributions for all sea states. It provides the approximate location of the quantile level for the distribution of hourly maxima corresponding to very low exceedance probability, not achievable in reasonable time from direct wave basin tests.

Extreme value analysis of tail distributions per sea state, performed using Bayesian inference, indicates that the GEV shape parameter is negative, implying a finite upper end point to the distribution of hourly maximum crest and wave heights as

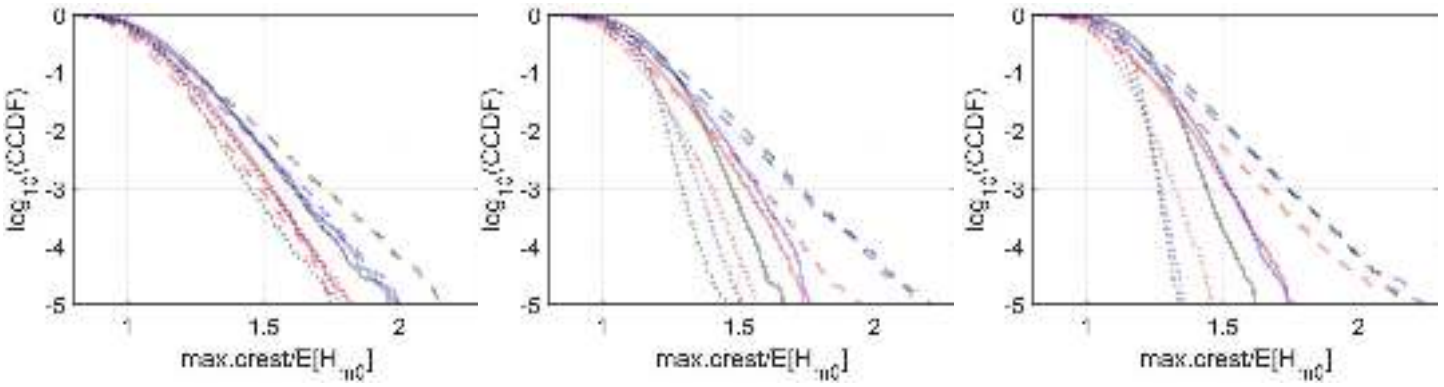


(a)



(b)

Figure 14: Tail model for long-term normalised hourly maximum wave height. Contour plots of (a) posterior predictive means for GEV shape ξ, scale σ and location μ, and (b) posterior predictive quantiles of GEV tails, from regression model with directional spreading $\theta_{sp} = 20°$.



(a) (b) (c)

Figure 15: Posterior predictive tails for long-term normalised hourly maximum crest height, for different combinations of Ursell number U and wave steepness ϵ, with directional spreading θ_{sp} of (a) 13°, (b) 20° and (c) 30°. Nine illustrative sea states chosen are shown in each panel corresponding $\epsilon = 0.025$ (dashed), $\epsilon = 0.04$ (solid), $\epsilon = 0.06$ (dotted), $U = 0.03$ (red), $U = 0.37$ (blue) and $U = 0.7$ (dotted).

might be expected. Response surface regression modelling provides good descriptions of the variation of GEV parameters as functions of sea state characteristics. The influence of each of Ursell number, wave steepness and directional spreading angle on tail distributions is quantified. Figure 15 illustrates the tail distributions for long-term normalised hourly maximum crest, for different combinations of Ursell number, wave steepness and directional spreading angle. For the cases considered, the normalised tail becomes longer with decreasing wave steepness, since the influence of wave breaking is reduced. The influence of increasing directional spreading angle is to exaggerate the effect of wave steepness already noted. The influence of Ursell number is less clear in general, consistent with the findings of Figure 12.

The influence of wave non-linearity affects the parameters of the distribution of hourly maximum wave height differently to that of hourly maximum crest, as can be seen from comparing Figure 12 and Figure 14.

The number n_I of independent locations estimated per sea state, shown in the final column of Table A.1 in Appendix Appendix A, reflects the extent of spatial dependence of hourly maximum crests for the sea state: as the value of n_I decreases, the extend of spatial dependence increases. In the table, the value of n_I shows a general increasing trend with sea state number; this is to be expected, since the sea states are arranged in order of decreasing peak periods and thus also in order of decreasing characteristic wave lengths. The relationship between the tail distribution of hourly maximum value of crest or wave height at a single location and that of the corresponding areal maximum for a given sea state is also indicative of the extent of spatial extremal dependence present in the sea state: informally, as the difference between the distribution of single location and areal maximum decreases, the extent of spatial extremal dependence increases.

The numerical pre-selection procedure used in this work allows the identification of sequences of random inputs (used in numerical simulation to produce extreme events) for use in controlling paddle motion in the wave basin tests, thereby likely to also produce large basin events. Numerical pre-selection might therefore be considered as a form of importance sampling, providing more efficient estimation of tail distributions.

The models estimated here are statistical, and data driven. The choice of GEV distribution to describe maxima events is natural, and supported by asymptotic statistical theory (e.g. Beirlant et al. 2004). However, estimation of GEV parameters, and the systematic variation of those parameters across sea states, is entirely data driven. The estimated trends with Ursell number, wave steepness and directional spreading angle should therefore only be used in an interpolative sense. Use of the current models to estimate distributions for hourly maximum crest and wave heights of sea states beyond the domain indicated in Table A.1 is not advisable.

The estimation strategy outlined in Section 3 adopts an approximate likelihood motivated by ad-hoc considerations to combine observations from multiple dependent locations reasonably without recourse to methods of spatial extreme value analysis. Elementary sensitivity analysis suggests that this approach is reasonable from a practical perspective. However, choice of approximate likelihood is clearly a source of bias and uncertainty which could be better understood.

It is interesting to compare the tails of crest distributions corresponding to some standard forms (Rayleigh, Forristall here, see Forristall 2000) with those estimated in the current work, for our illustrative sea states exhibiting breaking, and no breaking. The Rayleigh and Forristall distributions describe individual crests; the corresponding distributions $F_M(m)$ of hourly maxima were calculated from the distributions $F_{\text{Ind}}(m)$ for individual crests using $F_M(m) = F_{\text{Ind}}^{n_{\text{Wav}}}$, where n_{Wav} is an estimate for the number of waves per hour, calculated using $n_{\text{Wav}} = 3600/T_1$, where T_1 is the mean spectral period in seconds. Results are given in Figure 16. In particular, it is clear that breaking induces obvious differences between the tail of the Forristall distribution, and the tails estimated in the current work. The tail of the distribution of maximum hourly crest for the non-breaking case is considerably longer than that suggested by Forristall. This is inevitable given the value of NPM measured for the sea state, illustrated in Figures 8 and 10 (for long-term and local normalisation).

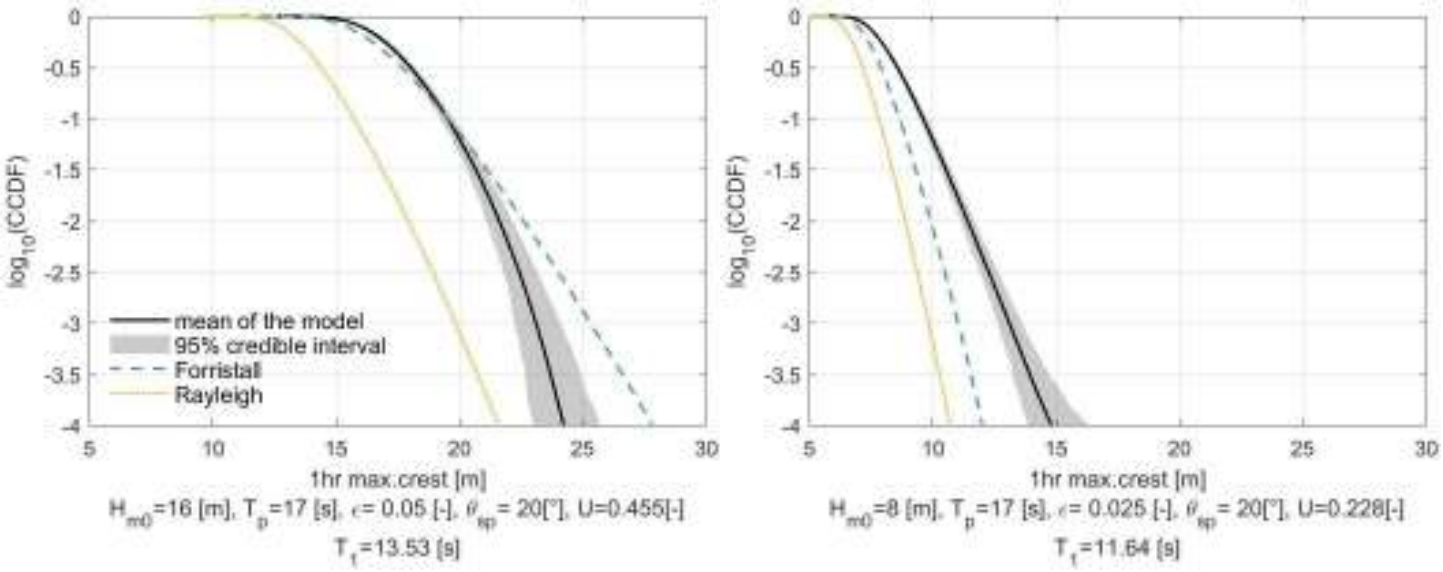


Figure 16: Comparison of estimated tails for sea states (a) exhibiting breaking and (b) not exhibiting breaking. Shown in each panel are posterior mean and 95% credible interval from the current work, together with Rayliegh and Forristall tails for the given sea state parameters.

Uncertainties from both marginal extreme value estimation, and response surface estimation, have been carefully characterised in this work. For this reason, the freely-available software accompanying this publication provides estimates for posterior predictive tail distributions for given sea state characteristics, i.e. including estimation of credible intervals for the tail as well as its expected form. Nevertheless, there are numerous sources of bias and uncertainty in the current work which are more difficult to quantify. We characterise the distribution of hourly maximum paddle-generated crests and wave heights in a laboratory-scale wave basin experiment. We assume that inferences made are relevant to describe the characteristics of full-scale wind-driven storm waves in the ocean, even for sea states exhibiting considerable wave breaking. It is likely that our planned experimental campaign in the wave basin provides considerably more efficient coverage of the domain of extreme events than an equivalent period of observation of the open ocean. Measurement uncertainties for offshore measurements of extreme events in harsh environments will be larger than in a controlled laboratory test environment. As in any measurement numerous sources of uncertainty are also present in the current work: e.g. how well can a wave basin with finite dimensions and given scale, and a given number of paddles (with given specifications) simulate a homogeneous sea state with given characteristics? How well does the wave basin setup mimic field conditions regarding wave non-linearity and breaking, with only a limited number of characteristic wave lengths between the wave-maker and measurement gauges? How does the fidelity of simulation vary with e.g. increasing directional spreading angle? How well can the wave gauges estimate the crest heights of breaking waves? Within the project rigorous efforts have been made in order to minimize these uncertainties, and to account for the effects of remaining uncertainty. Other research (CresT, ShorTCresT and Loads JIPs, see Buchner et al. 2011, Hennig et al. 2015) has concluded that wave basin testing is absolutely necessary to quantify tails of extreme waves. The current work attempts to combine numerical and basin modelling of extreme waves wisely to achieve this efficiently.

6. Acknowledgements

This research formed part of the Total “AWARE” project to which numerous partners have contributed greatly. We gratefully acknowledge all these contributions, noting in particular those from Kevin Ewans (MetOcean Research), Graham Feld (Shell), Hans Fabricius Hansen (Haw Metocean, previously DHI), Linas Mikalauskas (LIC Engineering), Hemming Schäffer (SchäfferWaves), Francesco Stevanato (LIC Engineering) and Graham Stewart (GS Offshore). We thank two anonymous reviewers for detailed comments on an early version of the manuscript.

References

Beirlant, J., Goegebeur, Y., Segers, J., Teugels, J., 2004. Statistics of extremes: theory and applications. Wiley, Chichester, UK.

Boccotti, P., 1989. On mechanics of irregular gravity waves. Atti Accademia Nazionale Lincei, Memorie VIII 19, 111–170.

Bredmose, H., Dixen, M., Ghadirian, A., Larsen, T., Schloer, S., Andersen, S., Wang, S., Bingham, H., Lindberg, O., Christensen, E., Vested, M., Carstensen, S., Engsig-Karup, A., Petersen, O., Hansen, H., Mariegaard, J., Taylor, P., Adcock, T., Obhrai, C., Gudmestad, O., Tarp-Johansen, N., Meyer, C., Krokstad, J., Suja-Thauvin, L., Hanson, T., 2016. DeRisk: accurate prediction of ULS wave loads. Outlook and first results. Energy Procedia 94, 379 – 387.

Buchner, B., Forristall, G., Ewans, K., Christou, M., Hennig, J., 2011. New insights in extreme crest height distributions (a summary of the CREST JIP). Proc. 30th Int. OMAE Symp., Rotterdam, The Netherlands .

DNV-RP-C205, 2010. Environmental conditions and environmental loads. Det Norske Veritas.

Ewans, K., 2019. A spectral description for extreme sea states offshore Denmark: Part I Power spectral density function. Proc. 2nd International Workshop on Waves, Storm Surges and Coastal Hazards, Melbourne, Australia .

Forristall, G.Z., 1978. On the statistical distribution of wave heights in a storm. J. Geophys. Res. 83, 2353–2358.

Forristall, G.Z., 2000. Wave crest distributions: Observations and second-order theory. J. Phys. Oceanogr. 30, 1931–1943.

Forristall, G.Z., 2007. The distribution for wave height in shallow water estimated from laboratory measurements. Unpublished (developed and reported as part of LoWiSh Phase 1). .

Gamerman, D., Lopes, H.F., 2006. Markov chain Monte Carlo: stochastic simulation for Bayesian inference. Chapman and Hall / CRC, Boca Raton, USA.

Glukhovskiy, B., 1966. Investigation of sea wind waves (in Russian), Leningrad. Proc. of Sea Climatology Conference , 51–71.

Hansen, H.F., Randell, D., Zeeberg, A.R., Jonathan, P., 2019. Directional-seasonal extreme value analysis of North Sea storm conditions. Accepted by Ocean Engineering, draft at www.lancs.ac.uk/∼jonathan .

Haring, R.E., Osborne, A.R., Spencer, L.P., 1976. Extreme wave parameters based on continental shelf storm wave records. Proc. 15th Int. Conf. on Coastal Engineering, Honolulu, HI .

Hennig, J., Scharnke, J., Swan, C., Hagen, O., Ewans, K., Tromans, P., 2015. Effects of short-crestedness on extreme wave impact: A summary of findings from the joint industry project ShorTCresT. Proc. 34th Int. OMAE Symp., St. John?s, NL, Canada .

Jonathan, P., Ewans, K.C., 2013. Statistical modelling of extreme ocean environments with implications for marine design : a review. Ocean Eng. 62, 91–109.

Karmpadakis, I., Swan, C., Christou, M., 2019. Laboratory investigation of crest height statistics in intermediate water depths. Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences 475, 20190183.

Katsardi, V., de Lutio, L., Swan, C., 2013. An experimental study of large waves in intermediate and shallow water depths. Part I: Wave height and crest height statistics. Coastal Eng. 73, 43 – 57.

Katsardi, V., Swan, C., 2011a. Omae2011-49957: An experimental study of shallow water wave statistics on mild bed slopes. Proc. 30st Int. Conf. of the OMAE, Rotterdam, NL .

Katsardi, V., Swan, C., 2011b. The evolution of large non-breaking waves in intermediate and shallow water. I. Numerical calculations of uni-directional seas. Proc. R. Soc. A 467, 778–805.

Krogstad, H., Barstow, S., 2004. Analysis and applications of second-order models for maximum crest height. J. Offshore Mech. Arctic Eng. 126, 66–71.

Longuet-Higgins, M.S., 1952. On the statistical distribution of the height of sea waves. J. Mar. Res. 11, 245–266.

Mackay, E., Johanning, L., 2018. Long-term distributions of individual wave and crest heights. Ocean Eng. 165, 164–183.

Mendez, F.J., Losada, I.J., Medina, R., 2004. Transformation model of wave height distribution on planar beaches. Coastal Eng. 50, 97–115.

Prevosto, M., Krogstad, H.E., Robin, A., 2000. Probability distributions for maximum wave and crest heights. Coastal Eng. 40, 329–360.

van Ravenzwaaij, D., Cassey, P., Brown, S.D., 2018. A simple introduction to Markov Chain Monte-Carlo sampling. Psychon. Bull. Rev. 25, 143–154.

Ribatet, M., Cooley, D., Davison, A., 2012. Bayesian inference from composite likelihoods, with an application to spatial extremes. Statistica Sinica 22, 813–845.

Schäffer, H., Steenberg, C., 2003. Second-order wave-maker theory for multidirectional waves. Ocean Eng. 30, 1203–1231.

Schubert, M., Jonathan, P., 2019. New crest and wave height distribution. `github.com/MatthiasSchubert/New_Crest_Wave_Height_Model`.

Sharma, J.N., Dean, R.G., 1981. Second-order directional seas and associated wave forces. SPE J. 21, 129?140.

Tayfun, M., 1990. Distribution of large wave heights. J. Waterw. Port Coast. Ocean Eng. 116, 686–707.

Torsethaugen, K., 1993. A two-peak spectral model. Proc. 12th Int. Conf. OMAE. 2, 175–180.

Torsethaugen, K., Haver, S., 2004. Simplified double peak spectral model for ocean waves. Proc. 14th Int. Offshore and Polar Eng. Conf., Toulon, France .

Tromans, P.S., Vanderschuren, L., 1995. Variable based design conditions in the North Sea: application of a new method. Offshore Technology Confernence, Houston (OTC–7683) .

Tychsen, J., 2016. Wave kinematics and hydrodynamic loads on intermediate water depth structures inferred from systematic model testing and field observations - tyra field extreme wave study 2013-2015. Proc. 3rd Offshore Structural Reliability Conference, OSRC2016, 14-16 September, Stavanger, Norway .

Tychsen, J., Risvig, S., Hansen, H.F., Hansen, N.E.O., Stevanato, F., 2016. Summary of the impact on structural reliability of the findings of the Tyra Field Extreme Wave Study 2013-15. Proc. 3rd Offshore Structural Reliability Conference, OSRC2016, 14-16 September, Stavanger, Norway .

van Vledder, G.P., 1991. Modification of the Glukhovskiy distribution. Technical Report. Delft Hydraulics Report H1203.

Wadsworth, J.L., Tawn, J.A., Davison, A.C., Elton, D.M., 2017. Modelling across extremal dependence classes. J. Roy. Statist. Soc. C 79, 149–175.

Wu, Y., Randell, D., Christou, M., Ewans, K., Jonathan, P., 2016. On the distribution of wave height in shallow water. Coastal Engineering 111, 39 – 49.

Appendix A. Achieved values of sea state parameters in the laboratory tests

Table A.1 provides a list of the spectral form, and achieved long-term values (full-scale) of significant wave height H_{m0}, peak wave spectral period T_P, wave steepness ϵ, Ursell number U and directional spreading angle θ_{sp} for the $n_S = 50$ sea states used in the current study. The table also gives the number n_I of independent sea states (Section 3.4) estimated per sea state.

Nr	H_{m0}	T_p	Steepness	Ursell	Spreading directional	Spreading s parameter in cos^2s distribution	Number of independent gauges
[#]	[m]	[s]	[-]	[-]	[°]		[#]
1	14.50	20.72	0.030	0.060	12.60	20.92	3
2	13.62	20.72	0.027	0.054	13.20	21.36	1
3	14.61	18.73	0.041	0.523	12.71	21.57	3
4	13.13	18.73	0.037	0.472	13.58	21.29	2
5	10.88	18.73	0.030	0.391	15.48	20.41	3
6	16.13	16.80	0.051	0.451	12.40	21.07	3
7	13.07	16.80	0.041	0.360	5.61	10.05	1
8	15.42	16.80	0.042	0.376	17.85	20.49	5
9	13.50	16.80	0.037	0.329	12.95	14.31	2
10	11.28	16.80	0.030	0.310	0.19	20.45	4
11	11.16	16.80	0.030	0.310	2.551	19.46	1
12	11.72	17.27	0.030	0.300	15.12	21.91	3
13	11.07	16.80	0.030	0.310	16.59	20.82	3
14	11.31	16.80	0.030	0.312	53.40	13.97	1
15	9.06	16.80	0.022	0.254	5.58	20.89	3
16	9.55	16.80	0.010	0.267	13.45	20.57	1
17	8.89	16.71	0.028	0.242	74.64	14.00	2
18	7.61	16.71	0.025	0.211	14.26	16.47	1
19	13.52	14.95	0.050	0.277	0.08	28.85	3
20	13.73	14.80	0.051	0.274	14.74	20.67	3
21	13.88	14.80	0.052	0.270	30.80	13.01	3
22	10.59	14.80	0.030	0.211	5.60	28.01	4
23	10.93	14.95	0.040	0.224	14.57	20.71	1
24	10.91	14.95	0.040	0.224	14.80	20.66	3
25	9.80	14.80	0.037	0.198	15.57	19.93	1
26	10.22	14.80	0.038	0.204	40.77	13.08	3
27	8.53	14.95	0.031	0.176	6.55	28.39	1
28	8.73	14.95	0.032	0.179	10.65	19.72	3
29	8.05	14.80	0.030	0.160	45.80	12.69	1
30	8.11	14.95	0.030	0.166	19.97	18.36	3
31	13.06	12.95	0.058	0.183	14.45	20.95	4
32	10.93	12.95	0.049	0.153	5.97	28.02	5
33	11.00	12.95	0.050	0.154	15.62	19.55	4
34	11.42	12.95	0.051	0.158	17.10	19.49	5
35	11.40	12.95	0.051	0.158	10.61	19.55	3
36	11.26	12.95	0.050	0.156	43.59	12.75	1
37	8.38	12.95	0.037	0.116	0.82	28.17	0
38	8.69	12.95	0.040	0.123	16.35	19.86	5
39	8.90	13.00	0.040	0.127	41.52	12.42	5
40	7.49	12.95	0.034	0.104	7.02	20.01	0
41	7.91	12.95	0.035	0.109	16.02	19.46	5
42	8.02	12.95	0.036	0.111	44.76	12.30	4
43	7.93	12.95	0.035	0.097	17.46	19.00	4
44	10.11	10.95	0.058	0.084	10.72	19.49	4
45	8.17	10.95	0.047	0.068	6.50	28.56	2
46	8.34	10.95	0.048	0.070	18.10	18.74	5
47	8.20	10.95	0.050	0.073	46.84	12.35	5
48	6.88	11.03	0.039	0.059	17.93	18.91	5
49	6.01	8.28	0.056	0.020	20.91	17.45	5
50	6.24	8.01	0.060	0.022	20.09	17.50	6

Table A.1: List of achieved values for sea state characteristics (full-scale) for the wave basin measurements.

The wave spectral form used for all sea states with the exception of 12, 24 and 34, and 13, 25 and 35, corresponds to that of Torsethaugen and Haver (2004) (and see Torsethaugen 1993) for swell-dominated sea states ($T_P > T_{Pf} = 6.6H_{m0}^{1/3}$) and to the JONSWAP form (with $f^{-4.5}$ spectral tail) for wind-sea dominated sea states; this composite spectrum has been referred to as the "Fabricius Hansen" form by members of the project team. The composite spectral form is also used for sea states 13, 25 and 35, but with a f^{-4} spectral tail. For sea states 12, 24 and 34, a standard JONSWAP spectrum is used with peak-enhancement factor (γ) expressed as a function of T_P/H_{m0} (as per DNV-RP-C205 2010 section 3.5.5.5) with f^{-5} spectral tail. In practice, despite careful experimentation, the measured spectral tail almost always exhibited the f^{-5} form,

suggesting that some of the high-frequency energy input via the wave-paddles is lost immediately when the target spectra tail is heavier than f^{-5}. There is speculation that wind friction (not present in the wave basin) is responsible for maintaining spectral tails heavier than f^{-5}. Further development of the composite spectral form has been performed as reported in Ewans (2019).

Table A.1 suggests that the number of independent gauge locations decreases with increasing T_P. This is to be expected since, as T_P increases, typical wavelengths increase, and hence the number of wavelengths between two gauges reduces. Hence measurements from different wave gauges become more correlated in general.

Appendix B. Numerical preselection (NPS) and splicing

The NPS algorithm

The procedure used for numerical pre-selection, described in Section 3.5, is given here algorithmically. In the notation of Section 2, we proceed as follows:

- Generate random sequences $\mathcal{S}$ of spectral amplitudes, phases and directions ("random inputs") necessary to generate $T = 10^4$ hours of wave time-series (in the wave tank or numerical simulator);
- Use the first portion $\mathcal{S}_0$ of $\mathcal{S}$ to generate $n_H = 105$ hours of continuous wave tank measurements in the target area;
- Use the whole of $\mathcal{S}$ to simulate T hours of waves in the numerical wave basin for the target area;
- Isolate the set (of size T) of hourly maximum crest elevations from the numerical simulation;
- Compare the first n_H hourly maximum crests from wave tank and numerical simulation (corresponding to sequence $\mathcal{S}_0$) to assess the fidelity of the numerical simulation;
- For each of the $n_M = 1000$ largest hourly maxima from the numerical simulation, isolate the corresponding interval $\mathcal{S}_j$ $(j = 1, 2, ..., n_M)$ of $\mathcal{S}$ responsible for generating the maximum (and a short period before and after it). Splice the intervals $\{\mathcal{S}_j\}_{j=1}^{n_M}$ together, creating a spliced sequence $\mathcal{S}_M$, with appropriate ramp-up, warm-up and ramp-down intervals for each $\mathcal{S}_j$ (see '**Splicing**' below). Use the spliced sequence to generate corresponding large crests in the physical wave tank.
- Use the value of the largest crest generated using sequence $\mathcal{S}_M$ (with associated estimate of uncertainty, as explained in the main text) as an estimate of the $\tilde{T}$ $(\approx T)$ hour maximum.

Identification of extreme events in numerical simulations is made on the basis of extreme crests not wave heights. The assumption is therefore made that the n_M intervals which generate extreme crests in numerical simulation will generate both extreme crests and wave heights in the physical tank.

The splicing procedure

Suppose we wish to generate extreme waves in the wave tank for a specific sea state with known wave spectrum, exploiting NPS. The n_M intervals $\mathcal{S}_j$ $(j = 1, 2, ..., n_M)$ of time-series each corresponding to an extreme crest event (identified by NPS) are spliced together into a 'daisy chain' sequence to be used to create waves in the physical wave tank. For each interval $\mathcal{S}_j$, the splicing methodology comprises of (a) ramp-up, (b) warm-up, (c) event and (d) ramp-down. In detail, the ramp-up interval $\mathcal{I}_{\mathrm{RU}}$ is used to ramp up the paddle signal with a cosine function, for 30 wave-maker time steps (where one wave-maker time step corresponds to 0.1s); hence the length of $\mathcal{I}_{\mathrm{RU}}$ is 3s. The warm-up interval $\mathcal{I}_{\mathrm{WU}}$ corresponds to the time required for the slowest frequency components generated to propagate from the wave-maker to the testing area. The duration of $\mathcal{I}_{\mathrm{WU}}$ is 330 time steps (corresponding to 33s); this was ample time even for a 2.5Hz free lab wave component to propagate obliquely from the far ends of the wave-maker to the reference point at the group velocity. The event interval $\mathcal{S}_j$ corresponds to the occurrence of the extreme crest event. The duration of $\mathcal{S}_j$ is 40 time steps (4s), and the extreme crest occurs half way through this interval. The ramp-down interval $\mathcal{I}_{\mathrm{RD}}$ is the time used to ramp down the signal with a sine function, with a duration of 30 time steps (3s).

The ramp-down interval $\mathcal{I}_{\mathrm{RD}}$ following the j^{th} event and the ramp-up interval $\mathcal{I}_{\mathrm{RU}}$ preceding the $(j+1)^{\mathrm{th}}$ event are overlapping tapers. The cosine ramp-up function and sine ramp-down function and their phases are specified such that there is conservation of energy locally, and that horizontal gradients at end points of intervals are avoided.

In addition, the full sequence of n_M intervals for each seat state is not executed as one long 'daisy chain' experiment. Instead, the full sequence is partitioned into sub-sequences or batches, each containing 200 events. Between each batch, an idle interval $\mathcal{I}_I$ of duration 370 wave-maker timesteps (37s) was introduced, where the wave-maker was idle and the water in the wave basin became calm.

Validation of the splicing procedure

Intensive investigations where performed to determine the optimal characteristics of the approach adopted to splicing. These included confirmatory tests that: large events can be identified from numerical simulation for a sea state and reproduced in the wave tank as single events; individual events can be reproduced from spliced time-series; ordering of spliced events does not affect results; and results are not sensitive to other experimental conditions. By means of a specific illustration, the following test was conducted to assess experimental repeatability and to explore the influence of reflections or return currents in the basin on crest characteristics. The test procedure is (a) execute base case (with $H_{m0} = 13.0$m, $T_P = 15$s) test with duration of 10 hours, (b) repeat base case (to assess repeatability), (c) execute base case test with (1h periodic) control signal shifted in time by half a basin resonance period, (d) repeat based case with control signal shifted by greater period (e.g. 0.5h). For each of (a)-(d), the 20 largest largest events would be expected to be common to all runs. Crest elevations and wave shapes for these events were compared and found to be in good agreement. In addition, (e) the largest 20 events from the based case (a) were identified and their corresponding wave-maker inputs used to create a daisy-chain using the splicing procedure described. Further, (f) wave-maker inputs for the 20 events identified were used to create single event runs executes in the basin. Moreover, all of experiments (a)-(f) were performed at five different reference locations in the basin, along the centre line of the wave maker. Reference points were located 3, 4, 5, 6 and 7m from the wave-maker. Therefore, in experiments (a)-(f), the 20 largest events in the base case experiment were executed as a single event, part of a spliced event, and as part of a random time series with and without time-shifting, for each of five different reference locations in the basin. Results indicated that the crest height and wave shape of each event were reproducible, and hence that the influence of any reflections or return currents could not be detected.

Diagnostic tests based on order statistics were also carried out to confirm that the statistical characteristics of the first crest in a batch of 200 from a daisy-chain (generated from calm conditions) are indistinguishable from those of any other crest in the batch, across all batches and sea states. It was also confirmed that the distributions of crests occurring immediately before and after the maximum crest in a batch were indistinguishable from those of any other crest for the batch, for all batches and sea states.

Appendix C. Bayesian tail inference

With reference to Section 3, we outline the Markov chain Monte Carlo procedure used to generate samples from the joint distribution of GEV parameters ξ, σ, μ, using the approximate likelihood L given in Section 3, independently per sea state. We assume we have isolated appropriate data (for normalised hourly crest of wave heights) to evaluate the likelihood components L_I, L_Q and $L_{\tilde{T}R}$ given values of ξ, σ, μ. For those wishing to read more about MCMC (and Metropolis-Hastings within Gibbs), Gamerman and Lopes (2006) provides a standard text, and van Ravenzwaaij et al. (2018) a simple introduction.

First we make a prior specification, that each of ξ, σ, μ is uniformly distributed on intervals $\mathcal{I}_\xi = [-0.2, -0.005]$, $\mathcal{I}_\sigma = (0.5, 1.5]$ and $\mathcal{I}_\mu = [-1, 1]$. Intervals $\mathcal{I}$ were selected based on preliminary GEV fits to data, ensuring that posterior distributions of parameters were not overly influenced by the choice of intervals. The upper limit of $\mathcal{I}_\xi$ was selected to impose a finite (but large) upper bound on the value of crest and wave height. Then we generate a chain of triads of parameter values which eventually, after a period of burn-in, yields a sample from the joint posterior distribution of parameters, using a Metropolis-Hastings within Gibbs inference scheme to sample from the full conditional densities

$$f(\xi|\text{data}, \sigma, \mu) \propto f(\text{data}|\xi, \sigma, \mu) f(\xi)$$
$$f(\sigma|\text{data}, \xi, \mu) \propto f(\text{data}|\xi, \sigma, \mu) f(\sigma)$$
$$f(\mu|\text{data}, \xi, \sigma) \propto f(\text{data}|\xi, \sigma, \mu) f(\mu)$$

in turn. Here $f(\xi)$, $f(\sigma)$ and $f(\mu)$ are prior densities, and $f(\text{data}|\xi, \sigma, \mu)$ is the sample GEV likelihood. In brief, given current state $(\xi, \sigma, \mu) = \{\eta_b\}_{b=1}^3$, we propose a new state $\{\eta_b^c\}_{b=1}^3$ with $\eta_b^c = \eta_b$ for all but one component b^* (with b^* taking values $1, 2, 3$ in cycles), and $\eta_{b^*}^c = \eta_{b^*} + \nu\zeta_{b^*}$, where ν is a standard Gaussian random variate, and $\{\zeta_b\}_{b=1}^3$ are a set of proposal standard deviations tuned to give acceptance rates of ≈ 0.25. We accept the candidate state in preference to the current state with probability

$$\min\left(1, \frac{f(\eta_{b^*}^c|\text{data}, \{\eta_b^c\}_{b\neq b^*}}{f(\eta_{b^*}|\text{data}, \{\eta_{b^*}\}_{b\neq b^*}}\right)$$

and otherwise retain the current state. At each iteration, since the value of $\tilde{T}$ is unknown in $L_{\tilde{T}R}$, we simply sample a value from the distribution $F_{\tilde{T}}$, the estimation of which is described in Section 3.5. We check convergence of the Markov chain to its stationary distribution by performing inference multiple times from different prior starting states, confirming that chains converge to a common joint distribution. Typically, a total of $n_{BI} \in [5000, 10000]$ burn-in iterations was used, and a subsequent chain of length 300,000 (uniformly thinned to n_R =3000) retained for regression modelling. Trace plots of parameter chains from multiple analyses with different starting values were inspected to ensure approximate convergence. For some sea states, considerably longer periods of burn-in were found to be necessary for reasonable convergence.

Appendix D. Regression model

Bayesian inference for the tail of normalised hourly maximum crest or wave height of the k^{th} sea state provides a sample $\{\xi_{rk}, \sigma_{rk}, \mu_{rk}\}_{r=1}^{n_R}$ of n_R correlated estimates from the posterior distribution of GEV parameters. We use the corresponding samples for all sea states as data to estimate a response surface regression model with which to predict GEV parameters in terms of covariates U, ϵ and θ_{sp} as introduced in Section 3. First we construct a sample for regression which consists of a single random drawing of a triad $\boldsymbol{y}$ from each of the posterior samples of GEV parameters (ξ, σ, μ), drawn independently for each of the n_S sea states. Next we construct the corresponding design vector $\boldsymbol{x}$ for each sea state consisting of intercept term ("`Int`" in Figure 13), linear terms in sea state covariates, squares of covariates and interaction terms for covariates. Then we estimate the regression model

$$\{\boldsymbol{y}|\boldsymbol{x}\} = \boldsymbol{x}'\boldsymbol{\beta} + \boldsymbol{\nu},$$

where we assume that error $\boldsymbol{\nu} \sim N(\mathbf{0}, \boldsymbol{\Sigma})$. We repeat the regression analysis a total of n_A times, for new random drawings of triads, to quantify the variability of the estimated regression vector $\boldsymbol{\beta}$ and the estimated distribution for error $\boldsymbol{\nu}$. We could use residuals from the regression as a random sample from the distribution of $\boldsymbol{\nu}$, but in practice find that residuals for ξ, σ and μ are relatively uncorrelated: we therefore choose instead to approximate $\boldsymbol{\Sigma}$ by a diagonal matrix, and estimate its diagonal elements.

Using the accumulated samples of regression vector estimates $\{\widehat{\boldsymbol{\beta}}_a\}_{a=1}^{n_A}$ and error matrices $\{\widehat{\boldsymbol{\Sigma}}_a\}_{a=1}^{n_A}$, prediction of GEV tail parameters for a new sea state with covariates represented by design vector $\boldsymbol{x}^*$ can be made. The empirical distribution of $\boldsymbol{y}^*$ is estimated from the sample $\{\boldsymbol{y}_a^*\}_{a=1}^{n_A}$ calculated using

$$\{\boldsymbol{y}_a^*|\boldsymbol{x}^*\} = \boldsymbol{x}^{*'}\widehat{\boldsymbol{\beta}}_a + \widehat{\boldsymbol{\nu}}_a, \text{ for } \widehat{\boldsymbol{\nu}}_a \sim N(\mathbf{0}, \widehat{\boldsymbol{\Sigma}}_a), \quad a = 1, 2, ..., n_A$$

The domains of some GEV parameters are restricted. For example, the domain of ξ is restricted to $(-0.5, 0)$ if the distribution is to have finite variance and finite upper end point, and we would expect $\sigma > 0$. More specifically, based on prior studies for a particular application, we might judge it advantageous to restrict the domains of ξ, σ and μ further. Since the linear regression described above assumes a response with unrestricted domain, it is advantageous to transform ξ, σ and μ such that the transformed responses have unrestricted domain for the regression. One suitable transformation is the logistic transformation, which maps a variable y with support (a, b) to y^* with support $(-\infty, \infty)$, given by

$$y^* = \log_e \left[\frac{y-a}{b-a}\right].$$

Examination of estimates for GEV parameters from prior analysis using wide domains (a, b) indicated that it would be reasonable to restrict the domains of σ to $(0, 2)$. Final regression models were therefore estimated using logistically-transformed ξ, logistically-transformed σ with restricted domain, and untransformed unrestricted μ.

The procedure described above ensures that uncertainty from estimation of extreme values models per sea state using finite samples of data and regression of extreme value model parameter estimates onto sea-state summary variables, is propagated to and captured within posterior estimates for the tails of predictive distributions of hourly maximum crest and wave height for unseen sea states.

Appendix E. Preliminary studies of modelling strategy

Section 3 provides a description of the modelling strategy adopted in this work to establish sea-state GEV models. The purpose of this section is to outline the hierarchy of modelling approaches considered before arriving at the composite approach of Section 3.6.

GEV models were established independently per sea state. 12 different models were considered per sea state, arranged in a hierarchy of four model levels L1-4, corresponding to inclusion of data from different measurements sources from Section 3. Level L1 considers only the 105 measurements of hourly maxima. Level L2 additionally considers the numerically pre-selected measurement (NPM), but ignores potential NPM uncertainty and bias (so that the influence of NPM uncertainty and bias can be assessed). Level L3 additionally considers NPM uncertainty and bias. Finally, level L4 also considers hourly areal maxima. Within each level, three different models M1-3 are estimated. In model M1, sample likelihood is based on independent wave gauge measurements. In model M2, sample likelihood is based on the samples means (across wave gauges) at a sequence of quantile levels. Model M3 uses a composite likelihood from M1 and M2. Posterior distributions for 600 sets of GEV parameters were estimated, corresponding to 12 model choices for each of 50 sea states.

The choice of the best combination of level L and model M was made heuristically from inspection of model diagnostic plots, and did not follow a stringent quantitative criterion. In particular it was decided that including areal maximum data was not beneficial, so that L4 could be omitted. However, inclusion of NPM had a clear influence, reducing uncertainties in the estimation of the shape parameter (as might be expected), implying that level L3 was important. It was concluded that the

combination L3-M3 provided the best overall balance, with reasonable fit performance across sea states, and approximately Gaussian-distributed posterior densities. The results shown in this paper are exclusively from L3-M3. Diagnostic plots are available from the authors on request for all 600 models. A schematic of the hierarchy of models considered is given in Figure E.1.

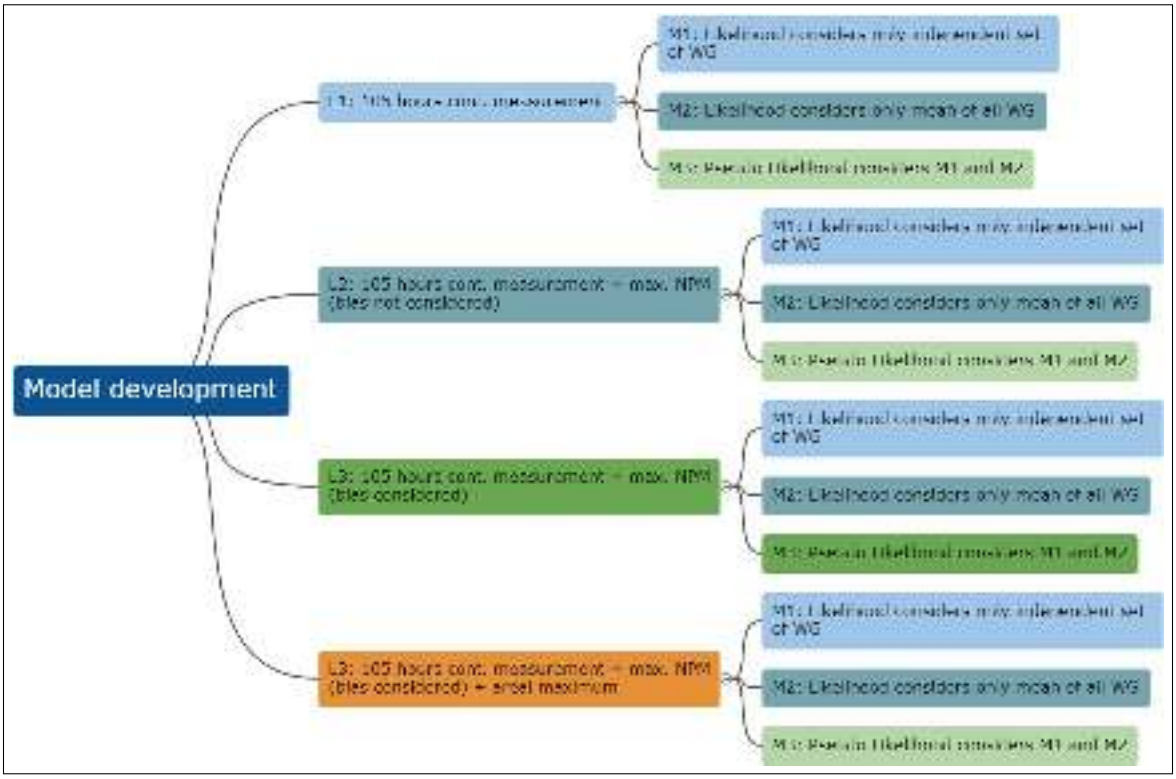


Figure E.1: Schematic of statistical models considered.

Appendix F. MATLAB code

Freely-available MATLAB code for estimation of the distribution of hourly maximum crest and wave height is provided in a GitHub repository. Schubert and Jonathan (2019) gives details. The software generates random hourly maximum crest and wave heights corresponding to a specified period of time, and can be called using either normalized sea state parameters (U, ϵ, θ_{sp},d; leading to normalized crest or wave heights) or non-normalized parameters (H_{m0}, T_P , θ_{sp}, d; leading to crest and wave height in metres) above mean sea level. Optionally, the software provides random realisations of the parameters of the GEV distribution and realisations of the distribution upper end point. The MATLAB function file (editable as an ASCII text file) provides detailed output from the regression analysis, sufficient to generate equivalent functionality in other software applications. A warning is output if input arguments are inconsistent with the domain of sea state parameters used for model building. Time-scale correction (see Section 3.8) is also performed.