Ocean Engineering 165 (2018) 164–183

Contents lists available at ScienceDirect

Ocean Engineering

journal homepage: www.elsevier.com/locate/oceaneng

Long-term distributions of individual wave and crest heights

Ed Mackay*, Lars Johanning

Renewable Energy Research Group, University of Exeter, Penryn, Cornwall, UK

ARTICLE INFO

Keywords:
Wave height
Crest height
Long-term distribution
Return period
Serial correlation

ABSTRACT

This paper considers three types of method for calculating return periods of individual wave and crest heights. The methods considered differ in the assumptions made about serial correlation in wave conditions. The long-term distribution of individual waves is formed under the assumption that either (1) individual waves, (2) the maximum wave height in each sea state or (3) the maximum wave height in each storm are independent events. The three types of method are compared using long time series of synthesised storms, where the return periods of individual wave heights are known. The methods which neglect serial correlation in sea states are shown to produce a positive bias in predicted return values of wave heights. The size of the bias is dependent on the shape of the tail of the distribution of storm peak significant wave height, with longer-tailed distributions resulting in larger biases. It is shown that storm-based methods give accurate predictions of return periods of individual wave heights. In particular, a Monte Carlo storm-based method is recommend for calculating return periods of individual wave and crest heights. Of all the models considered, the Monte Carlo method requires the fewest assumptions about the data, the fewest subjective judgements from the user and is simplest to implement.

1. Introduction

Estimating the long-term statistics of individual wave or crest heights is an important problem in the design of offshore and coastal structures. The long-term statistics of individual waves are dependent on both the long-term distribution of sea states and the short-term distribution of wave heights or crests heights, conditional on sea state. To produce an accurate estimate of the heights of extreme individual waves, information from the long-term and short-term distributions must be combined in an appropriate manner. The approaches that have been proposed for combining these distributions are equally applicable to both wave heights and crest heights, so to avoid referring to both throughout the text, the following discussion is presented in terms of wave heights.

The simplest approach for estimating extreme individual wave heights at a given exceedance probability, e.g. once in 100 years, is to estimate the significant wave height, H_s, at a return period of 100-years then calculate the most probable maximum wave height in that sea state, assuming a duration of somewhere between 3 and 6 h (see e.g. Hogben, 1990; Tucker and Pitt, 2001). There are several problems with this approach. Firstly, the appropriate duration of sea state to use for calculating the most probable maximum wave height is not clear. Secondly, this approach neglects the probability that the largest wave could occur in a sea state other than the 1 in 100-year H_s. This can lead to significant underestimates in predictions of extremes, since there will be several sea states with H_s close to the most severe value, either within the same storm or in separate storms (Carter and Challenor, 1990).

To overcome the problems of the simple approach, various methods have been proposed to combine the long-term and short-term distributions that account for the probability of large waves occurring in any sea state. Battjes (1970) calculated the total number of waves exceeding a level in a given interval and divided this by the total number of waves in the interval to derive an estimate of the long-term distribution of all individual wave heights. Krogstad (1985) proposed a method for calculating the long-term distribution of the maximum wave height in an interval, derived in terms of the distribution of the maximum wave height in each sea state during the interval (see also Prevosto et al., 2000; Krogstad and Barstow, 2004). Various methods have also been proposed for calculating return periods of individual wave heights from the distribution of the maximum wave height in each storm (Jahns and Wheeler, 1973; Ward et al., 1979; Haring and Heideman, 1980; Boccotti, 1986, 2000; Forristall et al., 1991; Tromans and Vanderschuren, 1995; Arena and Pavone, 2006; Fedele and Arena, 2010; Laface and Arena, 2016; Mackay and Johanning, 2018a; b).

Although the assumption is not always explicit in the derivations of each approach, the various methods proposed all have the common feature of calculating the distribution of the maximum wave height in a

* Corresponding author.
E-mail address: e.mackay@exeter.ac.uk (E. Mackay).

https://doi.org/10.1016/j.oceaneng.2018.07.047
Received 22 May 2018; Received in revised form 5 July 2018; Accepted 15 July 2018

random “event”, where the event is either a storm, a sea state, or a single wave. In the last case the maximum wave height in the event is just the individual wave height. Long-term statistics of individual waves are then calculated from the random event distribution, assuming that occurrences of events are independent.

The assumption of independence of these events is not true in general, with records of wave measurements exhibiting serial correlation at multiple scales, with correlation between successive individual wave heights, sea states and storms. In the statistical literature, the effect of serial correlation on estimates of extreme values is often quantified using the extremal index, which can be introduced as follows. Suppose that $X_1, \ldots, X_n$ are a sequence of n independent variables with common distribution function F. In this case, the distribution of the maximum observation is given by $\Pr(\max\{X_1,\ldots,X_n\} < x) = F^n(x)$. If instead $Y_1, \ldots, Y_n$ are a stationary process, also with common distribution function F, but with some level of serial correlation, then subject to certain regularity conditions, it can be shown that $\Pr(\max\{Y_1,\ldots,Y_n\} < x) = F^{\theta n}(x)$, where $\theta \in [0,1]$ is known as the extremal index (see e.g. Coles, 2001). Serial correlation effectively reduces the probability of large observations in a sequence of a given length. Therefore, for processes with $\theta < 1$, assuming that observations are independent will lead to a positive bias in the estimates of extreme values. Some studies have proposed estimating θ explicitly and using this in the estimation of extremes (e.g. Fawcett and Walshaw, 2012). However, obtaining a reliable estimate of θ is difficult in practice and subject to considerable uncertainties (Davis et al., 2013). For oceanographic data it is more common to adopt the peaks-over-threshold scheme which selects events that are approximately independent, ensuring that $\theta \approx 1$ (Jonathan and Ewans, 2013).

The methods proposed by Battjes (1970) and Krogstad (1985) (referred to as the ‘all-wave’ (AW) and ‘sea state maxima’ (SSM) methods respectively from here onwards) and the various storm-based methods make implicit assumptions about independence between events. The AW method assumes that all wave heights are independent, the SSM method assumes that sea state maxima are independent, and storm-based methods assumes that the maximum wave heights in separate storms are independent. Given the serial correlation in wave height time series, the three methods would be expected to give different results, with the AW method producing the highest estimates and storm-based methods producing the lowest estimates.

Forristall (2008) compared estimates of the long-term distribution of individual wave heights from the AW, SSM and storm-based methods. Forristall conducted Monte Carlo simulations of 100,000 years of individual wave heights from synthetic storms, assuming that the time series of H_s in a storm follows a triangular shape with a fixed relationship between the peak H_s and duration of the storm. The duration, D, of the storm was defined to be the time for which $H_s > 0$ and a value of $D = 8H_{s,peak}$ was used, where D is in hours. The zero-crossing period, T_z, was assumed to be constant at 10s and wave heights were assumed to follow a Rayleigh distribution. It was shown that AW and SSM methods produce a positive bias in estimates of the 100-year wave height, consistent with both models neglecting serial-correlation effects. Forristall showed that the storm-based method of Tromans and Vanderschuren (1995) gave the correct return values of individual wave heights when applied to the synthetic triangular storm data with Rayleigh distributed wave heights.

Forristall's study provides a useful insight into the differences between various methods for calculating return periods of individual waves. However, due to the assumptions about the fixed shape of the storms and constant wave period, it is difficult to draw conclusions about the accuracy of storm-based methods when applied to real data. The purpose of this paper is to compare methods for estimating return periods of individual wave heights based on more realistic simulations of synthetic storms, where the wave period varies throughout the storm and the temporal evolution of sea state parameters is based on measured data. In the current work, a block-resampling method is used to generate random time series of realistic storm histories, which are used to compare various methods of estimating return periods of individual wave heights. The block-resampling method divides a time series of measured wave data into discrete blocks consisting of storms where the peak wave heights can be considered approximately independent. The problem of determining time scales over which storm peak wave heights can be considered independent is also discussed in some detail.

The results presented in this paper also have implications for the estimation of extreme load values on marine structures. The “full long-term response analysis” method advocated by some authors (e.g. Sagrilo et al., 2011; Naess and Moan, 2012; Giske et al., 2017) is essentially a method for forming the long-term distribution of the maximum load in each sea state, analogous to the SSM method for wave and crest heights. This method is therefore likely to be subject to the same problems associated with neglecting serial correlation in sea states. Methods for calculating extreme loads which account for serial correlation in sea states have been proposed by other authors (e.g. Forristall et al., 1991; Tromans and Vanderschuren, 1995). However, the focus of this work is on wave and crest heights, and the effect of serial correlation on extreme load values is beyond the scope of the paper.

The paper starts in Section 2 with a brief review of how return periods and return values are defined in the context of the various types of long-term distributions considered. Section 3 presents a short discussion of models for the short-term distribution of wave and crest heights conditional on sea state. Section 4 presents the mathematical derivation of the methods for combining the long-term and short-term distributions, and highlights where various assumptions about independence are made – either implicitly or explicitly. The methods are compared in a simplified example in Section 5, which isolates the effects of serial correlation in sea states. The methods are then applied to measured data in Section 6, providing a quantitative comparison of the effect the various assumptions made in each method in a real situation. The accuracy of the methods is compared in Section 7 using Monte Carlo simulations of synthetic storms. Finally, conclusions are presented in Section 8.

2. Return periods & return values

The methods for estimating the long-term distribution of individual wave heights considered in this paper are used to define return periods and return values in slightly different ways. It is therefore useful to review how return periods and return values are defined in each context. Return values are defined in terms of the distribution of the maximum wave height in a year. We denote the probability that the maximum individual wave height, H_{max}, does not exceed h in any year selected at random as $\Pr(H_{max} \leq h|\ 1\ \text{year})$. The T-year return value of individual wave height, H_T, is then defined as the value which has an exceedance probability $1/T$ in any year:

$$\Pr(H_{max} > H_T|\ 1\ \text{year}) = \frac{1}{T}, \quad T > 1 \tag{1}$$

The duration T is referred to as the return period and is the average period between exceedances of H_T. Over the last few decades, the peaks-over-threshold (POT) method has gained popularity over the annual-maxima method (see Jonathan and Ewans, 2013, for a review of the use of POT in an oceanographic context). In this method, the distribution of the annual maximum is not estimated explicitly. Instead, the distribution of independent threshold exceedances is estimated. If each independent threshold exceedance is described as an ‘event’ then return values of wave heights can be defined in terms of the distribution of the maximum wave height in a random event, as the solution of:

$$\Pr(H_{max} > H_T|\text{event}) = \frac{1}{Tm}, \quad T > \frac{1}{m} \tag{2}$$

where m is the expected number of events per year (see e.g. Coles, 2001). In the present context, the event is either a storm, sea state or

individual wave.

The definitions (1) and (2) lead to slightly different return values for low return periods, but the differences become negligible at larger return periods. This can be seen as follows. The distribution of the maximum wave height in a year can be calculated by assuming independence of maxima in each event as:

$$\Pr(H_{max} \leq h|\ 1\text{ year}) = \Pr(H_{max} \leq h|\text{event})^m, \tag{3}$$

Substituting (3) into (1) and using a Taylor series expansion gives:

$$\Pr(H_{max} \leq H_T|\text{event}) = \left(1 - \frac{1}{T}\right)^{1/m} = 1 - \frac{1}{Tm} + O\left(\frac{1}{T^2}\right) \tag{4}$$

So, the two definitions are equivalent for large T. In practice the difference is negligible for $T \geq 10$, so both definitions will be used in the following discussions, depending on which is more appropriate.

3. Short-term distribution of wave heights conditional on sea state

A sea state, σ, is defined to be a period of time in which the wave conditions can be considered approximately stationary, with the duration normally defined in the range 30 min to 3 h. A sea state is defined in terms of the wave spectrum and summarised in terms of spectral parameters, such as significant wave height $H_s = 4\sqrt{m_0}$, zero up-crossing period $T_z = \sqrt{m_0/m_2}$, mean period $T_m = m_0/m_1$, where $m_n = \int_0^\infty f^n S(f)\mathrm{d}f$ is the n^{th} moment of the wave frequency spectrum, $S(f)$.

The short-term distribution of wave heights, H, conditional on sea state is denoted

$$\Pr(H \leq h|\sigma). \tag{5}$$

Models for (5) in terms of sea state parameters and water depth is the subject of ongoing research and it is beyond the scope of this work to provide a review of the extensive literature on this topic. The focus of this work is on how to combine the short-term distribution (5) with the long-term distribution of sea states. The methods described are applicable to any model for the short-term distribution. The short-term distribution is primarily dependent on H_s, but is also affected by the water depth, wave steepness, spectral bandwidth, directional spread and currents. In the following, it is assumed that the model for the short-term distribution adequately captures these effects.

4. Long-term distributions

4.1. Long-term distribution of all wave heights

Battjes (1970) proposed a method for calculating return values of individual wave heights in terms of the long-term distribution of all wave heights. The probability that any wave selected at random exceeds height h is calculated as the ratio of the expected number of waves exceeding h per year to the expected number of waves per year. The expected number of waves exceeding h in a sea state is

$$M_\sigma(h) = N(T_z)\Pr(H > h|\sigma), \tag{6}$$

where $N(T_z) = D_\sigma/T_z$ is the expected number of waves in a sea state, D_σ is the duration of the sea state, normally assumed to be somewhere in the region 1–6 h, and T_z is the zero up-crossing period. The expected number of waves exceeding h in a year, denoted $M_{yr}(h)$, is calculated by integrating $M_\sigma(h)$ over the probability of occurrence of sea states and multiplying by the number of sea states per year, NPY:

$$M_{yr}(h) = NPY\int_0^\infty\int_0^\infty p(H_s, T_z)N(T_z)\Pr(H > h|\sigma)dH_s dT_z. \tag{7}$$

The occurrence of sea states is specified in terms of $p(H_s, T_z)$, the joint probability density function of H_s and T_z. This implicitly assumes that the short-term distribution is only influenced by H_s and T_z, which, as discussed in Section 3, is not necessarily true. The effect of other sea state parameters on the short-term distribution could be captured by using the joint density function of multiple parameters, however, this quickly becomes very complicated as the number of variables is increased. In some applications of the AW method, the joint density function is used (e.g. Hagen et al., 2017). However, in most applications (7) is simplified as follows. To avoid having to model joint distribution, the joint density function is written as $p(H_s, T_z) = p(H_s)p(T_z|H_s)$, and the mean number of waves for a given H_s is defined as

$$\overline{N}(H_s) = \int_0^\infty p(T_z|H_s)N(T_z)\mathrm{d}T_z. \tag{8}$$

Substituting this back into (7) gives:

$$M_{yr}(h) = NPY\int_0^\infty p(H_s)\overline{N}(H_s)\Pr(H > h|\sigma(H_s))\mathrm{d}H_s. \tag{9}$$

In this approach, all the sea state variables that influence the short-term distribution must be modelled as functions of H_s.

In the original AW method proposed by Battjes, the expected number of waves in any sea state selected at random is calculated as

$$\overline{N_r} = \int_0^\infty p(H_s)\overline{N}(H_s)dH_s. \tag{10}$$

The expected number of waves per year is therefore $\overline{N_r}.\ NPY$ and the probability that any wave selected at random exceeds height h is

$$\begin{aligned}\Pr(H > h|\text{any wave}) &= \frac{M_{yr}(h)}{\overline{N_r}.\ NPY}\\ &= \frac{1}{\overline{N_r}}\int_0^\infty p(H_s)\overline{N}(H_s)\Pr(H > h|\sigma(H_s))\mathrm{d}H_s\end{aligned} \tag{11}$$

The distribution of the annual maximum wave height is then calculated as:

$$\Pr(H_{max} \leq h|\ 1\text{ year}) = \Pr(H \leq h|\text{any wave})^{\overline{N_r}\cdot NPY} \tag{12}$$

This calculation makes the implicit assumption that individual wave heights are independent, which, as discussed in the introduction, is not true.

Tucker (1989) pointed out that $\overline{N_r}$ is sensitive to the number of waves in low sea states and that it seems counter-intuitive that this should influence the extreme values. Tucker noted that the T-year return value of individual wave height is the value of h for which the expected number of waves exceeding h per year is $M_{yr}(h) = 1/T$. The annual distribution can therefore be calculated directly from (1) and (9) as:

$$\begin{aligned}\Pr(H_{max} \leq h|\ 1\text{ year}) &= 1 - M_{yr}(h)\\ &= 1 - NPY\int_0^\infty \Pr(H > h|\sigma(H_s))p(H_s)\overline{N}(H_s)\mathrm{d}H_s.\end{aligned} \tag{13}$$

Note that this definition is only valid for $M_{yr}(h) < 1$. Although it is not immediately obvious from (13), Tucker's expression for the annual distribution implicitly assumes that individual wave heights are uncorrelated. Expanding the original Battjes expression (11) as a Taylor series, we have

$$\begin{aligned}\Pr(H_{max} \leq h|\ 1\text{ year}) &= \Pr(H \leq h|\text{any wave})^{\overline{N_r}\cdot NPY}\\ &= \left(1 - \frac{M_{yr}(h)}{\overline{N_r}.\ NPY}\right)^{\overline{N_r}\cdot NPY}\\ &= 1 - M_{yr}(h) + O((M_{yr}(h))^2)\end{aligned} \tag{14}$$

In Tucker's definition $M_{yr}(h) = 1/T$, so for large T the terms of order $(M_{yr}(h))^2$ are negligible and the two expressions are equivalent. Therefore, both Battjes' and Tucker's expressions for the distribution of the maximum wave height in a year can be thought of as being formed under the assumption that there is no serial correlation in individual

wave heights. This assumption is discussed further in the following section.

4.2. Long-term distribution of sea state maxima

The long-term distribution of the maximum wave height in a sea state can be calculated in two ways, with one approach based on an average over time and the other based on an average over the probability of occurrence of sea states. Both methods give equivalent results, but it is not immediately obvious to see this from looking at the expressions. The temporal average method will be presented first and it will then be shown how it is equivalent to the probabilistic method.

The distribution of the maximum wave in a sea state is usually calculated by assuming that successive wave heights are independent, so that

$$\Pr(H_{max} \le h|\sigma) = \Pr(H \le h|\sigma)^N. \tag{15}$$

where $N = D_\sigma/T_z$ is the expected number of waves in the sea state, as defined in the previous section. In reality, consecutive wave and crest heights are correlated, with the largest waves occurring in groups. Naess (1985) showed that for linear waves from a JONSWAP spectrum, the effect of ignoring correlation between successive waves has a negligible influence on the distribution of the maximum wave height. Fedele (2016) considered the effect of correlation between successive wave crests using the linear model of Fedele (2005) and also concluded that correlation could be neglected for typical broadbanded waves from a JONSWAP spectrum. Fedele (2016) also noted that for the case of nonlinear waves, since the nonlinear harmonics are phase-locked to the linear components, the dependence between successive crests is unlikely to be affected by nonlinearities. Janssen (2015) suggested that it is more appropriate to characterise N as a function of the number of wave groups in the record, as it is more plausible that the maximum wave height in each group is independent than it is that successive wave heights are independent. However, this idea has not been pursued in this work.

The maximum wave heights in successive sea states can be considered independent, in the sense that the maximum height is dependent only on the sea state parameters and not the maximum height in adjacent sea states. The distribution of the maximum wave height in the time interval $[0,\ A]$ can therefore be written as:

$$\begin{aligned}\Pr(H_{max} \le h|[0,A]) &= \textstyle\prod_{i=1}^{k} \Pr(H_{max} \le h|\sigma(\tau_i)) \\ &= \textstyle\prod_{i=1}^{k} \Pr(H \le h|\sigma(\tau_i))^{D_\sigma/T_z(\tau_i)} \\ &= \exp\left(\textstyle\sum_{i=1}^{k} \frac{D_\sigma}{T_z(\tau_i)} \ln[\Pr(H \le h|\sigma(\tau_i))]\right)\end{aligned} \tag{16}$$

where $\tau_i \in [0, A]$. As the duration of each sea state, D_σ, tends to zero, the summation in (16) can be expressed as an integral (Borgman, 1973):

$$\Pr(H_{max} \le h|[0,A]) = \exp\left(\int_0^A \frac{1}{T_z(t)} \ln[\Pr(H \le h|\sigma(t))]\,\mathrm{d}t\right). \tag{17}$$

Equations (16) and (17) give expressions for the distribution of the maximum wave height in an interval $[0,\ A]$ if the time series of sea states is known, buy they contain no information about the long-term distribution of sea states. By applying the ergodic assumption (Naess, 1984) that, on average, the length of time that sea states are in the interval $[H_s,\ H_s + dH_s]$ during time $[0, A]$ is $\mathrm{d}t = Ap(H_s)\mathrm{d}H_s$, the distribution (17) can be rewritten as (Krogstad, 1985):

$$\Pr(H_{max} \le h|[0,A]) = \exp\left[A \int_0^\infty \frac{1}{T_z(H_s)} p(H_s) \ln[\Pr(H \le h|\sigma(H_s))]\,\mathrm{d}H_s\right] \tag{18}$$

Note that in moving from the integral over time to the integral over H_s, it must be assumed that T_z and other sea state parameters can be modelled as taking a mean value conditional on H_s, without affecting the calculation.

The application of the ergodic assumption in passing from (17) to (18) effectively counts the proportion of sea states which exceed a threshold level but does not account for temporal clustering of sea states in storms. To make this explicit, if $A = D_\sigma$, then, recalling $\overline{N}(H_s) = D_\sigma/T_z(H_s)$ and using (15), gives

$$\Pr(H_{max} \le h|\text{any sea state}) = \exp\left[\int_0^\infty p(H_s) \ln[\Pr(H_{max} \le h|\sigma(H_s))]\,\mathrm{d}H_s\right] \tag{19}$$

Putting $A = NPY{\cdot}D_\sigma = 1$ year gives

$$\begin{aligned}\Pr(H_{max} \le h|\ 1\text{ year}) &= \exp\left[NPY \int_0^\infty p(H_s) \ln[\Pr(H_{max} \le h|\sigma(H_s))]\,\mathrm{d}H_s\right] \\ &= \Pr(H_{max} \le h|\text{any sea state})^{NPY}\end{aligned} \tag{20}$$

So, in Krogstad's formulation (18), the annual distribution is identical to that formed under the assumption that the maxima in each sea state are independent.

A distinction should be made between the type of independence assumed to derive (16) and the type of independence assumed in (18). In the derivation of (16) it is assumed that the maximum wave height in each sea state is dependent only on the sea state parameters and not on the maximum height in adjacent sea states. In this expression, the values of the sea state parameters are assumed to be known, such as would be the case for a measured time series of sea states in a storm. In (18) the sequence of sea states over the year are not known, only their marginal distribution is known. By replacing the integral over a known time series of sea states (17) with an integral over the marginal distribution of sea states (18), the implicit assumption is made that there is no serial correlation in sea states. This implicit assumption is made clear by equations (19) and (20). As noted in the introduction, serial correlation reduces the probability of large observations in a sequence of a given length. This effect is not accounted for in the model of Krogstad (1985).

4.2.1. Connection to other formulations

Krogstad's model can be shown to be equivalent to other expressions for the long-term distribution of sea state maxima. Sagrilo et al. (2011) noted that by using Taylor series expansions for the exponential and logarithm terms, Krogstad's expression (18) can be shown to be approximately equal to the probabilistic average. Using the Taylor series expansion for the logarithm $\ln(x) = x - 1 + O(x^2)$ and noting that $\int_0^\infty p(H_s)\mathrm{d}H_s = 1$ gives

$$\Pr(H_{max} \le h|\text{any sea state}) \approx \exp\left[\int_0^\infty p(H_s) \Pr(H_{max} \le h|\sigma(H_s))\mathrm{d}H_s - 1\right] \tag{21}$$

Note that the approximation $\ln(x) \approx x - 1$ is an identity when $x = 1$, and for the values of h relevant to extremes, $\Pr(H_{max} \le h|\sigma(H_s)) \approx 1$, so this approximation of the logarithm term is accurate for this purpose. Similarly, using the Taylor series expansion $\exp(x) \approx 1 + x$ (which is an identity for $x = 0$ and a good approximation in this case) gives

$$\Pr(H_{max} \le h|\text{any sea state}) \approx \int_0^\infty p(H_s) \Pr(H_{max} \le h|\sigma(H_s))\mathrm{d}H_s \tag{22}$$

This expression, given by Tucker and Pitt (2001, Section 6.4.2), is the population mean, whereas Krogstad's expression (18) is the ergodic mean. These two expressions differ slightly for lower values of h, but are almost identical for the higher values of h relevant to extremes. However, Krogstad's expression is more accurate to compute as it uses the logarithm of $\Pr(H_{max} \le h|\sigma(H_s))$.

Krogstad's model can also be related to the AW model as follows. Taking $A = NPY{\cdot}D_\sigma = 1$ year, in (18) and using the same linear expansions for the exponential and logarithm terms gives:

$$\Pr(H_{max} \le h|\,1\text{ year}) \approx 1 + NPY \int_0^\infty p(H_s)\overline{N}(H_s)(\Pr(H \le h|H_s) - 1)\mathrm{d}H_s, \tag{23}$$

which is identical to (13). It is somewhat surprising that the AW and SSM methods result in approximately equivalent expressions for the distribution of the annual maximum wave height. However, if we consider that both models assume that there is no serial correlation between individual wave heights or sea states, then the reason for the equivalence becomes apparent. In the examples considered in Sections 5–7 the AW and SSM methods show differences for return periods less than 10 years but give nearly identical results for higher return periods.

4.3. Long-term distribution of storm maxima

A storm can be thought of as a sequence of sea states, where wave heights increase to a peak value and then decrease again. The distribution of the maximum wave height in a storm can be calculated using (16) where the time interval $[0, A]$ is defined to be the duration of the storm. The process of partitioning a time series into separate storms is analogous to defining "declustering" criteria in the peaks-over-threshold (POT) method (see Coles, 2001). The criteria used to define separate storms typically state that the time between the peak H_s of two adjacent storms must be larger than some minimum value. In some models an additional criterion is used, requiring that the minimum H_s between two adjacent peaks must be less than some multiple of the lower of the peak. The criteria for defining what constitutes separate storms can be defined more rigorously by considering the correlation between successive storm peak wave heights. This will be discussed in more detail in Section 6.2. For now, it is assumed that the time series has been partitioned into discrete blocks, constituting storms where the peak wave heights can be considered independent.

The distribution of the maximum wave height in any storm selected at random, denoted $\Pr(H_{max} \le h|\text{any storm})$, can be estimated in two ways. In the first method, the distributions of the maximum wave height in the measured storms can be parameterised in some way, in terms of a statistically equivalent storm, and the long-term distribution of storm parameters can be combined with the short-term distribution for the storm, in a manner that is analogous to the SSM and AW methods. Alternatively, $\Pr(H_{max} \le h|\text{any storm})$ can be estimated directly from the data using a Monte Carlo method. The two methods are described in Sections 4.3.1 and 4.3.2.

4.3.1. The equivalent storm method

The distribution of the maximum wave height in a measured storm will be denoted $\Pr(H_{max} \le h|\text{MS})$ and the distribution for the equivalent storm will be denoted $\Pr(H_{max} \le h|\text{ES})$. Two types of approach for parameterising $\Pr(H_{max} \le h|\text{MS})$ have been proposed. One approach is to model the temporal evolution of sea states in a storm using some simplified geometric form, such as a triangle (Boccotti, 1986, 2000; Arena and Pavone, 2006; Laface et al., 2017), trapezoid (Martin-Soldevilla et al., 2015), parabola (Tucker and Pitt, 2001, Section 6.5.4), power law (Fedele and Arena, 2010; Fedele, 2012; Arena et al., 2014) or exponential law (Laface and Arena, 2016). In most of the approaches, the parameters of the equivalent storm are fitted using an iterative procedure so that the distribution of the maximum wave height in the equivalent storm is matched as closely as possible to the measured storm (although Fedele (2012) provided an explicit formula for the duration of the equivalent storm). This approach is reasonably effective because the order of sea states in the storm does not affect the distribution of the maximum wave height. Therefore, the product in (16) can be re-ordered into a monotonically increasing series of sea states (in terms of the values of H_s), for which a linear, power or exponential law is a reasonable fit. The drawback of this approach is that only the temporal evolution of H_s is modelled. It is therefore necessary to establish models for the mean values of T_z and other sea state parameters as a function of H_s, in the same way as for the AW and SSM models.

The other approach is to model $\Pr(H_{max} \le h|\text{MS})$ directly and define the equivalent storm as a statistical distribution rather than a time series of H_s. This approach was adopted by Tromans and Vanderschuren (1995), who assumed that the square of the maximum wave height in the storm followed a Gumbel distribution. Mackay and Johanning (2018a) developed this idea and used the generalised extreme value distribution (GEV) to model the distribution of the maximum wave height in a storm. They demonstrated that using the GEV to model $\Pr(H_{max} \le h|\text{MS})$ improves the goodness-of-fit by an order of magnitude compared to both temporal evolution methods and the Tromans and Vanderschuren (TV) method. Moreover, the GEV model was shown to be more robust than both the TV and temporal evolution methods to uncertainties in the models for the distribution of storm parameters. The GEV model also results in a much simpler form for $\Pr(H_{max} \le h|\text{ES})$ than temporal evolution methods.

The TV method is recommend in several metocean design standards (ISO, 2015; DNV GL, 2017) and will therefore be considered as a reference model in comparison to the GEV model. The temporal evolution methods will not be considered further.

In the TV method the distribution of the maximum wave height in the equivalent storm is defined as

$$\Pr(H_{max} \le h|ES) = \exp\left(-\exp\left(-\ln(N)\left(\left(\frac{h}{H_{mp}}\right)^2 - 1\right)\right)\right), \tag{24}$$

where N is the number of waves in the storm and H_{mp} is the most probable maximum wave height in the storm (i.e. the mode of $\Pr(H_{max} \le h|\text{MS})$). Tromans and Vanderschuren (1995) noted that $\ln(N) \approx 8$ for Northern European winter storms, but did not state explicitly how they estimated $\ln(N)$. Mackay and Johanning (2018a) considered two options for calculating $\ln(N)$. In the first method, $\ln(N)$ was calculated in terms of the moments of $\Pr(H_{max} \le h|\text{MS})$, as $\ln(N) = \pi H_{mp}^2/(\text{std}(H_{max}^2)\sqrt{6})$. In the second method the value of $\ln(N)$ was determined numerically by finding the value that minimises the difference between $\Pr(H_{max} \le h|\text{MS})$ and $\Pr(H_{max} \le h|\text{ES})$, quantified in terms of the Cramér–von Mises goodness-of-fit parameter, ω:

$$\omega^2 = \int_0^\infty [\Pr(H_{max} \le h|\text{MS}) - \Pr(H_{max} \le h|\text{ES})]^2\mathrm{d}h \tag{25}$$

It was demonstrated that the second method gave a better fit to the data. However, the moment based estimators are recommended in DNV GL (2017) and will therefore be used here.

Mackay and Johanning (2018a) defined the distribution of the maximum wave height in the equivalent storm in terms of the GEV as:

$$\Pr(H_{max} \le h|\text{ES}) = \begin{cases} \exp\left(-\left(1 + k\left(\frac{h-a}{b}\right)\right)^{-\frac{1}{k}}\right) & \text{for } k \ne 0 \\ \exp\left(-\exp\left(-\left(\frac{h-a}{b}\right)\right)\right) & \text{for } k = 0 \end{cases} \tag{26}$$

where a, b and k are the location, scale and shape parameters, respectively. The GEV can be fitted to (16) by finding the parameters that minimise ω. The first guess for the parameters of the GEV is defined by assuming $k = 0$, which from the properties of the GEV, then implies $a = H_{mp}$ and $b = \sqrt{6}\ \text{std}(H_{max}|\text{MS})/\pi$. The parameters that minimise ω are then found using a simplex search technique (Lagarias et al., 1998). Mackay and Johanning (2018a) demonstrated that the GEV provides an excellent fit to $\Pr(H_{max} \le h|\text{MS})$, with a bias of less than 0.2% in quantiles up to an exceedance probability of 10^{-3}, meaning that there is very little inaccuracy introduced by parameterising the storm using the

GEV, compared to modelling all sea state variables as directly related to H_s.

Once each measured storm has been fitted with an equivalent storm, the distribution of the maximum wave height in a random storm can be calculated using the law of total probability, by integrating $\Pr(H_{max} \leq h|\text{ES})$ over the long-term joint density function of storm parameters, in a manner analogous to the SSM model (22). In the case of the TV model, $\ln(N)$ is assumed to be constant and the mean value over all storms is used. For the GEV model Mackay and Johanning (2018a) showed that the triple integral over the joint density function $p(a, b, k)$ could be reduced to a single integral over $p(a)$ using the mean value of k and a linear model $b = A + Ba$ established using regression. It was also shown that the GEV location parameter a is almost identical to H_{mp} (since the mean value of k was close to zero). Therefore, the distribution of the maximum wave height in a random storm can be written in the same way for both the TV and GEV models as:

$$\Pr(H_{max} \leq h|\text{any storm}) = \int p(H_{mp})\Pr(H_{max} \leq h|\text{ES})\, \mathrm{d}H_{mp}. \tag{27}$$

The similarity between (27) the SSM model (22) is clear. However, one advantage of this form is that no model is required for how other sea state parameters vary with H_s.

The density function $p(H_{mp})$ is modelled using the POT method, with the generalised Pareto distribution (GPD) fitted to exceedances of H_{mp} above some high threshold. The fitting of the GPD and the selection of the threshold is discussed in Section 6.3.

4.3.2. The Monte Carlo method

Monte Carlo methods have been applied to the estimation of extreme load statistics for various offshore structures. A Monte Carlo method for estimating the short-term mean upcrossing rate of high levels (which can be used in place of the short-term distribution) was presented by Naess et al. (2007a, b). This short-term mean upcrossing rate was then used in combination with a model for the long-term extremes that is equivalent to the SSM method (see e.g. Sagrilo et al., 2011), which ignores the effects of serial correlation in sea states. Brown et al. (2017) used a Monte Carlo approach to estimate the extreme response of an FPSO in squall conditions. Brown et al. (2017) conducted numerical simulations to create a database of the short-term extreme response in squall conditions, which was then resampled in order to estimate the long-term extreme response.

Mackay and Johanning (2018b) proposed a storm-based Monte Carlo (MC) method for calculating return periods of individual wave and crest heights. The motivating idea for the method is that if a long time series of individual wave heights were available, then $\Pr(H_{max} \leq h|\text{any storm})$ could be estimated directly from the data using a standard POT approach, without the need to calculate $\Pr(H_{max} \leq h|\text{MS})$ or fit equivalent storms. In the MC method, a random realisation of the maximum wave height in each sea state is simulated from the metocean parameter time series and the GPD is fitted to storm peak wave heights exceeding some threshold. This gives an estimate of $\Pr(H_{max} \leq h|\text{any storm})$, but subject to a high sampling variability, due to the random variability in the storm peak wave heights. To account for the random variability in the simulated wave heights, this process is repeated n times and the parameters of the fitted GPDs are averaged over the n trials to obtain a stable estimate of $\Pr(H_{max} \leq h|\text{any storm})$. Mackay and Johanning (2018b) showed that using $n = 1000$ is sufficient to obtain a stable estimate and noted that the computational time for a 30-year dataset is around 1 min on a standard contemporary laptop.

The random maxima in each sea state are simulated by generating a sequence of uniformly distributed random variables $P_i \in [0,1]$, $i = 1, \ldots, M$, where M is the number of sea states in the metocean dataset. The random maximum wave height in each sea state, $H_{max,i}$, is then obtained from (15) as the solution of

$$P_i^{1/N_i} = \Pr(H \leq H_{max,i}|\sigma_i), \tag{28}$$

where $N_i = D_\sigma / T_{z,i}$, as before. Many models for $\Pr(H \leq h|\sigma)$ can be inverted analytically to obtain a closed form solution for the random maximum wave heights. However, even if this is not possible, interpolation of $\Pr(H \leq h|\sigma)$ to the desired probability gives a simple means of for calculating the random maximum wave heights. The difference between the MC method of Brown et al. (2017) and that of Mackay and Johanning (2018b) is that in the latter study a new set of random maximum wave heights are generated on each trial via (28), whereas in the former study a pre-determined set of load maxima are resampled.

A threshold must be selected for fitting the GPD to storm peak wave heights. Selecting a threshold for each iteration of the MC method would be impractical. Mackay and Johanning (2018b) recommended that the threshold is selected prior to the MC simulations, using the median value of the maximum wave height in each sea state, denoted H_{med}, which is calculated from (28) by setting $P_i = 0.5$ in each sea state. The median values can then be declustered to select the peak values in each storm and the threshold can be selected, following the procedures outlined in Section 6.3. The choice of declustering criteria is discussed in Section 6.2.

The MC method could also be applied to estimate the SSM and AW distributions directly from the data, removing the need to model the variation of the short-term distribution parameters with H_s. However, for the AW method, simulating all waves for each sea state would be very computationally intensive and these ideas are not pursued further here.

5. Comparison in simplified example

There are two main differences between the AW and SSM methods and storm-based methods. Firstly, both the AW and SSM methods make the implicit assumption that there is no serial correlation in sea states. Secondly, the AW and SSM methods both require that models are established for the mean values of the short-term distribution parameters as a function of H_s. Before comparing the methods in a realistic setting where both of these differences become relevant, it is useful to use a simplified example to isolate the effect of serial correlation in sea states.

In this simplified example it is assumed that individual wave heights follow a Rayleigh distribution:

$$\Pr(H \leq h|H_s) = 1 - \exp\left(-2\left(\frac{h}{H_s}\right)^2\right) \tag{29}$$

It is also assumed that the wave period is constant at $T_z = 10$s, so that the distribution of the maximum wave height in a sea state is dependent only on H_s. Suppose that the time series of H_s is composed of discrete, independent storms, each comprising a fixed number, N, consecutive sea states. The number of storms per year is $M = NPY/N$, where NPY is the number of sea states per year. We further assume that H_s is constant during each storm, so that the marginal distribution of H_s is equal to the distribution of storm peak H_s. For the sea state-based model, the distribution of the maximum wave height in a year is calculated from (22) as

$$\Pr(H_{max} \leq h|1\text{ year}) = \left[\int_0^\infty p(H_s)\Pr(H_{max} \leq h|H_s)\mathrm{d}H_s\right]^{NM} \tag{30}$$

For the storm-based model, the distribution of the maximum wave height in a year can be calculated by integrating the distribution of the maximum wave height in a storm, conditional on storm peak H_s, over the distribution of storm peak H_s (which is equal to the marginal distribution of H_s in this example), which gives

$$\Pr(H_{max} \leq h|1\text{ year}) = \left[\int_0^\infty p(H_s)[\Pr(H_{max} \leq h|H_s)]^N \mathrm{d}H_s\right]^M \tag{31}$$

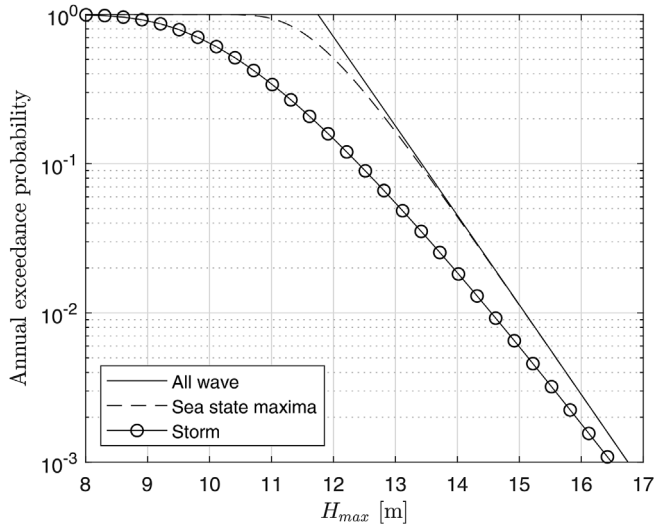

Fig. 1. Distribution of the maximum wave height in a year from the all-wave, sea-state-maxima and storm methods for hypothetical case of independent rectangular storms.

Comparison of (30) and (31) shows that the assumptions about independence of sea states and storms lead to differing expressions for the distribution of the annual maximum wave height. To quantify the difference between the two models, we assume that sea states have a 1-h duration, storms consist of $N = 100$ sea states and there are $NPY = 8760$ sea states per year. We also assume that H_s follows a Weibull distribution:

$$\Pr(H_s \leq x) = 1 - \exp\left(-\left(\frac{x}{\alpha}\right)^{\beta}\right) \tag{32}$$

where $\alpha = \beta = 2$. Fig. 1 shows the distributions of the annual maximum wave heights from the sea state (30) and storm-based models (31) for this simple example. Results from the AW model (13) are also included for comparison. As expected from (23) the AW and SSM models give almost identical results for lower exceedance probabilities. The AW and SSM method predicts larger wave heights than the storm method at a given exceedance level (i.e. larger return values), but the difference decreases at lower exceedance probabilities (i.e. larger return periods). In this example, the difference is due only to the differing assumptions about independence of sea states and storms. The results are consistent with the observation made in the introduction, that neglecting serial correlation results in a positive bias in estimates of extreme values. Strictly speaking, we have only shown that the sea state-based model gives larger return values, rather than a positive bias. However, the simulation studies presented in Section 7 will show that it is in fact a positive bias. The assumptions made in this example are clearly unrealistic. In the following two sections it will be shown that biases observed in this artificial example also occur under realistic conditions.

6. Application to measured data

In this section the methods used to fit the models to the data are discussed. Two sets of measured wave data are used to illustrate the procedures. The datasets used are described in Section 6.1. The selection of declustering criteria for the storm-based methods is discussed in Section 6.2. Parameter estimation and threshold selection for the GPD are discussed in Section 6.3. Finally, a comparison of return values of individual crest heights calculated using each method is presented in Section 6.4.

6.1. Datasets

The datasets used are long-term records from wave buoys in two distinct wave climate regimes and are the same as those used in Mackay and Johanning (2018a). The first dataset, referred to as Site A, is from NDBC buoy number 44025, located off the coast of Long Island, New York, in a water depth of 36 m. The dataset consists of 25 years of hourly records of wave spectra over the period April 1991–December 2016. The second dataset, referred to as Site B, is from NDBC buoy number 46014, located off the coast of Northern California in a water depth of 256 m. The dataset consists of 35 years of hourly records of wave spectra over the period April 1981–December 2016. Scatterplots of H_s against T_z for the two datasets are shown in Fig. 2. Both datasets have similar maximum observed H_s, with a maximum H_s at Site A of 9.64 m and a maximum at Site B of 10.34 m. In the shallower water site, the maximum values of H_s occur for the steepest seas with $0.06 < s_z < 0.07$ (where $s_z = 2\pi H_s/gT_z^2$), whilst at the deeper site the extreme values of H_s occur for a wider range of steepness. The range of T_z observed at the Californian site is also much larger than at the New York site.

In the examples presented in Sections 6.4 and 7, the Forristall (2000) model for the short-term distribution of crest heights in directionally spread seas, will be used to illustrate the application of the long-term models. The Forristall model assumes that crest heights, C, follow a Weibull distribution:

$$\Pr(C \leq \eta|\sigma) = 1 - \exp\left(-\left(\frac{\eta}{\alpha H_s}\right)^{\beta}\right). \tag{33}$$

The parameters α and β are defined in terms of the significant

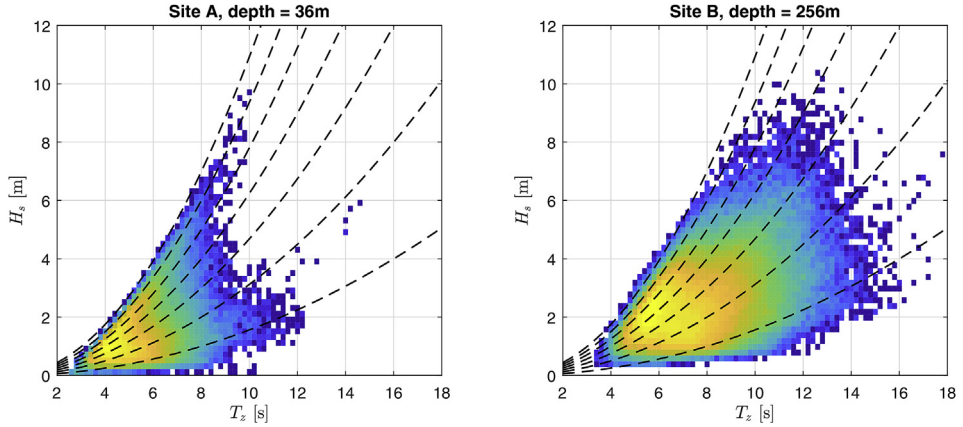

Fig. 2. Scatter plots of H_s against T_z for the two datasets used in this study. Colour indicates density of occurrence. Black dashed lines indicate s_z from 0.01 to 0.07 at intervals of 0.01. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

steepness, s_m, and Ursell number, U_{rs}, defined as

$$s_m = \frac{2\pi H_s}{g T_m^2}, \tag{34}$$

$$U_{rs} = \frac{H_s}{k_m^2 d^3}, \tag{35}$$

where k_m is the finite depth wave number corresponding to T_m and d is the water depth. The distribution parameters are given by

$$\alpha = 0.3536 + 0.2568 s_m + 0.0800 U_{rs} \tag{36}$$

$$\beta = 2 - 1.7912 s_m - 0.5302 U_{rs} + 0.2824 U_{rs}^2 \tag{37}$$

The AW and SSM methods both require models for the mean values of α, β and T_z as a functions of H_s. Mackay and Johanning (2018a) fitted linear models for the variation of the mean values of these parameters as a function of H_s, which are used here as well. The reader is referred to Mackay and Johanning (2018a) for details.

6.2. Selection of declustering criteria

The objective of this article is to establish the accuracy of storm-based methods for estimating return periods of individual wave heights and contrast this to methods which neglect serial correlation in sea states. It is therefore important to consider the correlation structure of sea state time series in some detail and to establish a rigorous method of estimating a minimum separation time between local maxima in the wave height time series, for which adjacent maxima can be considered statistically independent.

Most statistical methods developed to quantify dependence between extreme events assume that the time series is stationary. Time series of H_s exhibit a clear seasonal variation in storm peak wave heights, which introduces a long-range dependence structure between extreme events and a non-stationarity in the time series. The long-range dependence due to seasonal variability is deterministic in origin, with the variation having a period of exactly one year. The variation can therefore be filtered out and the normalised, stationary time series can be examined for shorter-range dependence of extreme events using established statistical methods.

The autocorrelation in storm peak H_s is shown in Fig. 3 for both datasets. In this example, storm peaks are defined as the maxima within a 5-day moving window. The lag shown is the number of storms separating adjacent peaks. The median time between storm peaks is 9.79 days for Site A and 9.86 days for Site B. The peaks in autocorrelation function occur around lags of ~36 and ~72 storms, which corresponds to the annual signal. The pattern in the autocorrelation in storm peak H_s is very similar for these two locations, which is due to both datasets being located at similar latitudes where the level of seasonal variability is similar. The slight difference between the autocorrelation functions for the two buoys is likely to be due to differences in the distribution of storm lengths and the effects of missing data in time series, which has not been accounted for in this simple analysis.

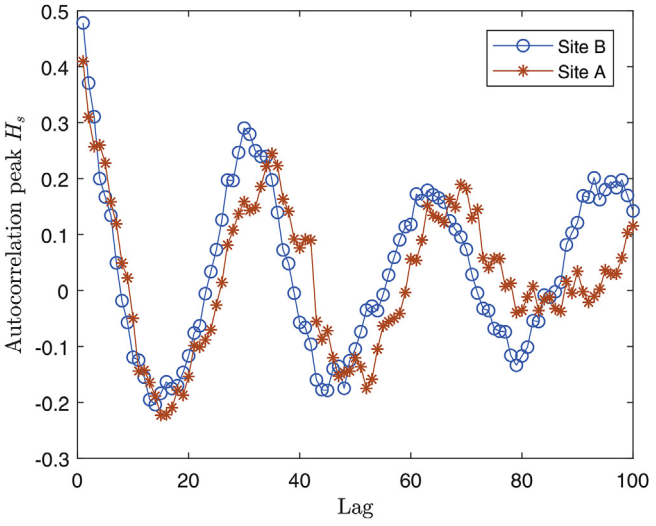

Fig. 3. Autocorrelation in storm peak H_s.

The seasonal variation in H_s can be modelled as:

$$H_s(t) = m(t) + s(t) H_s^{stat}(t) \tag{38}$$

where $H_s^{stat}(t)$ is assumed to be a stationary time series and $m(t)$ and $s(t)$ are deterministic periodic functions with period one year, representing the mean and standard deviation (STD) of the distribution of H_s at a given time in the year (see Monbet el al, 2007 and references therein). The functions $m(t)$ and $s(t)$ are estimated here using the method of Athanassoulis and Stefanakos (1995), with the mean and STD of H_s modelled using low order Fourier expansions. Figs. 4 and 5 show the mean and STD calculated in 10-day bins throughout the year for each dataset, together with a second order Fourier expansion of the binned values. The pattern in the seasonal variation differs between the two locations, but the second order model appears adequate for both locations. Figs. 4 and 5 also show the original data plotted against the day of year and the normalised data, H_s^{stat}, together with 90, 95 and 99% quantiles calculated in 15-day windows (the longer bin length has been used here for stability). It appears that the normalised data is approximately stationary for Site A, but there is some residual trend in the extreme values at Site B. However, as will be shown below, the normalisation seems to be sufficient to examine the short-range dependence (i.e. on time scales less than seasonal variation) in the extreme values.

Fig. 3 showed the autocorrelation function between storm peak H_s for peaks separated by at least 5 days. If we wish to assess the dependence structure in the extreme values without pre-selecting a declustering criterion to identify storm peaks, then the autocorrelation function is not appropriate since it measures the correlation between data at all levels. Davis and Mikosch (2009) introduced the concept of the extremogram as an analogue of the autocorrelation function for sequences of extreme events in a time series. The definition of the extremogram presented by Davis and Mikosch (2009) is used for assessing dependence between multivariate time series, but the concept is also applicable to univariate time series. To simplify the notation, the definition will be presented here for univariate time series only.

For a stationary univariate process (X_t), the extremogram is defined is defined as

$$\rho(\tau) = \lim_{u \to \infty} \Pr(X_{t+\tau} > u | X_t > u) \tag{39}$$

In the univariate case, the extremogram is the same as the tail dependence coefficient between X_t and $X_{t+\tau}$ (Beirlant et al., 2004). To test for correlation in exceedances at finite levels, an estimate of the extremogram at level u can be defined as

$$\rho_u(\tau) = \frac{\sum_{t=1}^{n-h} I_u(X_t) I_u(X_{t+\tau})}{\sum_{t=1}^{n} I_u(X_t)} \tag{40}$$

where I_u is the indicator function

$$I_u(X) = \begin{cases} 1 & \text{if } X > u, \\ 0 & \text{if } X \le u. \end{cases} \tag{41}$$

Two events A and B are independent if $P(A|B) = P(A)$. Therefore, if exceedances of threshold u at separation τ are independent then $\rho_u(\tau) - \Pr(X > u) = 0$. At extreme values $\Pr(X > u) \to 0$, and Davis and Mikosch (2009) showed that, in this limiting case, $\rho(\tau)$ has the properties of a correlation function. However, at finite levels $\Pr(X > u)$ is non-negligible and hence can be subtracted from $\rho_u(\tau)$ to assess independence.

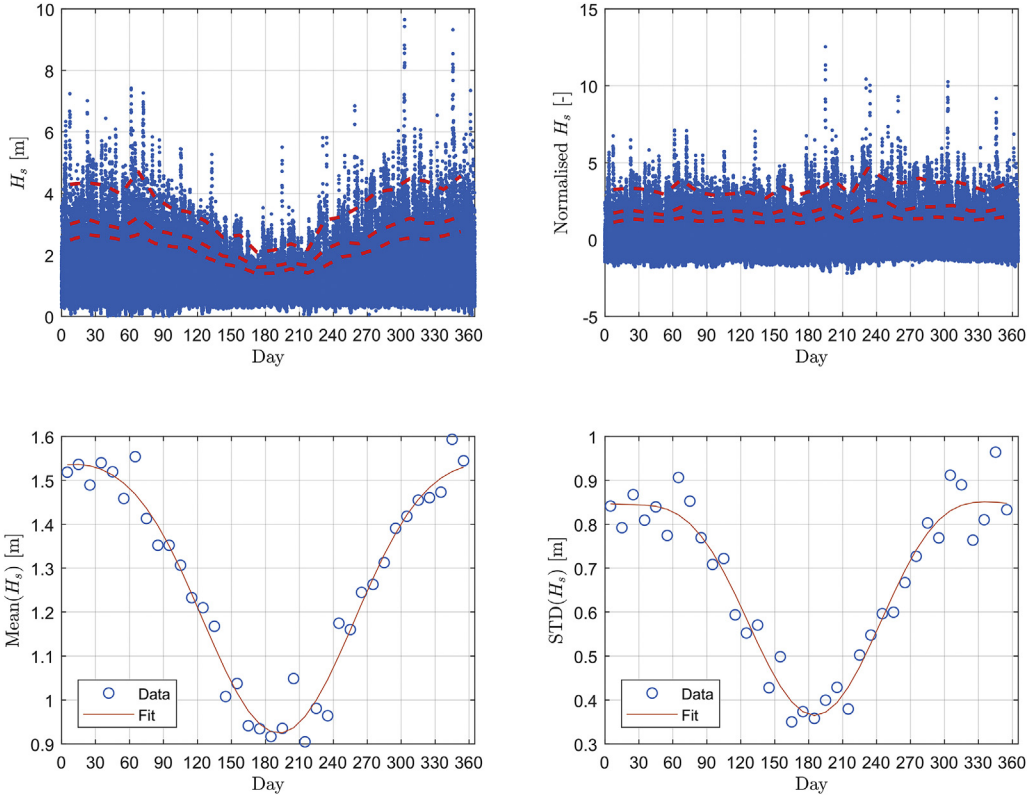

Fig. 4. Seasonal variation plots for Site A. Lower panels: 10-day mean and STD of H_s and second-order Fourier expansion. Upper panels: Hourly values of H_s (left) and normalised H_s (right) against day of year. Red dashed lines are 90, 95 and 99% quantiles in 15-day bins. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

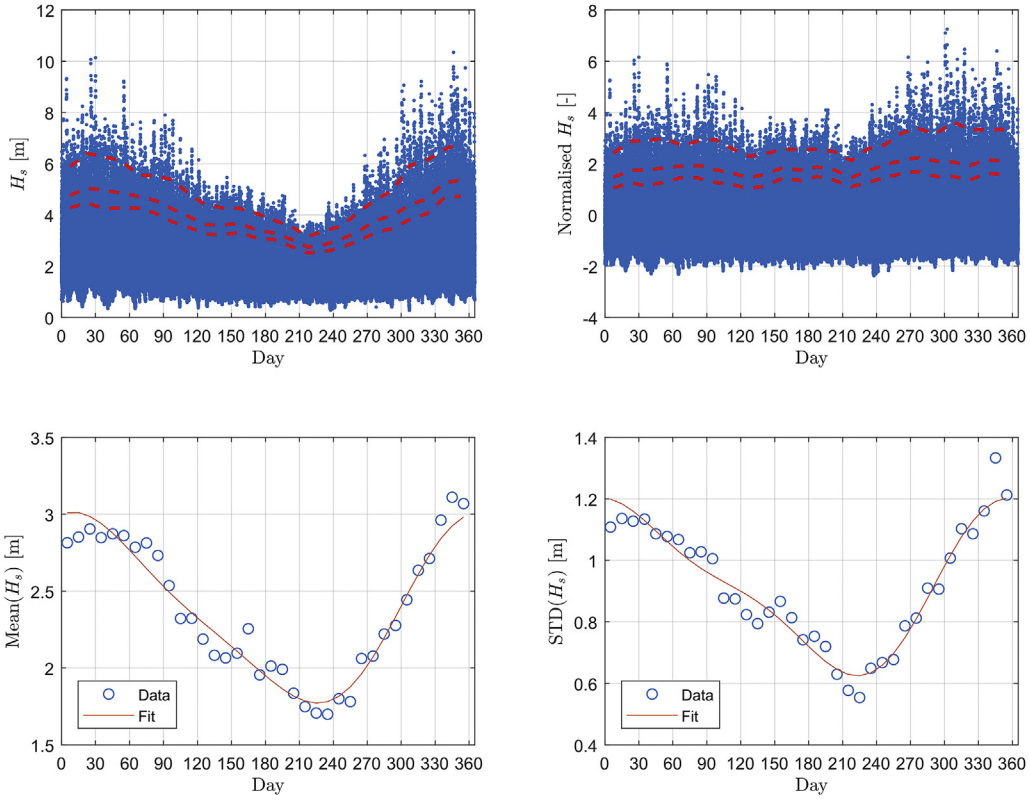

Fig. 5. As Fig. 4, but for Site B.

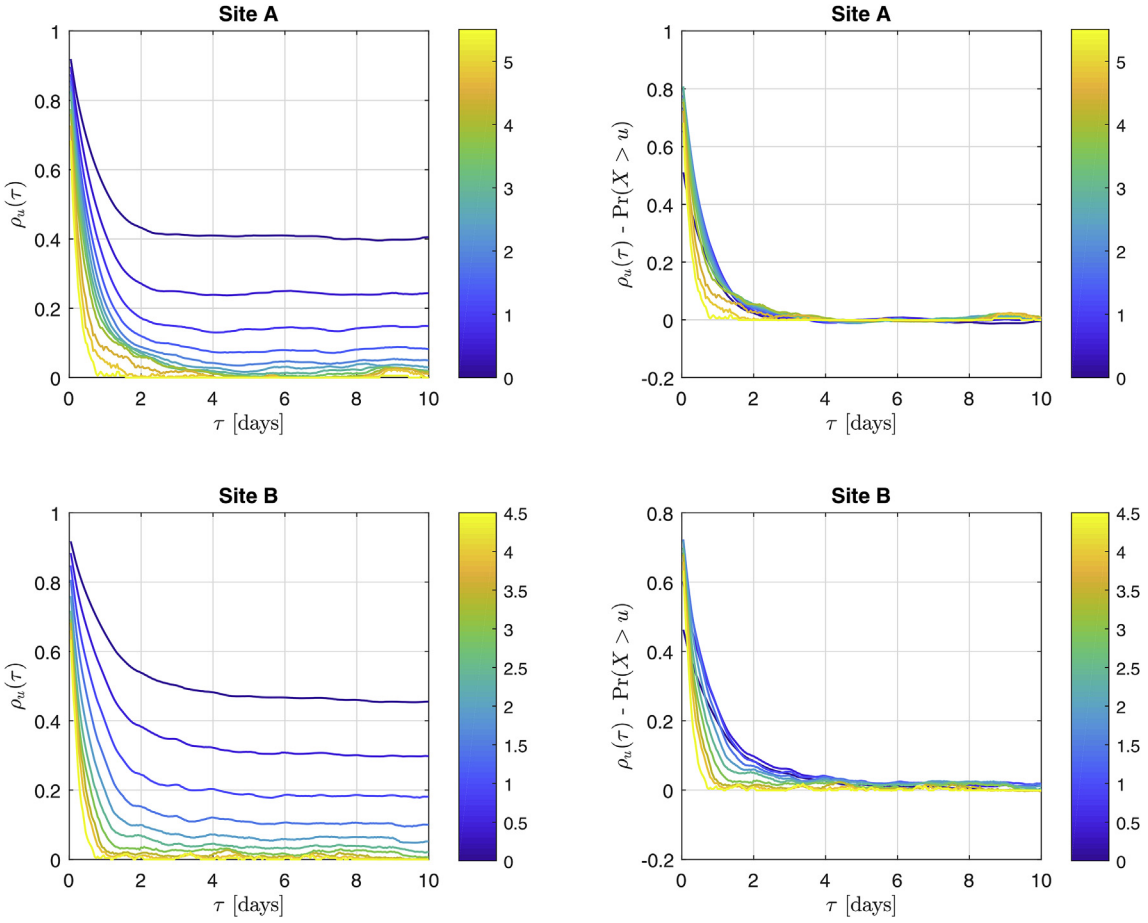

Fig. 6. Extremogram plots for normalised H_s data for both dataset. Colour of lines denote threshold value u from 0 to 4 at intervals of 5m. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Fig. 6 shows $\rho_u(\tau)$ and $\rho_u(\tau) - \Pr(X > u)$ for the normalised H_s data for both datasets. Values are shown for separation times between 0 and 10 days and a range of threshold levels. A threshold of 0 corresponds to just over the 50th percentile for both datasets, whilst the upper threshold level shown corresponds approximately the 99.9th percentile for both datasets. Obviously, the threshold value of 0 does not represent extreme values, but the range of thresholds is shown to illustrate the difference in the dependence structure between common and extreme values. It is also interesting to note that even for lower sea states the correlation decays to zero after about 5 days for both datasets. This means that selecting a declustering criteria that ensures independence at lower levels will also ensure independence at extreme levels.

For Site A, the dependence structure remains almost constant with threshold for most of the range, with events separated by ~4 days being effectively independent. For the highest thresholds, the correlation range decreases rapidly, with events separated by ~1 day being effectively independent for a threshold of 5.5. In contrast, Site B exhibits a more continuous gradual decrease in dependence with threshold, but a similar pattern in dependence at the highest levels.

By applying a seasonal normalisation and examining the dependence structure of the normalised time series, we are making the implicit assumption that the dependence structure of the smaller storms in the summer months is the same as that for the larger storms in the winter months, except for a shift in mean and STD of the time series. It is conceivable that this is assumption is invalid, and that the duration of high peaks differs in summer and winter. However, selecting a sufficiently large minimum separation time for declustering will ensure that the declustered events are effectively independent.

Given these observations, defining storms as local maxima in H_s in a 5-day window appears to be sufficient to ensure independence. This definition of storms will be used from here onwards. A separation time of 5 days is also reasonable based on physical arguments, since peaks separated by 5 days will correspond to waves generated from separate low pressure systems.

It is emphasised that the seasonal normalisation applied above has only been used to examine the dependence structure in the extreme values and define a declustering criterion. No model of the seasonal variation is required for estimating the long-term distribution of wave heights. The effect of seasonal variability on the simulated time series and extreme values is discussed in Section 7.1.

6.3. Fitting of distributions

So far, the method used to fit the long-term distribution models to the data has not been discussed. The AW and SSM methods both require a model for of $p(H_s)$, the equivalent storm method requires an estimate of $p(H_{mp})$, and the MC method requires multiple estimates of $\Pr(H_{max} \leq h|\text{any storm})$. For the storm-based methods, estimates of the distributions are only required for storms where the maximum wave height exceeds a threshold level, and the occurrence rate of storms exceeding the threshold level is accounted for in the calculation of return periods (2). This means that the standard POT approach can be adopted for the storm-based methods.

For the AW and SSM methods, a model for $p(H_s)$ is only needed for the tail of the distribution as well, since for high values of h and low values of H_s, $\Pr(H > h|\sigma(H_s)) \approx 0$. Therefore, the GPD can also be used to model $p(H_s)$ for the AW and SSM models and we can set the lower limit of integration in (13) and (18) as the threshold level used to fit the

GPD. This avoids the need to establish a model for the distribution of the entire range of values of H_s, which can be problematic, since fitting a model for the entire range of H_s does not guarantee a good fit to the highest values which have the strongest influence on the extreme wave heights. Ferreira and Guedes Soares (1999) showed that three models of $p(H_s)$ which all had a good fit over the bulk of the data, gave estimates of 100-year H_s which differed by over 5 m for a dataset from the Portuguese coast. Just modelling the tail of the distribution avoids this problem.

There are conflicting requirements in selecting the threshold used for modelling $p(H_s)$. The threshold must be low enough that the integrals (13) and (18) are not affected, but high enough to ensure a good fit of the GPD to the data. For the datasets considered here, it was found that the maximum percentile which could be used for the threshold level differed between the two datasets, with the integrals for Site A remaining invariant for thresholds up to the 99.5% quantile ($H_s \approx 8$m), but the integrals for Site B remaining invariant only up to the 95% quantile ($H_s \approx 7$m). The difference in sensitivity is due to the different shapes of the tails of the distribution (discussed below), with the distribution for Site A having a "longer tail" than for Site B, meaning that exceedance probabilities decrease with H_s at a lower rate in the tail of the distribution for Site A. Fortunately, for both datasets it was found that the GPD provides a good fit to the data for lower threshold values than this, as discussed below.

The GPD fitted to the tail of the distribution gives the conditional non-exceedance probability $\Pr(H_s \leq h | H_s > u)$, where u is the threshold level. The CDF of H_s can then be calculated as:

$$\Pr(H_s \leq h) = 1 - \Pr(H_s > u)[1 - \Pr(H_s \leq h | H_s > u)] \tag{42}$$

where $\Pr(H_s > u)$ is the fraction of samples where H_s exceeds the threshold. The GPD is defined as

$$\Pr(X \leq h | X > u) = \begin{cases} 1 - \left(1 + \xi\left(\frac{h-u}{\sigma}\right)\right)^{-\frac{1}{\xi}} & for\ \ \xi \neq 0\ and\ \sigma > 0, \\ 1 - \exp\left(-\left(\frac{h-u}{\sigma}\right)\right) & for\ \ \xi = 0\ and\ \sigma > 0. \end{cases} \tag{43}$$

Here, the variable X denotes either H_s, H_{mp} or H_{max}, depending on the application. The parameters σ and ξ are the scale and shape parameters respectively. When $\xi \geq 0$, the distribution is unbounded from above and referred to as heavy tailed or long tailed. When $\xi < 0$ the distribution has a maximum value, equal to $u - \sigma/\xi$, and is referred to as short tailed. The methods used to estimate the parameters of the GPD and select the threshold are discussed in the following two subsections.

It is worth noting that some practitioners opt to use the Weibull distribution for modelling the tail of the long-term distribution, rather than the GPD. Jonathan and Ewans (2013) discuss differences between the GPD and Weibull approach and note that using a Weibull distribution implicitly constrains the tail to be unbounded from above. It is reasonable to expect an upper bound on storm peak H_s or wave heights, due to physical constraints such as fetch limitations or water depth. The use of the GPD allows for the possibility of an upper bound on the tail of the distribution, whereas the Weibull distribution does not. The use of the GPD for modelling the tail of the distribution is also justified by asymptotic arguments (see e.g. Coles, 2001). For these reasons, the GPD has been used to model the tails of the long-term distributions in this work.

6.3.1. GPD parameter estimation

The method used to estimate the parameters of the GPD can have a significant influence on results, especially with relatively small sample sizes (Mackay et al., 2011). The metocean time series available for estimating extreme values are typically quite short in relation to the return periods of interest, resulting in small samples sizes. This means that using accurate estimators is important. The maximum likelihood (ML) method is commonly used in metocean applications (e.g. Jonathan and Ewans, 2013), due to its asymptotic properties of being unbiased and having the lowest possible variance. However, the ML estimators do not achieve these asymptotic properties until large sample sizes. Hosking and Wallis (1987) showed that the ML estimators are non-optimal for sample sizes up to 500, with higher bias and variance than other estimators, such as the moments and probability-weighted-moments estimators. Hosking and Wallis also noted that sometimes solutions to the ML equations do not exist and that at other times when a solution does exist there can be convergence problems with the algorithm they used to find them.

Estimation of the GPD parameters is the subject of ongoing research. A quantitative comparison of recent methods for estimating the parameters was presented by Kang and Song (2017). They concluded that the empirical Bayesian-likelihood method (EBM) of Zhang (2010) has one of the best performances in a wide range of cases. The EBM method is both computationally efficient and has low bias and variance compared to other methods, which is especially important for the MC method, where the fitting is performed for each random trial. The EBM method has therefore been used in this work. An algorithm for calculating the EBM estimators of the GPD parameters was provided by Zhang (2010), using the R statistical language, which can easily be translated into other codes. The reader is referred there for details.

6.3.2. Threshold selection

The selection of the threshold for fitting the GPD is a compromise between having a sufficient number of observations and violating the asymptotic arguments which justify the use of the GPD. Various methods have been proposed for selecting an appropriate threshold. An overview of graphical methods is presented by Coles (2001) and Jonathan and Ewans (2013) discuss some more recent studies. In this work the threshold has been selected by fitting the GPD for a range of thresholds and selecting the threshold as the lowest value for which the shape parameter ξ and estimates of high quantiles of the distribution reach an approximately stable value.

For each threshold value a bootstrap technique is used to establish confidence bounds on the estimates. At each trial of the bootstrap procedure, a random error is added to each resampled data point, corresponding to the sampling uncertainty in the measurement. This addition of a random error helps smooth the threshold plots, making trends easier to identify. The sampling distribution of H_s is approximately Guassian, with a coefficient of variation of $\mathrm{COV}(H_s) = \alpha\sqrt{T_m/D_\sigma}$ where $\alpha \approx 0.48$ for a Pierson Moskowitz spectrum and $\alpha \approx 0.61$ for a JONSWAP spectrum with peak enhancement factor of 3.3 (Forristall et al., 1996). In our datasets $D_\sigma = 1200s$ and the mean value of $\sqrt{T_m/D_\sigma}$ was 0.067 for Site A and 0.082 for Site B. A value of $\mathrm{COV}(H_s) = 0.04$ has been used for both datasets as a representative value, given the range of spectral shapes and values of T_m.

Threshold selection plots for the fit of the GPD to measured H_s for each dataset are shown in Fig. 7. Note that for the AW and SSM methods, the GPD is fitted to all values of H_s exceeding the threshold and no declustering is applied. For Site B the threshold has been selected as 7.5 m, since both the GPD shape parameter and 50-year return value of H_s tend to approximately constant values above this threshold. There are 44 storms where the peak H_s exceeds 7.5 m or ~1.4 storms per year. For Site A, the shape parameter and 50-year return values also appear approximately constant for thresholds above ~7.5 m. However, in this case there are only two storms where the peak H_s exceeds 7.5 m, which is insufficient to establish a reliable fit to the tail of the distribution. Instead, the threshold has been set at 5 m, where there is also a period of stability in the GPD shape parameter, and for which there are 54 storms where the peak H_s exceeds the threshold (2.4 storms per year). The fit of the GPD to the data at the selected thresholds is shown in Fig. 8. In both cases, the GPD provides a good visual fit to the data. Note that at Site A, for the two storms where the peak H_s exceeds 7.5 m, there are 17 hourly records with H_s exceeding 7.5 m.

The fitting of equivalent storms (defined in terms of the Gumbel and GEV distributions) to the two datasets was discussed in detail by Mackay and Johanning (2018a) and is not repeated here. The threshold

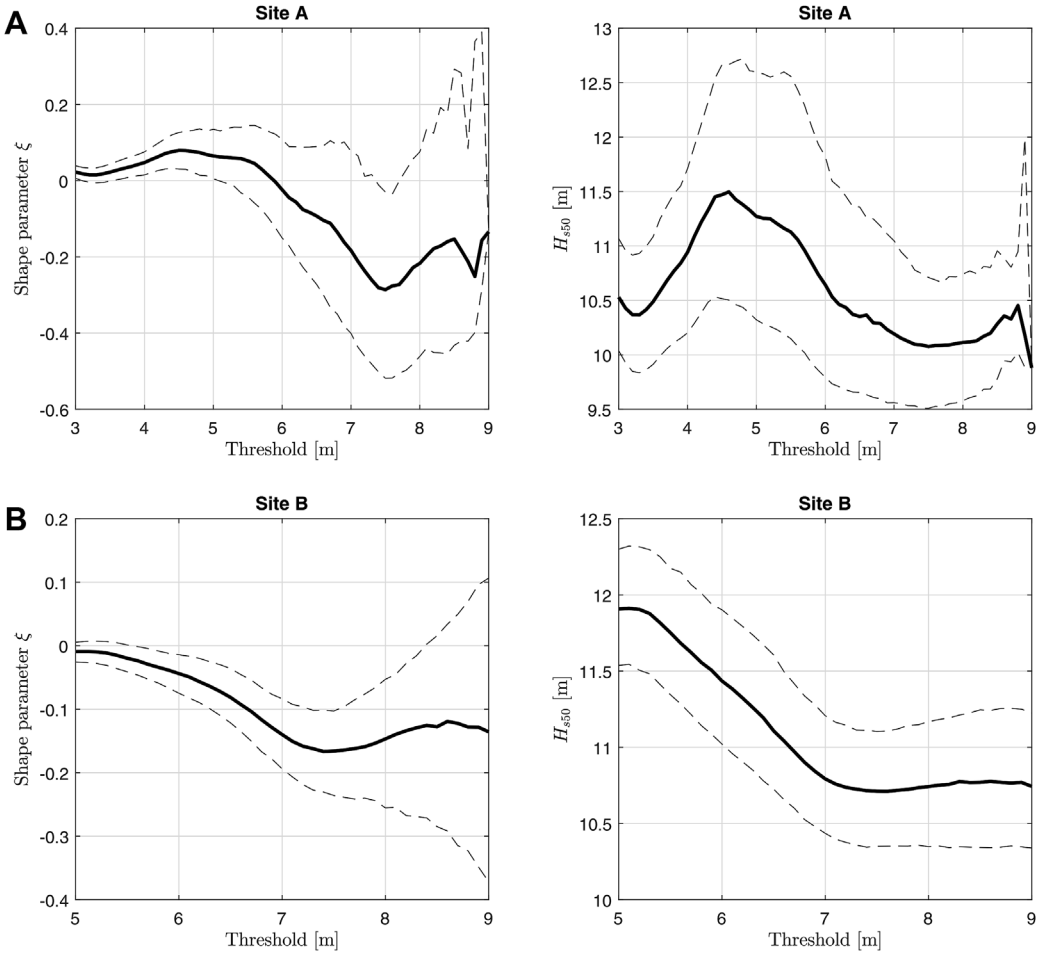

Fig. 7. Threshold plots for fit of GPD to H_s for both datasets.

selection plots for the fit of the GPD to the most probable maximum crest height in each storm, C_{mp}, are shown in Fig. 9. In this case, there is a clear choice of threshold for both datasets. For Site A both the shape parameter and 50-year return value of C_{mp} are approximately stable for thresholds above ~5 m and for Site B stability is reached at ~5.5 m. The fit of the GPD to each dataset at the selected threshold level is shown in Fig. 10 and appears satisfactory in both cases. For Site A the distribution is long-tailed due to the occurrence of two large storms, with C_{mp} ~2 m higher than all other storms. In contrast the distribution for Site B is short-tailed and there is a relatively close grouping in the values of C_{mp} for the largest storms.

For the MC method the threshold is selected by fitting to the declustered values of C_{med} (the median value of the maximum crest height in each storm, defined in Section 4.3.2). Fig. 11 shows the maximum value of C_{med} in each storm compared to C_{mp} for both datasets. The value of C_{mp} is slightly larger than max(C_{med}) on average, but the two

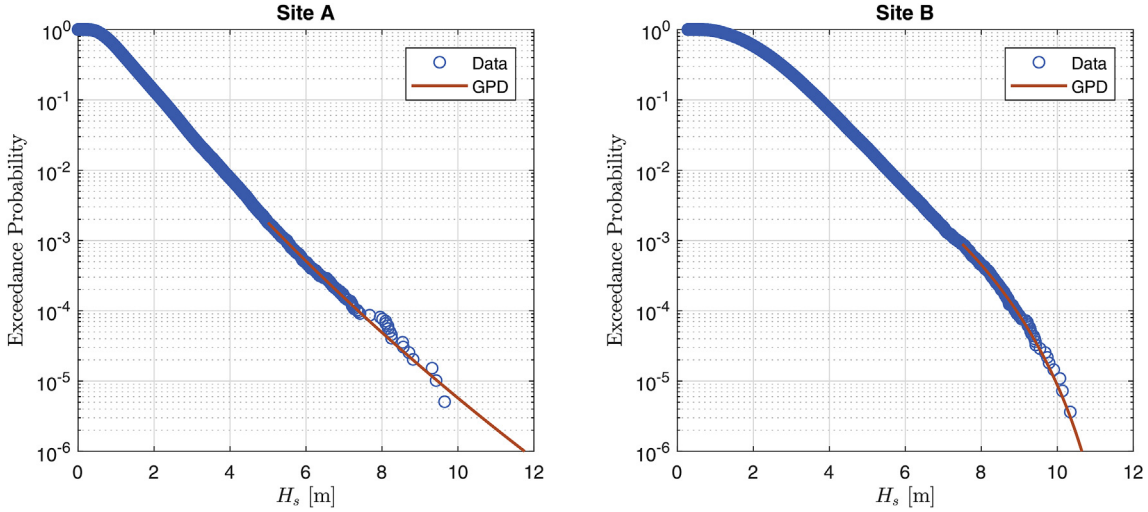

Fig. 8. Fit of the GPD to the tail of the distribution of H_s for both datasets.

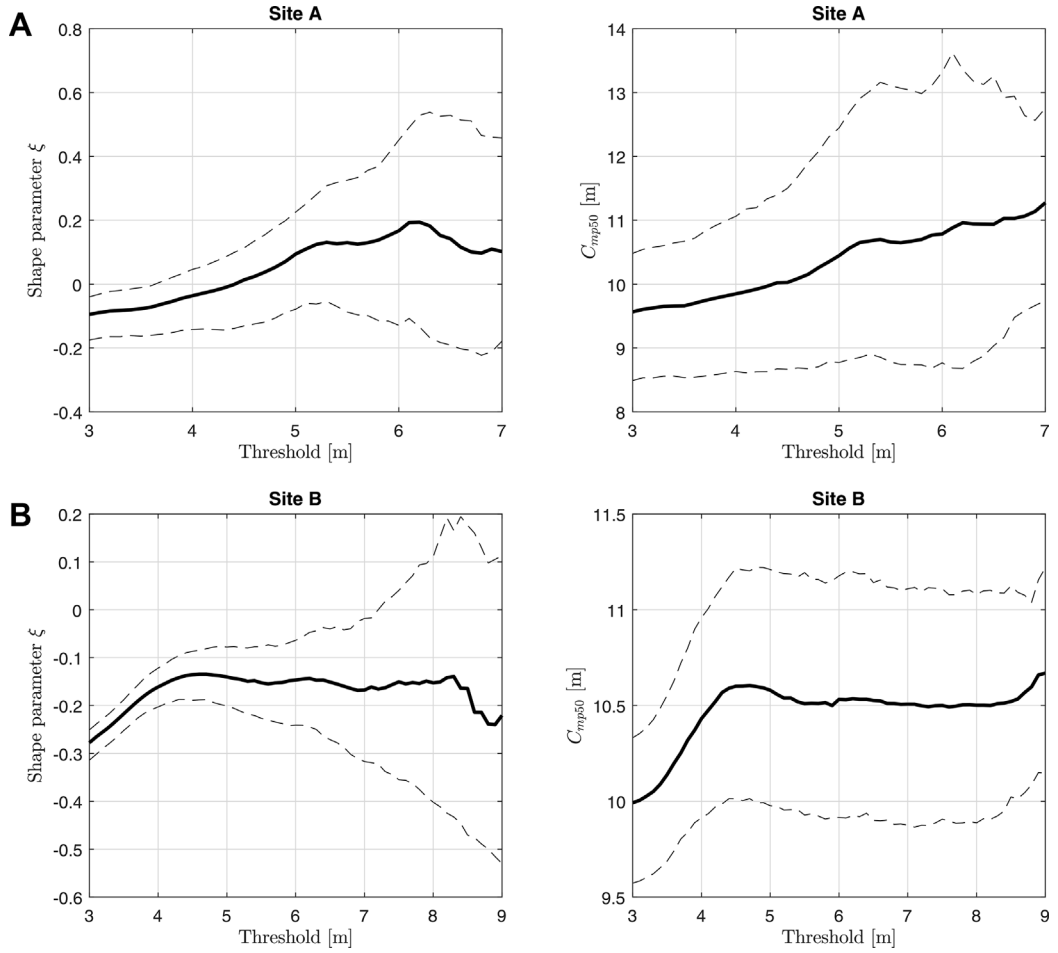

Fig. 9. Threshold plots for fit of GPD to C_{mp} for both datasets.

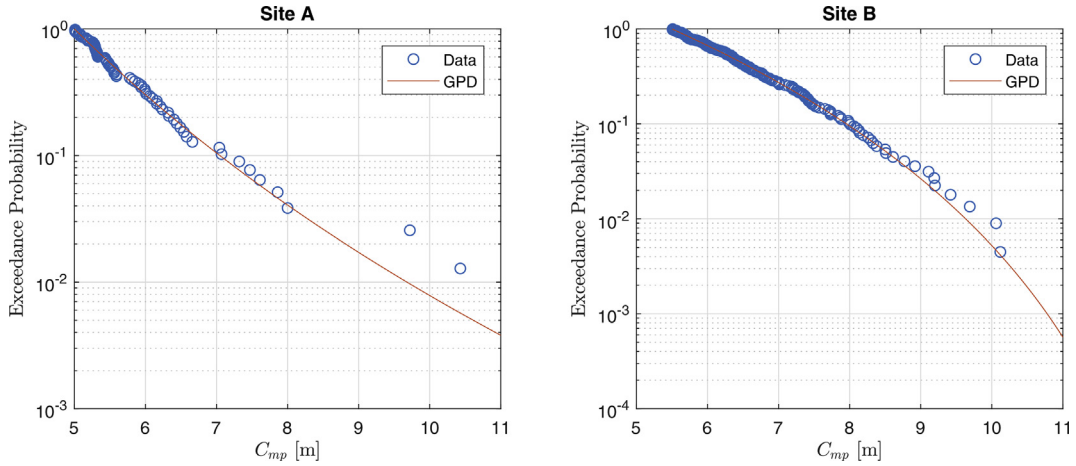

Fig. 10. Fit of the GPD to the tail of the distribution of C_{mp} for both datasets.

variables are strongly correlated with a Pearson correlation coefficient in excess of 0.99. Given the very strong correlation in the two variables, the choice of threshold for the equivalent storm method will also be appropriate for the MC method and has been applied here.

6.4. Comparison of return values of individual crest heights

Fig. 12 shows the return values of maximum crest heights calculated using each method. For both datasets the AW method gives the highest predicted crest heights at return periods of 1 year, but converges with the predicted crest heights from the SSM method for return periods of greater than 10 years, as expected from the theoretical considerations presented in Section 4. The storm-based methods agree well for both datasets. For Site A, the Gumbel and GEV methods are in close agreement over the entire range of return periods, whereas for Site B the Gumbel method produces slightly lower estimates at higher return periods. Mackay and Johanning (2018a) showed that for this dataset, the Gumbel and GEV methods agreed better if $\ln(N)$ is modelled as linearly dependent on C_{mp}, rather than constant. The difference in the estimate of C_{max} at the 100-year level between the Gumbel and GEV

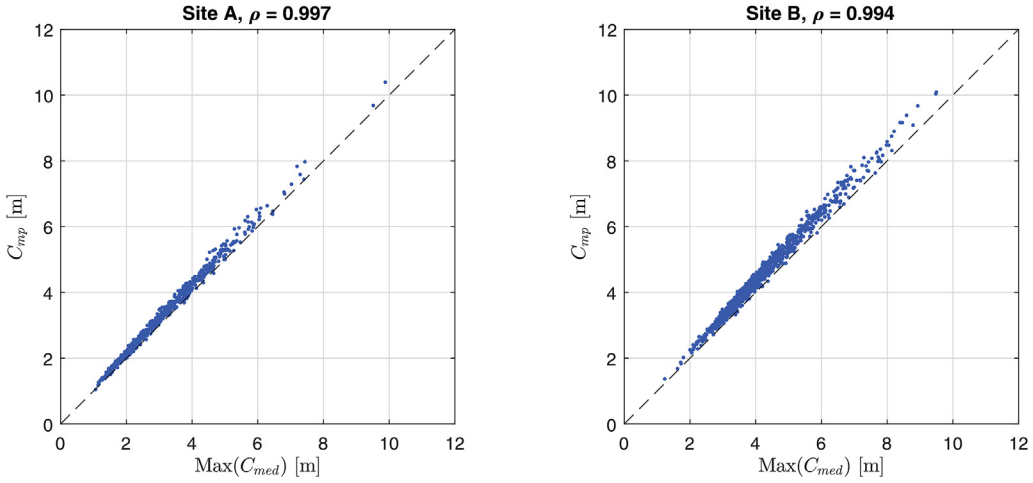

Fig. 11. Comparison of maximum value of C_{med} in each storm with C_{mp}. Dashed lines show 1:1 relation.

methods is around 0.2 m and remains similar at the 1000-year level. Given the uncertainty in the estimates due to the finite length of the metocean dataset (discussed further in Section 7.2.2), the difference between the two methods is not significant.

For Site A, the AW and SSM methods result in significantly higher predictions at lower return periods than the storm-based methods, with a difference of 1.5 m at a return period of 10years and 1.1 m at a return period of 100years. For Site B, the AW and SSM methods agree better with the storm-based methods and give slightly lower predictions at return periods above 50 years.

Given that the AW and SSM methods neglect serial correlation in sea states, it would be reasonable to expect that they would give higher estimates of return values than the storm-based methods for both datasets. However, the difference between the return values calculated using AW and SSM methods and storm-based methods will depend both on the level of serial correlation in the extreme values and also any differences in the fitted distributions of H_s and C_{mp}. Given the uncertainty in the fitted distributions, it is not possible to conclude from this example whether the differences in the return values observed for Site A are due to serial correlation or uncertainty in the fitted distributions. The simulations presented in the following section remove the influence of the uncertainty in the fitted distributions, to illustrate the differences that arise from serial correlation only.

7. Application to synthesised data

In this section, synthesised time series of sea states are used to investigate the accuracy of each method. The synthetic time series are not intended to be precise representations of the wave climates at either of the locations considered in the previous section, but are intended to capture realistic features of wave climates in order to assess the accuracy of the models for the long-term distributions of individual crest heights. Section 7.1 describes the method used to synthesise the time series and the application of the models to the synthetic data is described in Section 7.2.

7.1. Creation of synthetic time series

Numerous methods have been proposed for generating synthetic time series of sea states (see Monbet et al., 2007 for a review). The challenge here is to generate a synthetic time series where the extreme characteristics are representative of a real situation and the joint evolution of sea state parameters during storms follows realistic patterns, matching the observed dependence structures (both in terms of serial correlation and the joint distribution of parameters). A block-resampling method (Härdle et al., 2003) has been applied here, where the measured time series of sea states is divided into discrete blocks and these blocks are resampled with replacement. The blocks correspond to storms where the peak H_s can be considered independent. Independent values of storm peak H_s are defined in the same way as before, as maxima within a 5-day moving window. The dividing points between blocks are defined to be the minimum value of H_s between adjacent storm peaks. Blocks with more than 10% missing data have been discarded. Dividing the time series in this way ensures that the maximum

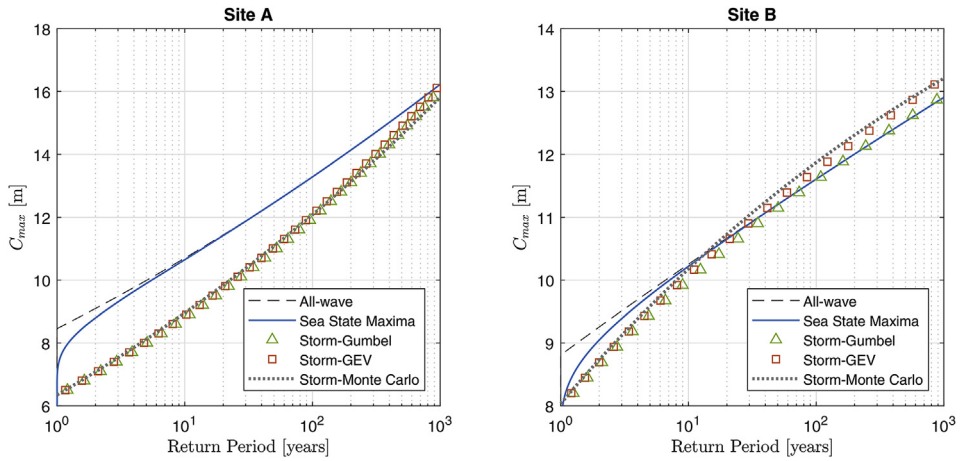

Fig. 12. Return values of maximum crest heights calculated from measured data using each method.

crest heights in each block are independent and that storm histories are realistic.

If the time series was resampled without modification, this would mean that the maximum H_s over any time period would be the maximum observed H_s. To overcome this, a distribution is fitted to the observed values of storm peak H_s, which can be used to extrapolate to extreme values (described below). To generate a storm with a random peak H_s, a random variable is generated from the fitted distribution and a block is selected at random from the 20 measured storms with the closest peak H_s. The values of H_s in the randomly selected measured storm are then linearly rescaled so that the peak H_s matches the randomly generated value. The values of T_m and T_z are scaled to maintain the observed steepness. If we denote the ratio of the random storm peak H_s to the measured storm peak H_s as $r = H^R_{s,peak}/H^M_{s,peak}$ (where the superscripts denote the random and measured values) then the rescaled values of T_m and T_z in the random storm are given by $T^R_m = T^M_m\sqrt{r}$ and $T^R_z = T^M_z\sqrt{r}$.

The reason for selecting the random storm from only the measured storms with the closest values of peak H_s is that the characteristics of the storms change with H_s, so sampling from the nearest neighbours preserves this relation. From Fig. 2 it is clear that the distribution of T_z is dependent on H_s. The distribution of H_s within a storm is also conditional on the value of the storm peak H_s. Fig. 13 shows the distribution of $H_s/H_{s,peak}$, binned by the value of storm peak H_s. It is evident that the distribution changes with the value of storm peak H_s and that the patterns differ between the two datasets, with the largest storms at Site A having a more peaked shape than the largest storms at Site B.

To avoid the problems with associated with initial distribution methods, discussed in Section 6.3, the distribution of storm peak H_s is fitted with a two-part model, using a lognormal model for the body of the data and a GPD model for the tail. A smooth transition between the models for the body and tail of the distribution is achieved using a mixing function, defined as

$$m(h) = \frac{1}{2}\left(1 + \tanh\left(\frac{h-v}{d}\right)\right) \tag{44}$$

where v is the mixing point and d is the mixing distance. The model for the CDF of storm peak H_s is then given by

$$\Pr(H_{s,peak} \le h) = (1 - m(h))P_{body}(h) + m(h)P_{tail}(h) \tag{45}$$

It was found that setting $d = 0.5$ and $v = u + d$, where u is the threshold at which the GPD is fitted, gave a smooth transitions for both datasets. The threshold for the GPD was selected as $u = 5.25$m for Site A and $u = 5.75$m for Site B. The fit of the models to the data are shown in Fig. 14. The models provide a good visual fit, both in the bodies and the tails of the distributions.

Scatter plots of H_s against T_z for example 500-year simulations are shown in Fig. 15, together with the measured H_s against T_z. The block-resampling method provides a realistic-looking joint distribution, which replicates the observed shapes reasonably well and preserves the relationship between H_s and significant steepness. However, as the method does not create new storm histories, but simply rescales measured storms, the storm with the largest value of T_z at Site A has been resampled several times, exaggerating this part of the distribution. It is reasonable to ask whether resampling of measured storms will give an adequate representation of the variability in storm characteristics. Mackay and Johanning (2018a) showed that the distribution of the maximum crest heights in storms can be very accurately represented in terms of the three-parameter GEV. Moreover, the joint distribution of the fitted GEV parameters exhibits remarkably regular characteristics, with the scale parameter (b) being strongly correlated to the location parameter ($a \approx C_{mp}$) and the shape parameter (k) taking a narrow distribution of values, approximately independent of a. Due to this regularity in the distribution of the maximum crest heights in measured storms, the block-resampling method should give an adequate representation of the variability in storm characteristics.

No attempt has been made to capture the seasonal variation in sea states in the synthetic time series. The rationale for this is that sampling all storms at random will capture the distribution of storms throughout the year. Over a one year section of simulated time series the seasonal variation in storms should be captured on average. Moreover, as we have assumed that the time series is composed of independent blocks, it is only the distribution of storms within the year that determines the distribution of the maximum crest height in the year and not the order in which storms occur. This argument is analogous to the argument that the order of sea states within a storm does not affect the distribution of the maximum crest height in the storm.

7.2. Results and discussion

A 100,000-year synthetic time series of sea states was generated for each dataset. For each sea state a random maximum crest height was simulated from the Forristall (2000) distribution using the method described in Section 4.3.2. The largest crest in each year of the synthetic time series was extracted and used to form an empirical estimate of $\Pr(C_{max} \le h|\ 1\text{ year})$, from which return values were calculated using (1). The return periods estimated from the 100,000-year simulation have approximately a 3% STD at the 100-year level and a 10% STD at the 1000-year level (see Appendix A). The synthetic time series were used to assess the accuracy of the SSM, ES and MC methods in two ways, as described in the following sections. The AW method was not considered in this section as it gives identical results to the SSM method for return periods above 10years.

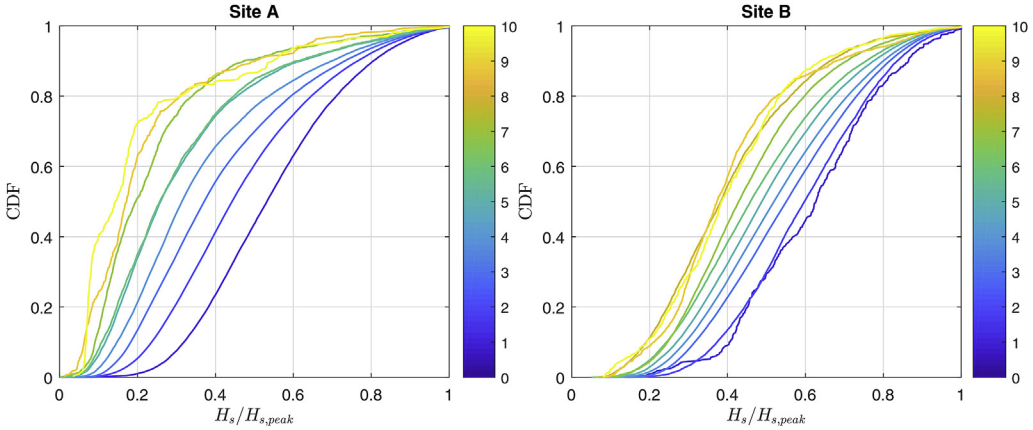

Fig. 13. CDF of normalised H_s, binned by peak H_s. Colour of line indicates peak H_s in metres. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

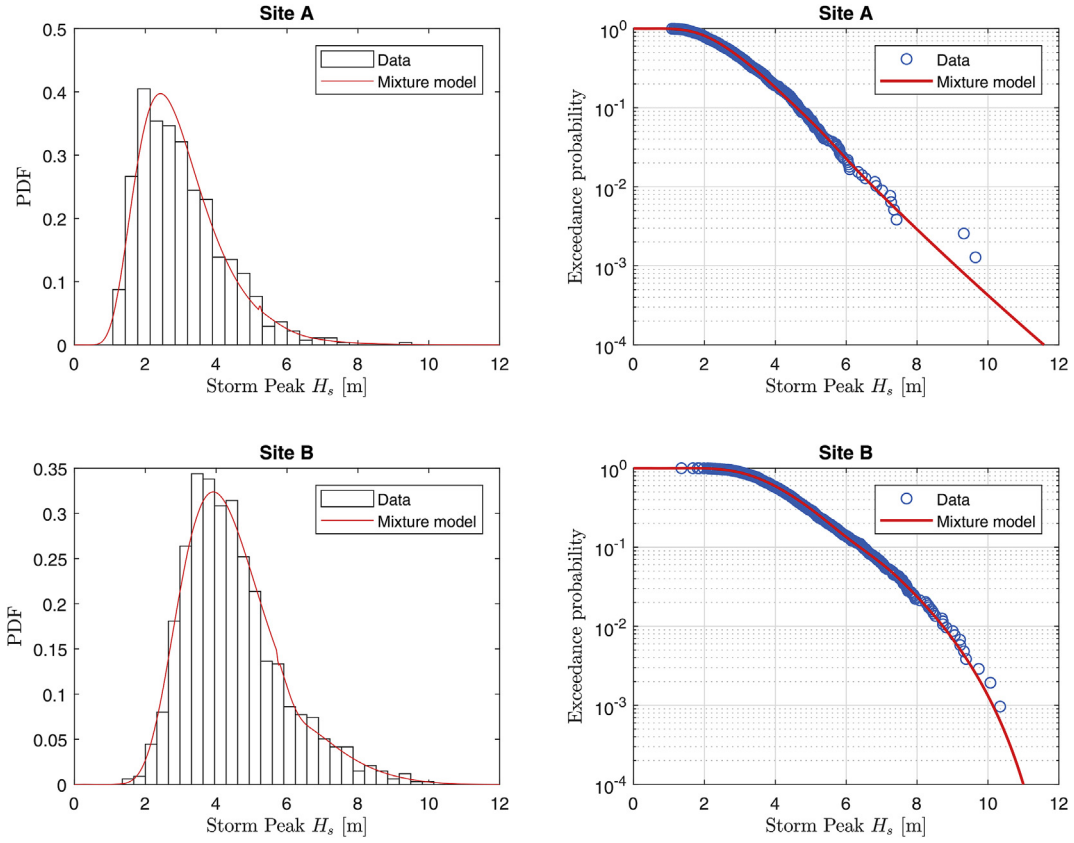

Fig. 14. Mixture models fitted to measured values of storm peak H_s.

7.2.1. Example 1 – distributions derived from 100,000-year simulation

In the first example, the long-term distributions $\Pr(H_s \leq h)$ and $\Pr(C_{mp} \leq h)$ have been derived from the full 100,000-year simulated time series. This effectively eliminates the uncertainty associated with fitting the distributions and isolates the effect of serial correlation on the estimated return values. The values of C_{mp} were calculated explicitly for each simulated storm, since the variation of C_{mp} with $H_{s,peak}$ is nonlinear due to the dependence of the short-term distribution parameters on the Ursell number.

To reduce computational time, the Gumbel and GEV distributions were not fitted to the distribution of the maximum crest height in the simulated storms. Instead the Gumbel and GEV parameters that were calculated from the measured data were applied in the simulations (i.e. the mean value of $\ln(N)$ for the Gumbel method and for the GEV the mean value of k and linear model $b(H_{mp})$). Similarly, the linear models for the variation of the short-term distribution parameters $\alpha(H_s)$ and $\beta(H_s)$ derived from the measured data were also used here for the calculations using the SSM method.

The MC method is not considered in this example. As the long-term distribution is assumed to be known, there is no fitting stage involved for the MC method. This means that the application of the MC method to the simulated sea states would be identical to the method used to derive the simulated crest heights.

The return values of C_{max} calculated using the SSM and storm-based methods are compared to the simulated values in Fig. 16. The estimates are very similar to those shown in Fig. 12, indicating that the models

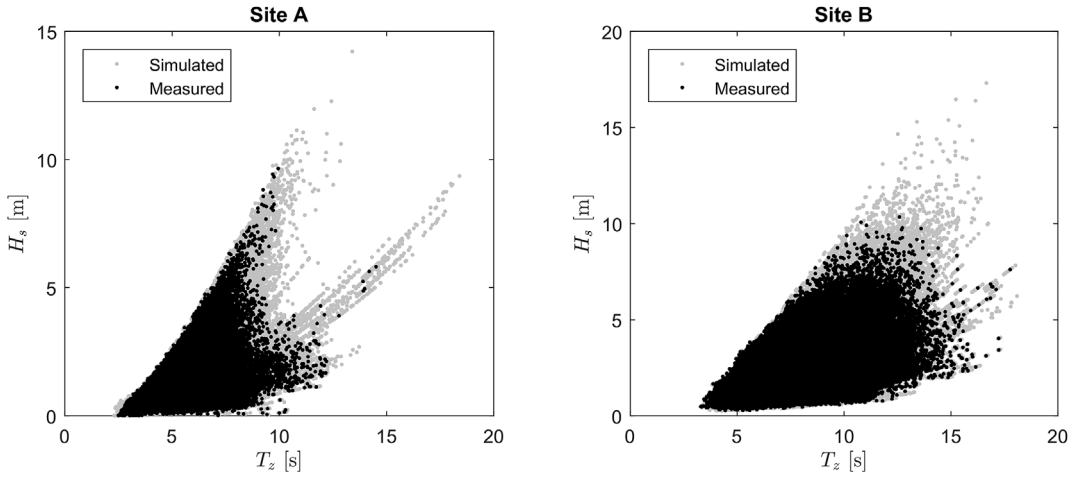

Fig. 15. Scatter plots of H_s against T_z for measured data and 500-year simulation.

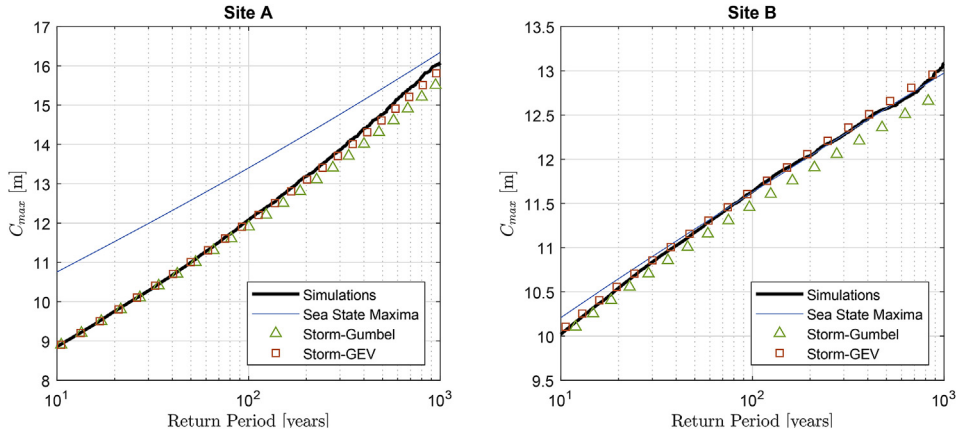

Fig. 16. Return values of maximum crest heights from simulations and calculated from sea-state-maxima and storm-based methods for case when the long-term distribution is known.

derived for $\Pr(H_s \leq h)$ and $\Pr(C_{mp} \leq h)$ for the measured data were reasonably accurate. The return values from the GEV method are very close to the simulated values for both datasets, whereas the Gumbel method produces slightly lower estimates than the simulated values. As mentioned in Section 6.4, this is likely to be due using a constant value of $\ln(N)$ in the Gumbel model rather than a value that is linearly dependent on C_{mp} (see Mackay and Johanning, 2018a). The SSM method gives return values that are close to the simulated values for Site B, but significantly overestimates return values for Site A.

In Section 6.2 the serial-dependence in sea states during high storms was shown to be similar for both datasets. Why then does neglecting serial correlation give reasonable results for one dataset but not the other? The fitted distribution of storm peak H_s for Site B was short-tailed, with the GPD shape parameter $\xi = -0.23$, whereas the distribution of storm peak H_s for Site A was (marginally) long-tailed, with $\xi = 0.03$. This means that the difference between the 10-year and 100-year H_s is larger for Site A than for Site B. So when a storm occurs with $H_s > H_{s,100}$ at Site B there are only a few sea states with $H_s > H_{s,10}$, whereas a storm with $H_s > H_{s,100}$ for Site A will contain many sea states with $H_s > H_{s,10}$. Therefore, neglecting serial correlation will have a larger impact when the distribution of storm peak H_s is long-tailed.

To test this explanation, the simulations for Site A were repeated with GPD shape parameter set at values of $\xi = -0.2, -0.1,\ 0$ and 0.05. All other variables were left unchanged. In particular, the same set of measured storms was resampled. The assumed distributions of $H_{s,peak}$ are shown in Fig. 17(a) and the bias in the return values of C_{max} calculated using the SSM method for each set of simulations are shown in Fig. 17(b). As expected, the bias increases with increasing shape parameter. This confirms that the tail shape has the dominant influence on the level of bias caused by neglecting serial correlation.

7.2.2. *Example 2 – distributions derived from 20-year simulations*

The second example is designed to examine the robustness of the estimated return values when the long-term distributions are fitted to the data. The 'true' return periods are calculated in the same way as in Example 1, using 100,000-year simulations for each dataset. The distributions $\Pr(H_s \leq h)$ and $\Pr(C_{mp} \leq h)$ are fitted to 20-year simulations, which is representative of a real situation using measured or hindcast data (analogous to that described in Section 6). The thresholds identified in Section 6.3.2 for fitting the GPD have also been used here. The fitting process was repeated 1000 times on different 20-year simulations to establish the mean and STD of the return periods calculated using each method. The mean and STD of the return values over the 1000 simulations are shown in Figs. 18 and 19.

The storm-based methods give very similar predictions for both datasets. They produce a small positive bias in the predicted return values for both datasets. This positive bias is a result of the estimators used for the GPD parameters, which is a known feature of the EBM method (Kang and Song, 2017). However, of the estimators considered, the EBM method performed the best in terms of bias and STD. In particular, using ML estimators results in larger bias and much slower computations. The Gumbel and GEV methods have very similar STDs for both datasets. The STD for the MC method is similar to that for the Gumbel and GEV methods at return periods up to 100 years, but differs slightly at longer return periods, with a slightly higher STD for Site A and slightly lower STD for Site B.

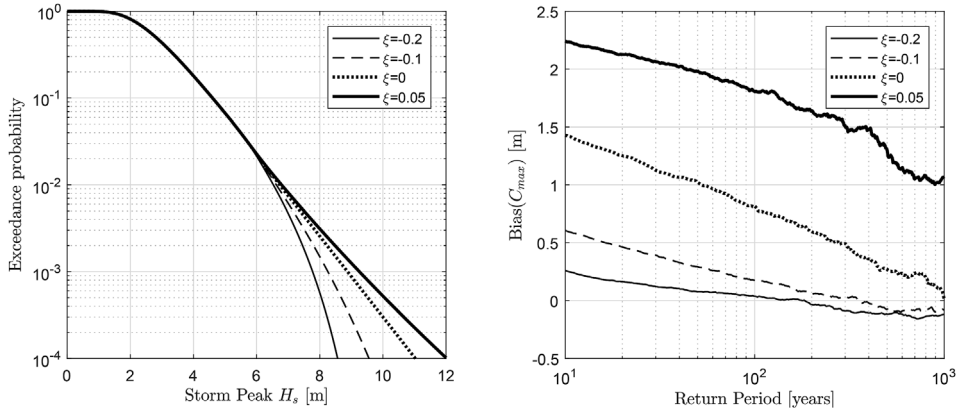

Fig. 17. Left: Assumed distributions of storm peak H_s with various GPD tail shapes. Right: Bias in return values of crest heights using sea-state-maxima method for simulations with various GPD tail shapes.

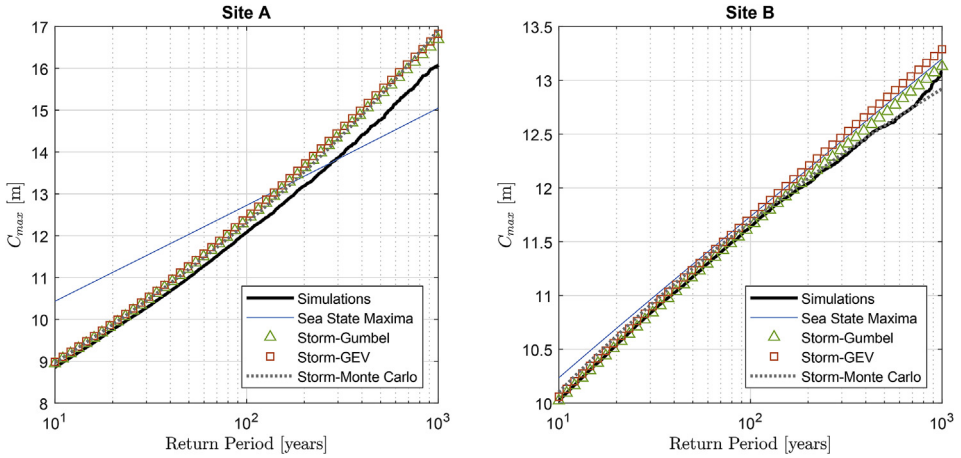

Fig. 18. Mean return values of maximum crest heights from 100,000-year simulations and calculated from 20-year simulations using sea-state-maxima and storm-based methods.

The SSM method performs similarly here to the previous examples, with a significant positive bias for low return periods at Site A and a negative bias at high return periods. For Site B the estimates from the SSM method are in close agreement with the simulated return values. Interestingly, the SSM method gives a higher STD in return value estimates than the storm-based methods at return periods less than 200 years, but lower STD at longer return periods. The higher STD at low return periods is likely related to the SSM method neglecting serial correlation. When a single large storm occurs which has multiple sea states above the 100 year level, the inclusion of all these sea states in the long-term distribution affects return values at low return periods. In contrast, the occurrence of such a storm only affects a single data point used in the inference for the storm-based methods, which results in greater stability in the estimates at lower return periods.

8. Conclusions

This paper has considered three methods for combining the long-term distribution of sea states with the short-term distribution of wave or crest heights conditional on sea state. It has been demonstrated both theoretically and numerically that methods which treat all waves as independent events give the same long-term distributions as methods that treat the maximum wave height in each sea state as independent events. Both of these methods neglect the effects of serial correlation in sea states. The numerical simulations presented in this work show that this can lead to significant positive bias in estimates of return values of individual wave and crest heights. The size of the positive bias is dependent on the shape of the tail of the distribution of storm peak H_s, with longer tails (larger values of the GPD shape parameter) leading to larger bias.

It was shown that storm-based methods give accurate predictions of return periods of individual wave heights. The method of Tromans and Vanderschuren (1995) which models the distribution of the maximum wave height in a storm using a Gumbel distribution, gives very similar results to the GEV method of Mackay and Johanning (2018a). The Gumbel method produces a slight underestimate of return values when the number of waves in the storm ($\ln(N)$) is assumed to be constant.

The Monte Carlo method proposed by Mackay and Johanning (2018b) was shown to compare well to the Gumbel and GEV storm-based methods for both the measured and simulated data. Of all the models considered, the Monte Carlo method requires the fewest assumptions and fitting stages with subjective judgements from the user. It does not require calculating the distribution of the maximum wave height in measured storms, fitting equivalent storms, estimating joint distributions of storm parameters or combining distributions via integration. Also, in comparison to the all-wave and sea-state-maxima methods, there is no requirement to form models of for the mean values of the short-term distribution parameters as a function of H_s. It is therefore recommended that the Monte Carlo method should be used for calculating return periods of individual wave and crest heights.

Declarations of interest

None.

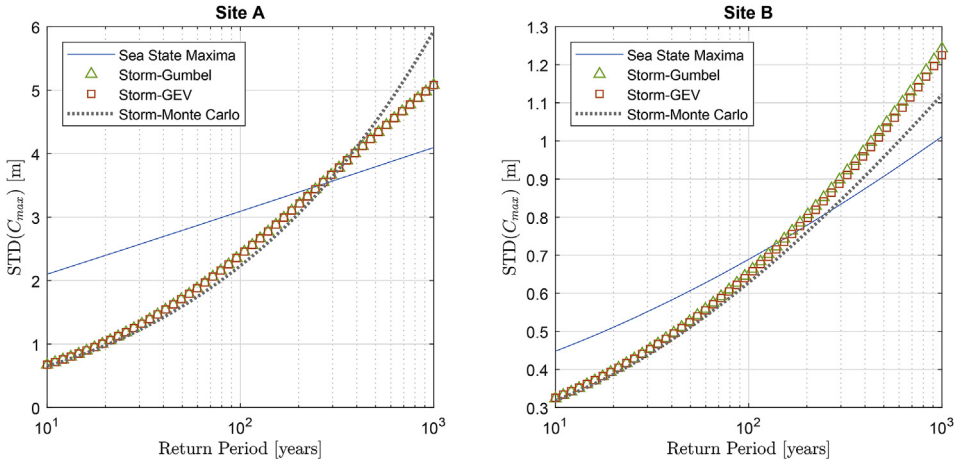

Fig. 19. STD in return values of maximum crest heights from 20-year simulations using sea-state-maxima and storm-based methods.

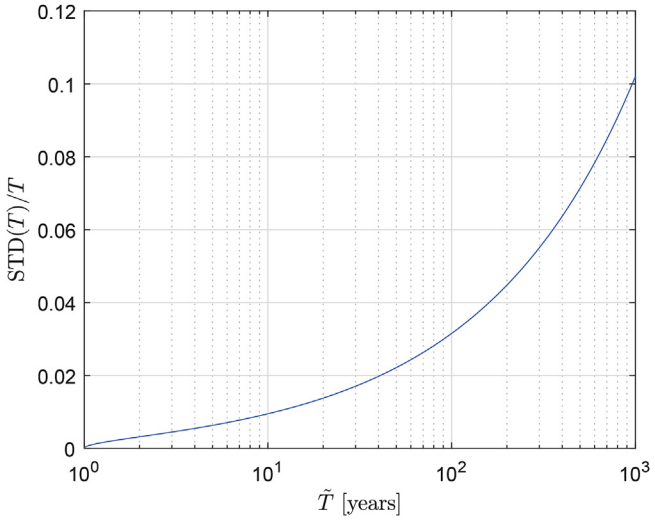

Fig. 20. STD in empirical estimates of return periods for simulation length of $n = 10^5$ years.

Acknowledgement

This work was partly funded through EPSRC grant EP/R007519/1.

Appendix A. Sampling properties of return periods

The sampling properties of the return periods estimated from the data can be derived as follows. Denote the ordered sample of annual maximum wave (or crest) heights as $H_{(1)} \leq H_{(2)} \leq ...\leq H_{(n)}$, where n is the length of the dataset in years. The empirical non-exceedance probability for the k^{th} order statistic is defined as

$$\widetilde{P} = \frac{k}{n+1} \tag{A.1}$$

Define

$$P_{(k)} = \Pr(H_{(k)} \leq h) \tag{A.2}$$

Note that $P_{(k)}$ is a random variable, but $\widetilde{P}$ is not – for any sample $\widetilde{P}$ is always given by (A.1). The empirical return period associated with the k^{th} order statistic is

$$\widetilde{T} = \frac{1}{1-\widetilde{P}} \tag{A.3}$$

which is also fixed for any sample. However, the true return period, $T = 1/(1 - P_{(k)})$, associated with the k^{th} order statistic is a random variable, whose sampling properties can be derived in terms of the distribution of $P_{(k)}$. The variable $P_{(k)}$ follows a beta distribution (David and Nagaraja, 2003), with density function given by:

$$f(P_{(k)}) = \frac{n!}{(k-1)!(n-k)!}(P_{(k)})^{k-1}(1-P_{(k)})^{n-k} \tag{A.4}$$

The expected values of T and T^2 are:

$$\mathrm{E}(T) = \frac{n!}{(k-1)!(n-k)!}\int_0^1 \frac{P^{k-1}(1-P)^{n-k}}{1-P}dP = \frac{n}{n-k} \tag{A.5}$$

$$\mathrm{E}(T^2) = \frac{n!}{(k-1)!(n-k)!}\int_0^1 \frac{P^{k-1}(1-P)^{n-k}}{(1-P)^2}dP = \frac{n(n-1)}{(n-k)(n-k-1)} \tag{A.6}$$

The variance of T is given by:

$$\mathrm{Var}(T) = \mathrm{E}(T^2) - [\mathrm{E}(T)]^2 = \frac{kn}{(n-k)^2(n-k-1)} \tag{A.7}$$

From (A.1) and (A.3) we have:

$$k = (n+1)\frac{(\widetilde{T}-1)}{\widetilde{T}} \tag{A.8}$$

Substituting this into (A.7) we arrive at:

$$\frac{STD(T)}{T} = \frac{1}{(n-\widetilde{T}+1)}\sqrt{\frac{n(n+1)(\widetilde{T}-1)}{n-2\widetilde{T}+1}}, \quad \widetilde{T} < \frac{n+1}{2} \tag{A.9}$$

The function in (A.9) is plotted in Fig. 20 using a value of $n = 10^5$ years. The estimated return periods from the 100,000 year simulations have a 3% STD at the 100-year level and a 10% STD at the 1000-year level.

Substituting (A.8) into (A.5) gives the expected value of the sample return periods as:

$$\mathrm{E}(T) = \left(\frac{n}{n-\widetilde{T}+1}\right)\widetilde{T} \tag{A.10}$$

When n is large and $n \gg \widetilde{T}$, the term in brackets is close to unity, but when $\widetilde{T}$ is close to n, there can be a significant positive bias in the sample return periods. For the 100,000 year simulations the bias is approximately 0.1% at the 100-year level and 1% at the 1000-year level. The bias is therefore effectively negligible compared to the STD.

References

Arena, F., Pavone, D., 2006. The return period of non-linear high wave crests. J. Geophys. Res. 111 C08004.

Arena, F., Malara, G., Romolo, A., 2014. On long-term statistics of high waves via the equivalent power storm model. Probabilist. Eng. Mech. 38, 103–110.

Athanassoulis, G.A., Stefanakos, C.N., 1995. A nonstationary stochastic model for long-term time series of significant wave height. J. Geophys. Res. 100 (C8), 16149–16162.

Battjes, J.A., 1970. Long-term Wave Height Distribution at Seven Stations Around the British Isles. National Institute of Oceanography Internal Report No. A44, 31pp.

Beirlant, J., Goegebeur, Y., Segers, J., Teugels, J., 2004. Statistics of Extremes: Theory and Applications. Wiley, New York MR2108013.

Boccotti, P., 1986. On coastal and offshore structure risk analysis. Excerpta of the Italian Contribution to the Field of Hydraulic Eng., 1. pp. 19–36.

Boccotti, P., 2000. Wave Mechanics for Ocean Engineering. Elsevier, New York.

Borgman, L.E., 1973. Probabilities for highest wave in hurricane. J. Waterw. Harb. Coast. Eng. Div. 99 (WW2), 185–207.

Brown, A., Gorter, W., Vanderschuren, L., Tromans, P., Jonathan, P., Verlaan, P., 2017. Design approach for Turret Moored vessels in highly variable Squall conditions. In: ASME. International Conference on Offshore Mechanics and Arctic Engineering, Volume 3A: Structures, Safety and Reliability:V03AT02A049, https://doi.org/10.1115/OMAE2017-61005.

Carter, D.J.T., Challenor, P.G., 1990. Metocean Parameters – Wave Parameters. Part 2: Estimation of Extreme Waves. UK Department of Energy, OTH89–300. HMSO, London.

Coles, S., 2001. An Introduction to the Statistical Modelling of Extreme Values. Springer-Verlag, London.

David, H.A., Nagaraja, H.N., 2003. Order Statistics. Wiley Series in Probability and Statistics.

Davis, R.A., Mikosch, T., 2009. The extremogram: a correlogram for extreme events. Bernoulli 15, 977–1009.

Davis, R.A., Mikosch, T., Zhao, Y., 2013. Measures of serial extremal dependence and their estimation. Stoch. Process. their Appl. 123, 2575–2602.

DNV GL, 2017. Environmental Conditions and Environmental Loads. Recommended Practice DNVGL-RP-C205, August 2017.

Fawcett, L., Walshaw, D., 2012. Estimating return levels from serially dependent extremes. Environmetrics 23, 272–283.

Fedele, F., 2005. Successive wave crests in Gaussian Seas. Probabilist. Eng. Mech. 20, 355–363.

Fedele, F., 2012. Space-time extremes in short-crested storm seas. J. Phys. Oceanogr. 42 (9), 1601–1615.

Fedele, F., 2016. Are rogue waves really unexpected? J. Phys. Oceanogr. 46, 1495–1508.

Fedele, F., Arena, F., 2010. Long-term statistics and extreme waves of sea storms. J. Phys. Oceanogr. 40 (5), 1106–1117.

Ferreira, J.A., Guedes Soares, C., 1999. Modelling the long-term distribution of significant wave height with the Beta and Gamma models. Ocean Eng. 26, 713–725.

Forristall, G.Z., Heideman, J.C., Leggett, I.M., Roskam, B., Vanderschuren, L., 1996. Effect of sampling variability on hindcast and measured wave heights. J. Waterw. Port, Coast. Ocean Eng. 122 (5), 216–225 Discussion: Goda Y 124(4): 214.

Forristall, G.Z., Larrabee, R.D., Mercier, R.S., 1991. Combined oceanographic criteria for deepwater structures in the Gulf of Mexico. In: 23rd Offshore Technology Conference, Houston, Texas, OTC-6541-MS.

Forristall, G.Z., 2000. Wave crest distributions: observations and second order theory. J. Phys. Oceanogr. 30 (8), 1931–1943.

Forristall, G.Z., 2008. How should we combine long and short term wave height distributions? In: Proc 27th Int. Conf. Offshore Mech. Arctic Eng, (OMAE2008–58012). Estoril, Portugal, 15–20 June 2008.

Giske, F.I.G., Leira, B.J., Øiseth, O., October 2017. Full long-term extreme response analysis of marine structures using inverse FORM. Probabilist. Eng. Mech. 50, 1–8.

Hagen, O., Grue, I.H., Birknes-Berg, J., Lian, G., Bruserud, K., 2017. Analysis of Short-term and Long-term Wave Statistics by Time Domain Simulations Statistics. Proc OMAE. Paper 61510.

Haring, R.E., Heideman, J.C., 1980. Gulf of Mexico rare wave return periods. J. Petrol. Technol. 32 (1), 35–47.

Hardle, W., Horowitz, J., Kreiss, J.P., 2003. Bootstrap methods for time series. Int. Stat. Rev. 71 (2), 435–459.

Hogben, N., 1990. Long term wave statistics. In: Le Mehaute, B., Hanes, D. (Eds.), In the Sea v.9, Pt. A Ocean Engineering Science. Wiley, New York, pp. 293–333.

Hosking, J.R.M., Wallis, J.R., 1987. Parameter and quantile estimation for the generalized Pareto distribution. Technometrics 29, 339–349.

ISO, 2015. Petroleum and Natural Gas Industries – Specific Requirements for Offshore Structures - Part 1: Metocean Design and Operating Considerations. ISO 19901-1:2015.

Jahns, H.O., Wheeler, J.D., 1973. Long term wave probabilities based on hindcasting of severe storms. J. Petrol. Technol. 25 (4), 473–486.

Janssen, P.A.E.M., 2015. Notes on the maximum wave height distribution. European centre for medium-range weather forecasts. Tech. Memo. 755 June 2015.

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

Kang, S., Song, J., 2017. Parameter and quantile estimation for the generalized Pareto distribution in peaks over threshold framework. J. Korean Statistical Society 46 (4), 487–501.

Krogstad, H.E., 1985. Height and period distributions of extreme waves. Appl. Ocean Res. 7, 158–165.

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

Laface, V., Arena, F., 2016. A new equivalent exponential storm model for long-term statistics of ocean waves. Coast Eng. 116, 133–151.

Laface, V., Arena, F., Maisondieu, C., Romolo, A., 2017. On long term statistics of ocean storms starting from partitioned sea states. In: Proc. 36th Int. Conf. Oce., Offshore Arctic Eng., June 25-30, 2017, Trondheim, Norway, OMAE2017-61750.

Lagarias, J.C., Reeds, J.A., Wright, M.H., Wright, P.E., 1998. Convergence properties of the Nelder-Mead simplex method in low dimensions. SIAM J. Optim. 9 (1), 112–147.

Mackay, E.B.L., Johanning, J., 2018a. A Generalised Equivalent Storm Model for Long-term Statistics of Ocean Waves. Accepted for publication in Coastal Engineeringhttps://doi.org/10.1016/j.coastaleng.2018.06.001.

Mackay, E.B.L., Johanning, J., 2018b. A simple and robust method for calculating return periods of ocean waves. In: ASME 37th International Conference on Ocean, Offshore and Arctic Engineering OMAE2018, Madrid, Spain, 17th - 22nd Jun 2018.

Mackay, E.B.L., Challenor, P.G., Bahaj, A.S., 2011. A comparison of estimators for the generalised Pareto distribution. Ocean Eng. 38 (11–12), 1338–1346.

Martín-Soldevilla, M.J., Martín-Hidalgo, M., Negro, V., López-Gutiérrez, J.S., Aberturas, P., 2015. Improvement of theoretical storm characterization for different climate conditions. Coast Eng. 96, 71–80.

Monbet, V., Ailliot, P., Prevosto, M., 2007. Survey of stochastic models for wind and sea state time series. Probabilist. Eng. Mech. 22 (2), 113–126.

Naess, A., 1984. Technical Note: on the long-term statistics of extremes. Appl. Ocean Res. 6 (4), 227–228.

Naess, A., 1985. The joint crossing frequency of stochastic processes and its application to wave theory. Appl. Ocean Res. 7 (1), 35–50.

Naess, A., Gaidai, O., Haver, S., 2007a. Efficient estimation of extreme response of drag dominated offshore structures by Monte Carlo simulation. Ocean Eng. 34 (16), 2188–2197.

Naess, A., Gaidai, O., Teigen, P.S., 2007b. Extreme response prediction for nonlinear floating offshore structures by Monte Carlo simulation. Appl. Ocean Res. 29, 221–230.

Naess, A., Moan, T., 2012. Stochastic Dynamics of Marine Structures. Cambridge University Press, Cambridge ISBN: 9781139021364.

Prevosto, M., Krogstad, H.E., Robin, A., 2000. Probability distributions for maximum wave and crest heights. Coast Eng. 40 (4), 329–360.

Sagrilo, L.V.S., Naess, A., Doria, A.S., July 2011. On the long-term response of marine structures. Appl. Ocean Res. 33 (3), 208–214.

Tromans, P.S., Vanderschuren, L., 1995. Response based design conditions in the North Sea: application of a new method. In: Proc. Offshore Tech. Conf., OTC, pp. 7683.

Tucker, M.J., 1989. Improved 'Battjes' method for predicting the probability of extreme waves. Appl. Ocean Res. 11 (4), 212–218.

Tucker, M.J., Pitt, E.G., 2001. Waves in Ocean Engineering. Elsevier, Amsterdam.

Ward, E.G., Borgman, L.E., Cardone, V.J., 1979. Statistics of hurricane waves in the Gulf of Mexico. J. Petrol. Technol. 632–642.

Zhang, J., 2010. Improving on estimation for the generalized Pareto distribution. Technometrics 52, 335–339.