Coastal Engineering xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Coastal Engineering

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

A generalised equivalent storm model for long-term statistics of ocean waves

Ed Mackay*, Lars Johanning

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

ARTICLE INFO

Keywords:
Wave height
Crest height
Long-term statistics
Equivalent storm
Return period

ABSTRACT

To calculate the return periods of individual wave or crest heights, the long-term distribution of sea states must be combined with the short-term distribution of individual wave or crest heights conditional on sea state. This is normally achieved using an equivalent storm model to parameterise the distribution of the maximum wave or crest height in a storm. A new equivalent storm model is introduced that generalises the approach of Tromans and Vanderschuren (1995). The generalised equivalent storm (GES) method is significantly simpler than equivalent storm methods that model the temporal evolution of the significant wave height in a storm. The GES method is applied to long time series of wave buoy measurements for deep and shallow water sites and demonstrated to be more accurate than existing methods at representing the statistical characteristics of measured storms. Return periods of crest heights from the GES method are shown to be more robust to uncertainties in the fitted models of the equivalent storm parameters than estimates from temporal evolution methods such as the equivalent triangular storm and equivalent power storm model.

1. Introduction

Calculating the long-term statistics of extreme individual wave and crest heights is an important problem in coastal and offshore engineering. This topic has received considerable attention in the literature, with various approaches proposed. The problem involves combining the short-term distribution of wave or crest heights conditional on sea state with the long-term distribution of sea states. The methods for combining the long-term and short-term distributions are identical for both wave and crest heights, so to avoid referring to both throughout the work, the following discussion will be presented in terms of wave heights rather than crest heights.

The proposed approaches can be grouped into three categories. In the first approach, originally proposed by Battjes (1970) and subsequently refined by Tucker (1989), the long-term distribution of all wave heights is formed by calculating the proportion of waves or crests exceeding a value in a given sea state and weighting by the probability of occurrence of the sea state.

Variants of the second approach have been proposed by Krogstad (1985) and Tucker and Pitt (2001, Section 6.4.2). In this approach the long-term distribution of wave heights is formed by integrating the short-term distribution of the maximum wave height in each sea state weighted by the probability of occurrence of each sea state.

In the third approach, the metocean data is divided into non-overlapping storms and the distribution of the maximum wave height in each storm is calculated. This distribution is then characterised in terms of an 'equivalent storm' which has a simple parametric form. The equivalent storm is fitted so that the distribution of the maximum wave height in the equivalent storm matches that in the measured storm. The long-term distribution of wave heights is then formed by integrating the short-term distribution of the maximum wave height in a storm, weighted by the distribution of storm parameters.

A comparison of these three types of methods was presented by Forristall (2008), using Monte Carlo simulations of individual wave heights over long time series of synthetic storms. Forristall demonstrated that Battjes' method gives higher estimates of return values for individual waves than Krogstad's method. He explains that the reason for this discrepancy is that when a storm occurs in which the significant wave height, H_s, exceeds the N-year return value, there can be many very large individual waves during that interval. All of these individual waves go into the distribution given by Battjes, whereas the Krogstad method only counts the highest wave in each r-hour interval. Forristall goes on to note that the Krogstad method gives higher estimates than methods that only consider the highest wave in each storm, since there can be several high waves in separate r-hour periods within each storm.

For many engineering applications, it seems more appropriate to calculate the probability that a single storm occurs in which a wave height exceeds a certain threshold. If such a storm occurs, usually it does not matter if that height is exceeded more than once within that storm, as the damage to the structure will occur the first time the threshold is exceeded.

Storm-based methods for calculating extreme wave conditions can

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

https://doi.org/10.1016/j.coastaleng.2018.06.001
Received 23 January 2018; Received in revised form 23 May 2018; Accepted 3 June 2018

be grouped into two categories, based on how the distribution of the maximum wave in the storm is parameterised. 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; Martín-Hidalgo et al., 2014; Laface et al., 2017), trapezoid (Martín-Soldevilla et al., 2015), parabola (Tucker and Pitt, 2001, Section 6.5.4), power law (Fedele and Arena, 2010; Arena et al., 2014) or exponential (Laface and Arena, 2016). The parameters of the equivalent storm are fitted so that the distribution of the maximum wave height in the equivalent storm is matched as closely as possible to the measured 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, so the measured storm can be re-ordered into a monotonically increasing series of sea states, 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, so if the short-term distribution of wave height is dependent on with multiple sea state parameters (such as steepness or directional spread) then a relation between other sea state parameters and H_s must be assumed in order to calculate the distribution of the maximum wave height in the equivalent storm.

The other approach, which has been less widely studied, is to model the distribution of the maximum wave height in the storm directly. This approach was adopted by Tromans and Vanderschuren (1995), who assumed that the maximum wave height in the storm followed a Gumbel distribution, with the parameters given as a function of the most probable wave height in the measured storm. This forces a fixed relation between the parameters of the distribution, which is not necessarily applicable in all locations.

In this paper, a generalisation of this approach is proposed, where the distribution of the maximum wave height in a storm is modelled as following a Generalised Extreme Value (GEV) distribution (for which the Gumbel distribution is a special case) and the parameters of the GEV are determined for each measured storm. This provides a more flexible approach than the Tromans and Vanderschuren (1995) method. Compared to temporal evolution models, the GEV method requires fewer assumptions and is simpler to implement.

The paper is organised as follows. The short-term distribution of individual wave and crest heights conditional on sea state is discussed in Section 2. The distribution of the maximum wave height in a storm is derived in Section 3. Existing equivalent storm models are reviewed in Section 4 and the generalised equivalent storm model is presented. The estimation of the long-term distribution of the parameters of the equivalent storms is discussed in Section 5. The method for combing the long-term and short-term distributions to calculate return periods of individual wave heights is presented in Section 6. An example application of the equivalent storm methods to measured wave data is presented in Section 7 and the accuracy and robustness of the methods is compared. The differences between the equivalent storm models considered are summarised in Section 8. Finally, conclusions are presented in Section 9.

2. The distribution of wave or crest 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 or crest heights conditional on sea state is denoted

$$\Pr(H \le h|\sigma) \tag{1}$$

Models for (1) 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 an exhaustive review. Wu et al. (2016) reviewed models for the short-term distribution of wave heights in shallow water and noted that many studies have shown dependence of the short-term distribution on steepness, relative water depth and spectral bandwidth (see also Tayfun and Fedele, 2007). Latheef and Swan (2013) demonstrated that the distribution of crest heights is dependent on both the steepness and directional spread of the sea state. The presence of currents also influences the wave and crest statistics, with opposing currents believed to be a contributing factor to the formation of rogue waves (Toffoli, 2016).

The focus of this work is on how to combine the short-term distribution (1) with the long-term distribution of sea states. The methods described are applicable to any model for the short-term distribution. The example presented in Section 7 uses the Forristall (2000) model for the short-term distribution of crest heights in directionally spread seas. The Forristall model assumes that crest heights follow a Weibull distribution:

$$\Pr(H \le h|\sigma) = 1 - \exp\left(-\left(\frac{h}{\alpha H_s}\right)^\beta\right). \tag{2}$$

The parameters α and β defined in terms of the significant steepness, s_m, and Ursell number, U_{rs}, defined as

$$s_m = \frac{2\pi H_s}{g T_m^2}, \tag{3}$$

$$U_{rs} = \frac{H_s}{k_m^2 d^3}, \tag{4}$$

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

$$\alpha = 0.3536 + 0.2568 s_m + 0.0800 U_{rs} \tag{5}$$

$$\beta = 2 - 1.7912 s_m - 0.5302 U_{rs} + 0.2824 U_{rs}^2 \tag{6}$$

3. The distribution of the maximum wave height in a measured storm

A storm can be thought of as a sequence of discrete sea states, σ_i, where the wave heights increase to a peak level and then decrease again. To calculate the distribution of the maximum wave height in a measured storm, we first need to calculate the distribution of the maximum wave height in each sea state. If it is assumed that individual wave heights are independent, then the probability that the maximum wave height in a sea state, H_{max}, does not exceed a level h in an interval Δt, is simply the product of the probability that each individual wave does not exceed h. In the interval Δt, there are $N = \Delta t/T_z$ waves, so

$$\Pr(H_{max} \le h|\sigma_i) = \Pr(H \le h|\sigma_i)^N. \tag{7}$$

In reality, consecutive wave and crest heights are correlated, with the largest waves occurring in groups. However, the assumption of independence in (7) is not restrictive. Krogstad and Barstow (2004) note that for Gaussian processes, there is an analytical theory for the distribution of the maximum value during a given period of time, and that this theory gives identical results to (7). They argue that this observation makes it reasonable to use (7) in the case of nonlinear waves, where in general no analytical theory exists.

The probability that the maximum wave height in a measured storm, MS, does not exceed level h, is calculated as the product of the probabilities that the wave height is not exceeded in any of the sea states over the course of the storm:

$$\Pr(H_{max} \le h|MS) = \prod_{i=1}^{k} \Pr(H_{max} \le h|\sigma_i), \tag{8}$$

where the storm is defined as sea states σ_1, ...,σ_k. There is some subjectivity in how a time series is partitioned into separate storms. The process is analogous to defining a declustering algorithm in the peaks-over-threshold (POT) method (see Section 5 for discussion of the POT method). 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 and that the minimum H_s between two adjacent peaks must be less than some multiple of the lower of the peak. For the present work, storms have been defined using a minimum temporal separation of 5 days between adjacent peaks and a requirement that the minimum value of H_s between adjacent peaks is less than 50% of the lower of the two peaks.

4. Definition of the equivalent storm

In order to analyse the statistics of the maximum wave height in storms in general, the distribution of the maximum wave height in the measured storm, $\Pr(H_{max} \le h|MS)$, must be parameterised in some way. Models for $\Pr(H_{max} \le h|MS)$ are referred to as equivalent storm models and will be denoted $\Pr(H_{max} \le h|ES)$. The quality of the fit of the equivalent storm model will be assessed in terms of the Cramér–von Mises goodness-of-fit parameter, ω, which quantifies the difference between two distributions:

$$\omega^2 = \int_0^{\infty} [\Pr(H_{max} \le h|MS) - \Pr(H_{max} \le h|ES)]^2 dh. \tag{9}$$

Sections 4.1 reviews equivalent storm methods that model the time series of H_s in the storm. Section 4.2 describes the Tromans and Vanderschuren (1995) (TV) method for modelling the distribution $\Pr(H_{max} \le h|MS)$ directly and introduces the generalised equivalent storm method as a generalisation of the TV method.

4.1. Temporal evolution methods

As noted in the introduction, there have been several models proposed for the temporal evolution of H_s over the course of a storm. To derive these models, equation (8) is rewritten as

$$\begin{aligned}\Pr(H_{max} \le h|MS) &= \prod_{i=1}^{k} [\Pr(H \le h|\sigma_i)]^{n_i} \\ &= \exp\left(\sum_{i=1}^{k} \frac{\Delta t}{T_{z,i}} \ln[\Pr(H \le h|\sigma_i)]\right).\end{aligned} \tag{10}$$

The last expression on the RHS of (10) is obtained by using the identity $x.\ y = \exp(\ln(x) + \ln(y))$. As the duration of each sea state, Δt, tends to zero, the summation in (10) can be expressed as an integral:

$$\Pr(H_{max} \le h|MS) = \exp\left(\int_0^{D} \frac{1}{T_z(t)} \ln[\Pr(H \le h|\sigma(t))]\,\mathrm{d}t\right), \tag{11}$$

where D is the duration of the storm (in seconds). The form of the integral in (11) was originally derived by Borgman (1973). The distribution of the maximum wave in the equivalent storm is then obtained by substituting the assumed model for $H_s(t)$ into (11) and changing the differential from $\mathrm{d}t$ to dH_s. For the equivalent power storm (EPS) model the time series of H_s is modelled as:

$$H_s(t) = a\left[1 - \left(\frac{2|t - t_0|}{b}\right)^k\right], \quad t_0 - \frac{b}{2} \le t \le t_0 + \frac{b}{2} \tag{12}$$

where a is the storm peak H_s at $t = t_0$, b is the storm duration and k is the storm shape parameter. The equivalent triangular storm (ETS) is a special case of the EPS model for $k = 1$ and the parabolic storm model is the case $k = 2$. Substituting (12) into (11) gives the distribution of the maximum wave height in the equivalent storm (ES) for the EPS model as:

$$\Pr(H_{max} \le h|ES) = \exp\left(\frac{b}{ka}\int_0^{a} \frac{\ln[P(H \le h|\sigma(H_s))]}{T_z(H_s)}\left(1 - \frac{H_s}{a}\right)^{\frac{1}{k}-1} \mathrm{d}H_s\right). \tag{13}$$

To apply the change in the differential from $\mathrm{d}t$ to dH_s, the variables $T_z(t)$ and $\sigma(t)$ must be changed to models $T_z(H_s)$ and $\sigma(H_s)$, which describe the mean values of T_z and other sea state parameters for a given value of H_s. The fitting of these models is described in Section 7.1.

To fit the EPS model to $\Pr(H_{max} \le h|MS)$, the peak value of H_s in the equivalent storm (a) is defined to be equal to that in the measured storm. The duration of the equivalent storm is defined as the value for which the expected value of the maximum wave height in the equivalent storm is equal to that in the measured storm. The duration is obtained using and iterative process.

In the EPS storm model, there is an additional free parameter, the exponent k, which describes the storm shape. Fedele and Arena (2010) and Arena et al. (2014) fitted the EPS for a range of values of k and compared the return values for each fixed value of k. In the present work, the value of k will be selected for each measured storm as the value which minimises the Cramér–von Mises parameter ω. This approach has been adopted to allow a fairer comparison with the generalised equivalent storm method, which has three free parameters.

In the equivalent exponential storm (EES) model (Laface and Arena, 2016) the storm is separated into growth and decay phases before and after the peak H_s. The peak H_s in the EES is defined to be equal to the peak H_s in the measured storm and duration of the growth and decay phases is determined in the same way as for the ETS and EPS models, by matching $\mathrm{E}(H_{max}|ES)$ with $\mathrm{E}(H_{max}|MS)$ for the growth and decay phases separately. However, Laface and Arena (2016) conclude that although the EES model can provide a better representation of the storm duration than the ETS and EPS models, the separation of the storm into separate growth and decay phases does not change the estimates of return periods of individual wave heights. The EES model will therefore not be considered further here.

Temporal evolution methods have been demonstrated to give a good fit to the distribution of the maximum wave height in the measured storm in most cases. Considering that measured storms have highly irregular shapes and the equivalent storms are fitted using only the first moment of the distribution (8), it is somewhat surprising that modelling a storm using these simple shapes can provide a good match to the measured distributions. Mackay (2012) noted that the reason for the good match is that the order of the sea states in (8) does not affect the calculation. The sea states in a storm can therefore be reordered into an ascending sequence where H_s increases from a minimum to a maximum value. Only the largest few sea states in a storm affect the distribution of the maximum wave height, and the temporal evolution of these largest few sea states of the reordered storm can be well-modelled by a simple shape.

One of the disadvantages with time series methods is that they only model the temporal evolution of H_s and relations between other sea state parameters and H_s must be assumed. Fortunately, it is not necessary to model the temporal evolution of sea state parameters. The distribution of the maximum wave height in a measured storm can be modelled directly, as explained in next section.

4.2. Distribution modelling methods

Instead of modelling the temporal evolution of sea states in a storm, an equivalent storm can be defined as a statistical distribution which matches the distribution of the maximum wave height in the measured storm (8). Tromans and Vanderschuren (1995) took this approach and proposed that (8) could be modelled using a Gumbel distribution in

H^2_{max} with fixed relations between the distribution parameters. The assumption of a fixed form for $\Pr(H_{max} \le h|ES)$ can be relaxed and a more general model can be applied. The generalised extreme value distribution (GEV) is a natural candidate for modelling (8), given that it is the asymptotic form of the distribution of the maximum of a large sample of independent, identically distributed observations (Coles, 2001). The maximum wave heights in each sea state in a storm are effectively independent, but are not identically distributed, as the distribution parameters vary with sea state. However, the GEV provides a flexible formulation, capable of representing a range of distribution shapes, so the main justification for the use of the GEV is in its fit to the data rather than asymptotic arguments. In the generalised equivalent storm method, the distribution of the maximum wave height in the storm is assumed to follow the GEV distribution:

$$\Pr(H_{max} \le h|ES) = \begin{cases} \exp\left(-\left(1 + k\left(\frac{h-a}{b}\right)\right)^{-\frac{1}{k}}\right), & \text{for } k \neq 0 \text{ and } b > 0 \\ \exp\left(-\exp\left(-\left(\frac{h-a}{b}\right)\right)\right), & \text{for } k = 0 \text{ and } b > 0 \end{cases} \tag{14}$$

where a, b and k are the location, scale and shape parameters, respectively. The variable names a, b and k have been reused here for analogy with the temporal evolution models. Using the same notation, simplifies the presentation later and the interpretation should be clear from the context.

The Gumbel distribution is a special case of the GEV when $k = 0$. However, the TV model assumes that H^2_{max} follows a Gumbel distribution rather than H_{max}. The TV model is defined as:

$$\Pr(H_{max} \le h|ES) = \exp\left(-\exp\left(-\left(\frac{h^2 - a}{b}\right)\right)\right). \tag{15}$$

The TV and GEV models are fitted using the mode and standard deviation of the distribution. The mode of the GEV (i.e. the most probable value) is given by:

$$\text{mode}(H_{max}|ES) = \begin{cases} a + b\frac{(1+k)^{-k} - 1}{k} & \text{for } k \neq 0 \\ a & \text{for } k = 0 \end{cases} \tag{16}$$

The standard deviation of the GEV is:

$$\text{std}(H_{max}|ES) = \begin{cases} b\pi/\sqrt{6}, & \text{for } k = 0 \\ b\sqrt{(g_2 - g_1^2)}/k, & \text{for } k \neq 0,\ k < 1/2 \\ \infty & \text{for } k \ge 1/2 \end{cases} \tag{17}$$

where $g_n = \Gamma(1 - nk)$ and Γ is the gamma function.

In the TV model, the location parameter is therefore defined as $a = H^2_{mp}$, where H_{mp} is the most probable maximum wave height in the storm (i.e. the mode of (8)). The scale parameter is defined to be in fixed relation to the location parameter, with $b = a/\ln(N)$, where $N = D/T_z$ is the number of waves in the storm and D is a measure of the duration of the storm. Tromans and Vanderschuren (1995) do not specify how D is defined, but note that $\ln(N) \approx 8$ for Northern European winter storms. The DNV GL guidelines (DNV GL, 2017) recommend the TV model for estimating return periods of individual wave and crest heights, however there is no recommendation for how to calculate $\ln(N)$. Two options for calculating $\ln(N)$ will be considered here. The first method is to use the standard deviation of the distribution, which gives

$$\ln(N) = \frac{\pi H^2_{mp}}{\text{std}(H^2_{max})\sqrt{6}} \tag{18}$$

In the second, the value of $\ln(N)$ is determined numerically by finding the value that minimises the Cramér–von Mises goodness-of-fit parameter, ω.

The GEV model is fitted numerically, by finding the values of a, b and k that minimise ω. Efficient numerical optimisation of three parameters requires a starting point that is close to the optimum. The first guess for the parameters of the GEV is defined by assuming $k = 0$, which then implies $a = H_{mp}$ and $b = \sqrt{6}\ \text{std}(H_{max})/\pi$. In this work, the parameters that minimise ω have been found using the MATLAB function 'fminsearch' which uses a simplex search technique (Lagarias et al., 1998) and first-guess values noted above.

5. The distribution of equivalent storm parameters

Once each measured storm has been fitted with an equivalent storm, the next step is to establish the long-term joint distribution of the equivalent storm parameters. The joint probability density function (PDF) of the equivalent storm parameters, $p(a, b, k)$, can be written as the product of marginal density, $p(a)$ and the conditional density function $p(b, k|a)$:

$$p(a, b, k) = p(a)p(b, k|a). \tag{19}$$

For the ETS and TV models there is no shape parameter k, so the joint pdf of the storm parameters is written as $p(a, b) = p(a)p(b|a)$.

The details of how the conditional PDFs are modelled is dependent on the equivalent storm method used, so the discussion of this will be postponed until examples are presented in Section 7. Here we consider the PDF $p(a)$. In the case of the ETS and EPS models, a is the peak H_s in the storm. In the case of the TV model, a is the square of the most probable maximum wave height in the storm, H^2_{mp}. For the GEV model, it will be shown in Section 7 that, to a very good approximation, $a = H_{mp}$. So, depending on the equivalent storm model used, $p(a)$ is the PDF of either storm peak H_s, H_{mp} or H^2_{mp}. In all cases, the parameter a defines some measure of the peak intensity of the storm. Two methods for estimating $p(a)$ will be described here. The first approach, described in Section 5.1, uses a peaks-over-threshold (POT) analysis, and is applicable to both the time series models (ETS/EPS) and the distribution models (TV/GEV). The second method, described in Section 5.2, is only applicable to the time series models and is derived from the mean time when H_s is above a threshold level in the equivalent storm.

5.1. The distribution of storm peak intensity using the POT method

The POT method is now commonly adopted in extreme value analysis for ocean engineering (see e.g. Coles, 2001; Jonathan and Ewans, 2013). In the POT method, the generalised Pareto distribution (GPD) is fitted to observations exceeding a high threshold. The CDF of the GPD is defined conditional on a exceeding some high threshold u:

$$\Pr(a \le x|a > u) = \begin{cases} 1 - \left(1 + \xi\left(\frac{x-u}{\sigma}\right)\right)^{-\frac{1}{\xi}} & \text{for } \xi \neq 0 \text{ and } \sigma > 0, \\ 1 - \exp\left(-\left(\frac{x-u}{\sigma}\right)\right) & \text{for } \xi = 0 \text{ and } \sigma > 0. \end{cases} \tag{20}$$

The parameters σ and ξ are the scale and shape parameters respectively. When $\xi \ge 0$, 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 method used to estimate the parameters of the GPD can have a significant influence on results, especially with relatively small sample sizes. 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. Hosking and Wallis (1987) showed that the maximum likelihood (ML) estimators for the GPD are non-optimal for sample sizes up to 500, with higher bias and variance than other estimators, such as the moments (MOM) and probability weighted moments (PWM) 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. Dupuis (1996) found that the MOM and PWM estimators are sensitive to threshold choice and sometimes result in non-feasible estimates, where $\xi < 0$ so that the estimated distribution has a maximum value of $u - \sigma/\xi$, but the maximum observation is larger than this value for the estimated parameters. A comparison of estimators for the GPD was presented by Mackay et al. (2011). The authors concluded that a modified version of the likelihood-moment (LM) estimator (Zhang, 2007), using an initial estimate of the shape parameter from the hybrid PWM method of Dupuis and Tsao (1998), gave the lowest bias and variance in estimates of high quantiles. Subsequent work has improved the estimation of parameters for the GPD. Kang and Song (2017) provide a quantitative comparison of more recent estimation methods for the GPD and conclude that the empirical Bayesian method (EBM) of Zhang (2010) performs best in a wide range of cases. The EBM method will be used here.

The threshold for the GPD is selected by fitting the distribution for a range of threshold values and selecting the threshold as the lowest value for which the shape parameter of the distribution and estimates of high quantiles converge to steady values (see Coles, 2001, for further details).

It should be noted that the distribution of storms where the peak intensity does not exceed the threshold is not modelled, as it does not influence the return values at high return periods. The occurrence rate of storms where the peak intensity exceeds the threshold is accounted for in the calculation of return periods, described in the next section.

5.2. The distribution of storm peak intensity for temporal evolution methods

For the temporal evolution methods, an alternative method has been proposed for estimating the PDF $p(a)$, as a function of the CDF of H_s (denoted $\Pr(H_s < a)$). By considering the mean time that $H_s > h$ in an equivalent storm, it can be shown that

$$p(a) = g(\bar{b}(a), \nu, \Pr(H_s < a)), \tag{21}$$

where $\bar{b}(a)$ is the mean storm duration as a function of a, ν is the occurrence rate of storms and the function g is dependent on the time series model assumed (see Arena and Pavone, 2006; Fedele and Arena, 2010 and Laface and Arena, 2016 for the specific forms of g for the ETS, EPS and EES models respectively). In the examples presented in the references above, the authors assume that $\Pr(H_s < a)$ follows a Weibull distribution.

The problem with this approach is that 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 and crest heights. This point has been well illustrated by Ferreira and Guedes Soares (1999), who fit three types of model to the distribution of H_s, where the three models have tails which tend to different forms of the GPD, with shape parameters $\xi < 0$, $\xi = 0$ and $\xi > 0$ respectively. They show that all three types of distribution fit H_s data from the Portuguese coast very well, with Kolmogorov–Smirnov tests for goodness-of-fit not rejecting any of the distributions as not fitting the data. The three distributions are very close to each other over the range of the bulk of the data but differ remarkably at high probability quantiles. This leads to estimates of 100-year return values of H_s differing by over 5 m between the models with $\xi < 0$ and $\xi > 0$. Mathiesen et al. (1994) noted that since there is no a priori reason to suppose that H_s follows one type of distribution rather than another, fitting a model to the entire range of H_s is not a reliable method for estimating extremes. Instead, Mathiesen et al. recommend the POT method, since in this method the quality of the fit is dependent on only the extreme values and not the bulk of the data. The POT/GPD method is also justified by asymptotic arguments (Coles, 2001).

Laface and Arena (2016) note that to improve the fit of the model for $\Pr(H_s < a)$, a piecewise model could be adopted, with e.g. a log-normal model for the bulk of the distribution and a Weibull model for values above some high threshold level. This introduces further complication into the calculation of $p(a)$, since a smooth transition from the model for the bulk to the tail must be ensured so that $\Pr(H_s < a)$ has a continuous derivative (the derivative is used in the function g). Whilst this is possible to achieve, it seems unnecessarily complicated compared to the POT/GPD model for $p(a)$.

Estimating $p(a)$ from $\Pr(H_s < a)$ in fact provides no new information about the distribution of a. In the case of ETS and EPS methods, $p(a)$ is the density function of storm peak H_s. The additional information provided in $\Pr(H_s < a)$ compared to $p(a)$ is related to the distribution of lower values of H_s in a storm, given the peak value a. The use of (21) to estimate $p(a)$ is therefore not expected to result in a lower variance than the POT/GPD method.

Laface et al. (2016) examined the sensitivity of estimates of $p(a)$ from the POT/GPD method and (21) to the choice of threshold level. Laface et al. (2017) found that return values of storm peak H_s calculated from $p(a)$ are more sensitive to threshold level when $p(a)$ is estimated from the POT/GPD method than when it is estimated from (21). However, the authors do not state what method was used to estimate the GPD parameters in their study. As noted in the preceding section, commonly used estimators for the GPD such as the maximum likelihood or probability weighted moments estimators are not optimal, especially for smaller sample sizes, and give a higher sensitivity to threshold level than other estimators such as the EBM estimator used in this study. Without knowing the GPD parameter estimation method used by Laface et al. (2017), it is not possible to conclude whether the use of (21) to estimate $p(a)$ is more stable with threshold level than the POT/GPD method in general, or whether their findings were a result of using non-optimal estimators for the GPD.

Given that the use of (21) to estimate $p(a)$ appears to offer several disadvantages and no clear advantages over the POT/GPD method, the POT/GPD has been adopted in this study for both the time series equivalent storm models (ETS, EPS) and the distribution equivalent storm models (TV, GEV).

6. Return periods of individual wave heights

The distribution of the maximum wave height in a random storm where the peak intensity, a, exceeds the threshold value, u, is calculated by integrating the short-term distribution of the maximum wave in a storm specified by the parameters a, b and k over the long-term joint distribution of a, b and k:

$$\Pr(H_{max} \le h|RS) = \int_k \int_b \int_a p(a)p(b, k|a)\Pr(H_{max} \le h|ES)\,\mathrm{d}a\,\mathrm{d}b\,\mathrm{d}k \tag{22}$$

For the ETS and TV models, the expression above reduces to a double integral as there is no shape parameter, k.

Once the distribution $\Pr(H_{max} \le h|RS)$ has been calculated, the return values of individual wave heights can be calculated. The wave height that is exceeded once every m storms on average, is the solution of:

$$\Pr(H_{max} \le h|RS) = 1 - \frac{1}{m}. \tag{23}$$

The wave height that is exceeded once every T years on average, is known as the T-year return value and is said to have a return period of T years. If there are on average ν storms exceeding the threshold level every year, then the T-year return value, H_T, is the solution of:

$$\Pr(H_{max} \le H_T|RS) = 1 - \frac{1}{\nu T}. \tag{24}$$

The value ν is estimated as $\nu = n/\tau$ where τ is the length of the dataset in years and n is the number of storms in the dataset which exceed the threshold value used in the POT analysis.

7. Example – application to measured data

The equivalent storm models discussed in Section 4 are compared using two long datasets of wave buoy measurements. Return periods of individual crest heights are calculated using the short-term crest height distribution of Forristall (2000), described in Section 2. The datasets used are described in Section 7.1. The fit of the equivalent storm models to the distribution of the maximum crest height in measured storms is compared in Section 7.2. The joint distributions of equivalent storm parameters are described in Section 7.3 and return periods from each model are compared in Section 7.4. Finally, the reduction of the double or triple integral (22) to a single integral for each model is discussed in Section 7.5.

7.1. Datasets

Two datasets have been selected for comparison of the equivalent storm models, representing two distinct wave climate regimes. The first dataset 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. The other dataset 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. Scatterplots of H_s against T_z for the two datasets are shown in Fig. 1. Both datasets have similar maximum observed H_s, with a maximum H_s at buoy 44025 of 9.64 m and a maximum at buoy 46014 of 10.34 m. In the shallower water site, the maximum values of H_s occur for the steepest seas with $0.06 < s_z < 0.07$, whilst at the deeper site the extremes H_s occurs for a wider range of steepness, with $0.03 < s_z < 0.06$. The range of T_z observed at the Californian site is also much larger than at the New York site.

To apply the time series equivalent storm methods, a model is needed for the mean value of T_z and the assumed short-term distribution parameters as a function of H_s. Fig. 2 shows the values of T_z and the Forristall (2000) distribution parameters α and β for individual sea states and the mean values, binned by H_s. The shallower location, Buoy 44025, shows a greater variation in α and β with H_s than Buoy 46014, due to the increase in nonlinear effects related to wave steepness and water depth (the parameters α and β are defined as functions of steepness and Ursell number – see Section 2). The mean values have been modelled as linearly dependent on H_s in sea states with $H_s > 2$m. The reason for the distinct change in the variation of α and β around $H_s = 2$m is not clear, however the values of α and β in lower sea states are irrelevant for extremes and will not be considered further here. The linear models provide a reasonable fit to the observed mean values for $H_s > 2$m. It should be noted that although the short-term distribution must be evaluated to calculate $\Pr(H_{max} \leq h|ES)$, the accuracy of the prediction of return values of wave heights is dependent on how closely $\Pr(H_{max} \leq h|ES)$ matches $\Pr(H_{max} \leq h|MS)$ and not the accuracy of the model for the variation of the distribution parameters with H_s. As $\Pr(H_{max} \leq h|ES)$ is fitted to $\Pr(H_{max} \leq h|MS)$, the fitted storm duration (and storm shape in the case of the EPS model) can compensate to some extent for inaccuracies in the models for $\overline{T_z}(H_s)$, $\overline{\alpha}(H_s)$ and $\overline{\beta}(H_s)$. The complications involved in the indirect modelling of $\Pr(H_{max} \leq h|MS)$ in the ETS and EPS models highlight the advantages of the TV and GEV models, which do not require this preliminary step.

7.2. Accuracy of equivalent storm models

For each dataset, independent storm peaks were identified, defined as local maxima in H_s separated from adjacent maxima by a minimum of 5 days and a requirement that the minimum value of H_s between two adjacent peaks is less than 50% of the lower of the two peaks. The time series of sea states was divided into discrete sections, with the dividing points located at the time of the minimum H_s between each storm peak. These blocks were defined as constituting a storm. This resulted in 780 storms for buoy 44025 and 1036 storms for buoy 46014.

The distribution of the maximum crest height in each storm was calculated and the ETS, EPS, TV and GEV equivalent storm models were fitted using the methods described in Section 4. For the TV model the parameters were estimated using both the moment method and numerical optimisation method described in Section 4.2. The results from these methods are denoted TV(mom) and TV(num) respectively.

Examples of the fits of the equivalent storm models to the largest measured storm in each dataset are shown in Figs. 3 and 4. The upper panels show the measured time series of H_s in the storm and the fitted ETS and EPS, together with the same series sorted into descending order. The ETS approximates a linear fit to the highest values of H_s in the storm, whilst the EPS provides a slightly better fit using a nonlinear variation with time. The density functions of the highest crest in the storm are shown in the lower left panels. All the equivalent storm methods provide a good approximation to the shape of the density functions.

To highlight the differences between the methods, the lower right panels show the ratio between the quantile of the equivalent storm and the measured storm at a given probability level. If the equivalent storm was a perfect fit to the measured storm, then this ratio would be equal to one for the entire probability range. The trends for both the examples are similar. The GEV model has the lowest bias in the quantiles, with quantiles with exceedance probabilities below 10^{-2} being slightly underestimated. The EPS model provides the next best fit to the measured storm, with a very similar performance to the GEV for the example in

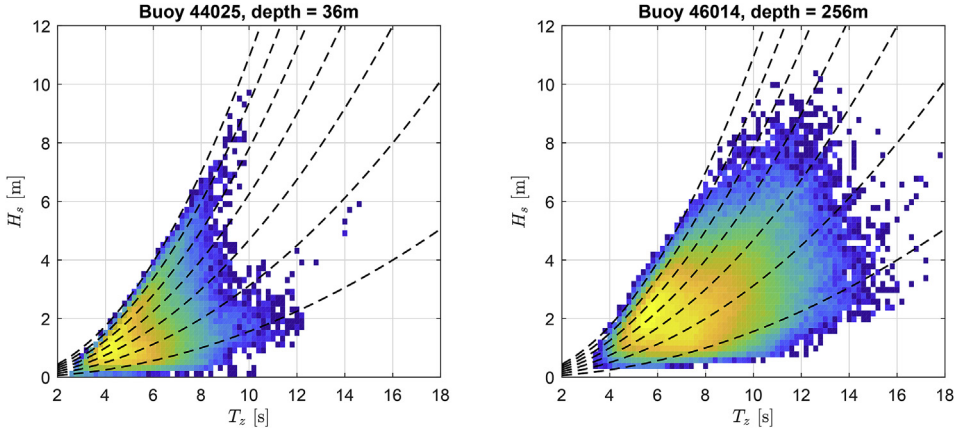

Fig. 1. Scatter plots of H_s against T_z for the two datasets used in this study. Colour indicates density of occurrence. Black dashed lines indicate s_z from 0.01 to 0.07 at intervals of 0.01. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

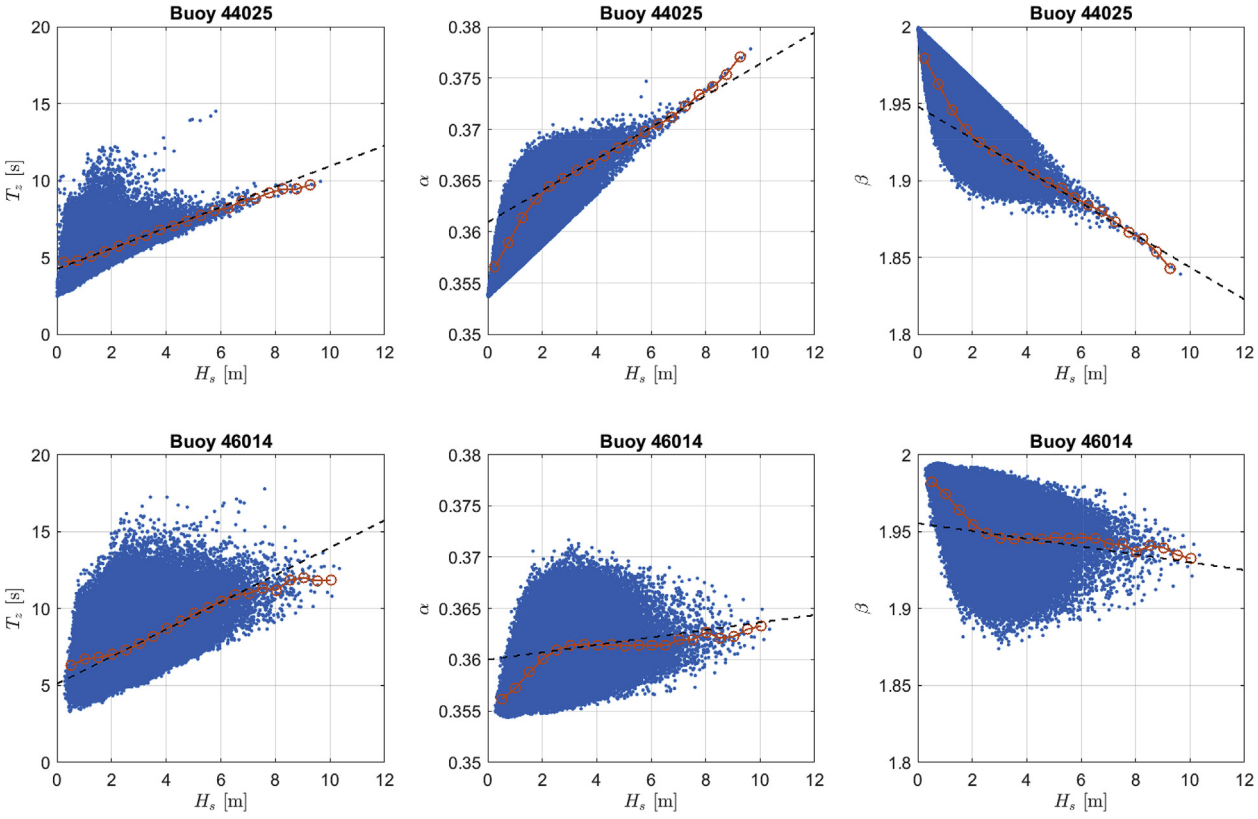

Fig. 2. Mean value of T_z and Forristall (2000) distribution parameters α and β as a function of H_s. Blue dots show values for individual sea states. Red circles are mean values binned by H_s. Black lines are linear models for the mean value as a function of H_s in sea states with $H_s > 2$m. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

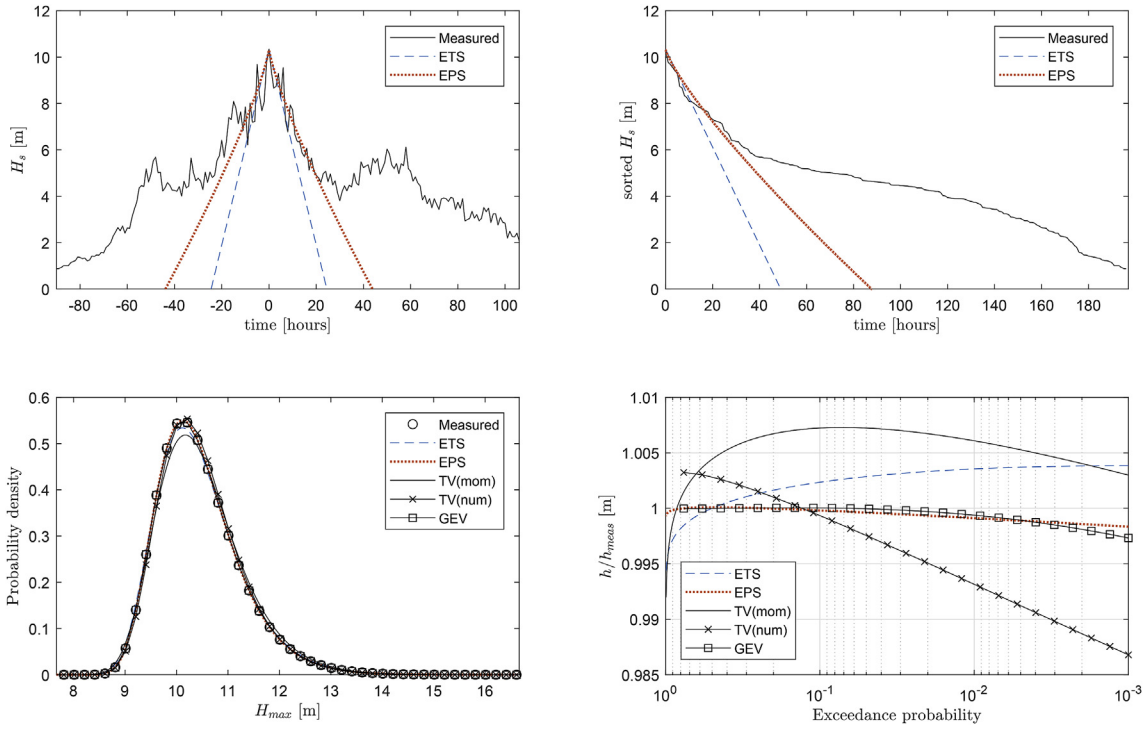

Fig. 3. Fit of the equivalent storm models for the largest storm observed at Buoy 46014. Upper left: Measured time series of H_s and fitted ETS and EPS models. Upper right: Time series of H_s sorted in descending order. Lower left: density function of maximum crest height in the storm. Lower right: Ratio of quantiles from the distributions in the measured and equivalent storms.

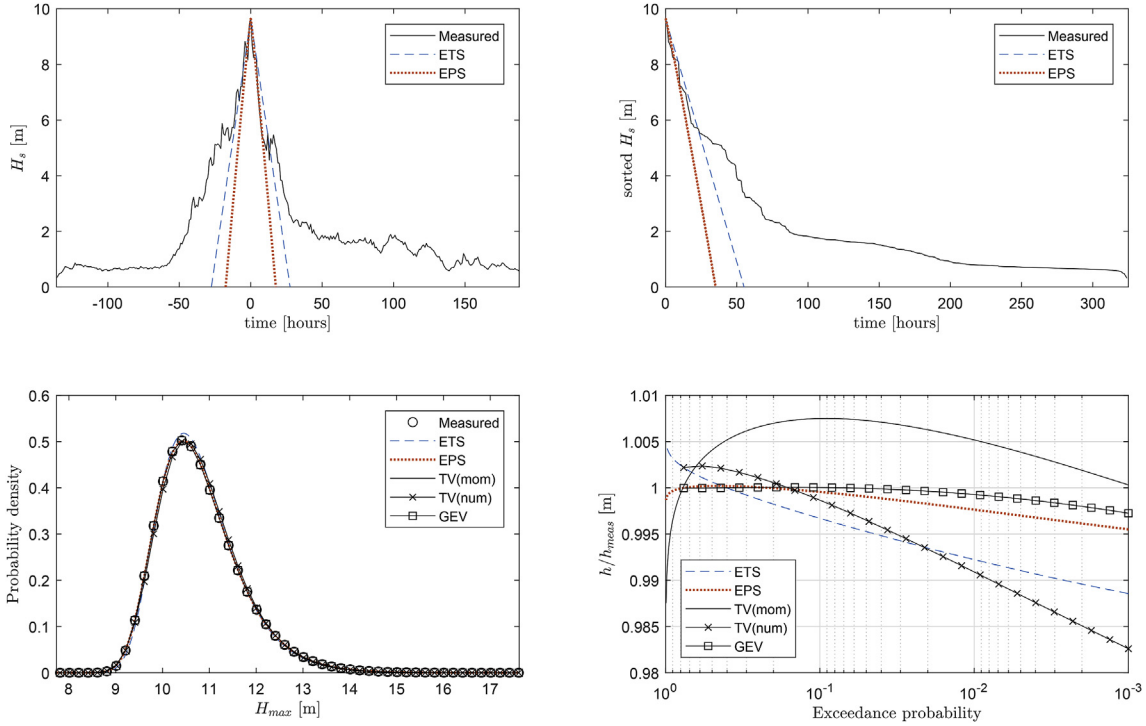

Fig. 4. As Fig. 3, but for the largest storm observed at Buoy 44025.

Fig. 3 and a slightly higher bias for the example shown in Fig. 4. The ETS model is biased high in the example shown in Fig. 3 and biased low for the example in Fig. 4. The TV model has the largest deviation from the measured storm, exhibiting the same pattern in the two examples. The TV model estimated using the moment method is biased high over the range of exceedance probabilities shown, whereas the TV model fitted numerically is biased high for exceedance probabilities less than 10^{-1} and low for the rest of the range.

The mean and standard deviation of the quantile ratios over all measured storms are shown in Figs. 5 and 6 for the two datasets. The trends in the biases (i.e. the mean value of the ratios) are consistent between the two datasets and similar to those shown in the examples for individual storms in Figs. 3 and 4. The GEV has the lowest bias and standard deviation of all models, indicating a consistently good fit to the measured storms. The EPS model has the next best performance, with a tendency to be biased slightly low at high probability quantiles, but with a low STD, indicating consistent performance. The ETS model has a lower bias than the EPS model for buoy 46014, but a slightly higher bias than the EPS model for buoy 44025. The ETS model has the highest STD of all the models, indicating the most variable quality of fit to the data. The TV models have a low STD indicating that the biases compared to the measured storms are consistent, following the same patterns as in the examples in Figs. 3 and 4.

The distribution of the Cramér–von Mises goodness-of-fit parameter, ω, over all measured storms is shown in Fig. 7 for each dataset. As with the quantile ratios, the GEV model has the best performance in terms of ω, showing almost an order of magnitude reduction in the mean value of ω compared to the EPS model. The TV model based on the moment estimates is the worst performing, almost two orders of magnitude higher ω than the GEV model. The ETS and numerically fit TV models have similar performance, but with the ETS model showing slightly better fit for buoy 44025. No correlation between ω and storm peak H_s was found for any of the models at either site, with the GEV having the best performance for all conditions.

It is not surprising that the GEV and EPS models provide better fit than ETS and TV models, considering that the former have three free parameters to describe the distribution, and the latter have two. However, the findings show that GEV method is more accurate than time series models as well as being simpler. The TV model based on the moment estimators is the simplest model but worst performing. The

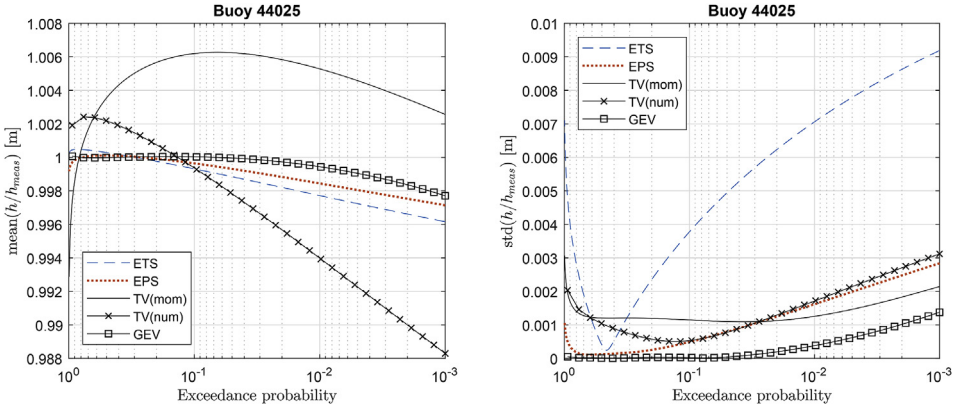

Fig. 5. Mean and STD of quantile ratios for Buoy 44025.

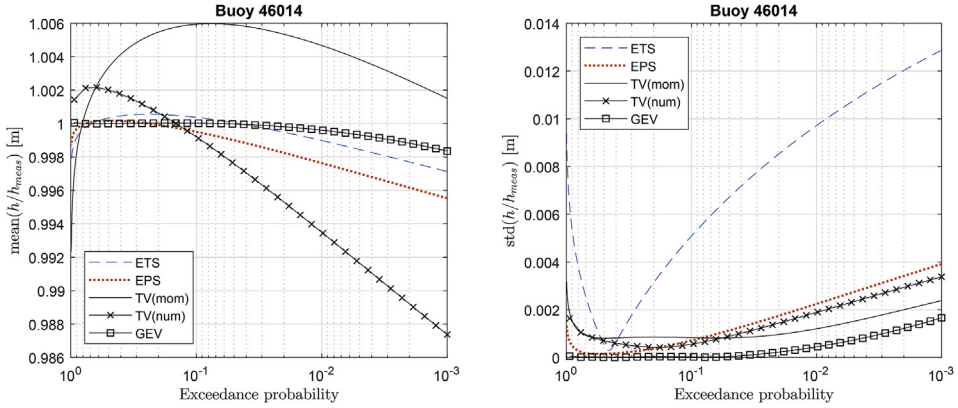

Fig. 6. Mean and STD of quantile ratios for Buoy 46014.

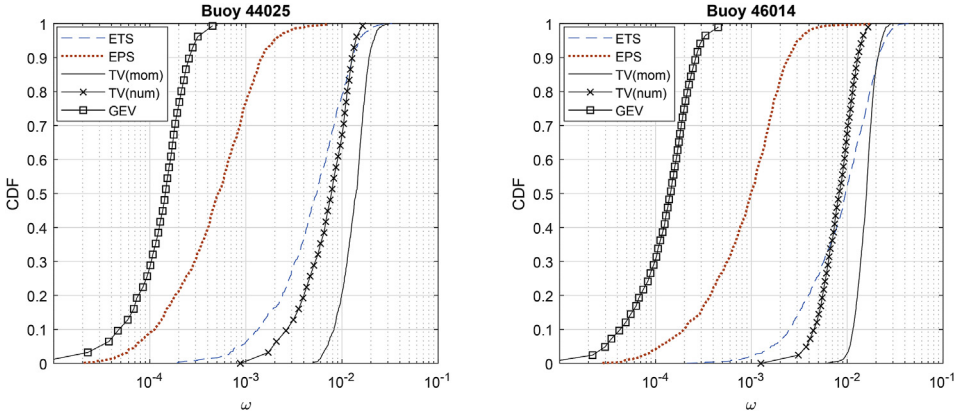

Fig. 7. Distribution of goodness of fit parameter ω for each equivalent storm model.

GEV model is only marginally more complex to implement than the numerical TV model, but shows a large improvement in the accuracy.

7.3. Distributions of equivalent storms parameters

7.3.1. ETS parameter relations

The relationship between the fitted ETS height, a, and duration, b, is shown in Fig. 8 for each dataset. There is an approximately linear log-log relation, giving a model for the parameters as $\ln(b) = A + B\ln(a)$ or equivalently $b = \exp(A)a^B$. The residuals of the model for b, defined as $r = \ln(b) - (A + B\ln(a))$ are approximately normally distributed about the regression line. This gives a model for $p(\ln(b)|a)$ as

$$p(\ln(b)|a) = N(\mu(a), \sigma) \tag{25}$$

where $\mu = A + B\ln(a)$ and σ is assumed to be constant.

The threshold for the GPD fit to the ETS heights a (storm peak H_s) was selected as 5 m for both datasets. For Buoy 44025 this gave 54 storms exceeding the threshold value and 304 storms exceeding the threshold for Buoy 46014. The ordered sample of threshold exceedances is denoted $a_{(1)} \leq a_{(2)} \leq ...\leq a_{(n)}$, where n is the number of storms exceeding the threshold value. The empirical non-exceedance probabilities of the threshold exceedances are defined as

$$\widetilde{P}_{(k)} = \Pr(a_{(k)} < x|a_{(k)} > u) = \frac{k}{n+1}. \tag{26}$$

The fit of the GPD is compared to the empirical exceedance probabilities in Fig. 9, for each dataset. The GPD appears to be very good fit to the data for Buoy 46014, whereas for buoy 44025 there are two large observations that lie further from the fitted GPD curve. One of these large storms was Hurricane Sandy in October 2012, with a peak H_s of 9.6 m. The other large storm was an extratropical cyclone, which occurred in December 1992, with a peak H_s of 9.3 m. All other storms at Buoy 44025 had peak H_s less than 7.5 m. These two large storms result in a positive estimate for shape parameter for the GPD of $\xi = 0.089$. For Buoy 46014, the greater number of observations close to the maximum result in a negative estimate for the shape parameters of the GPD of $\xi = -0.157$.

Confidence bounds can be added to the empirical exceedance probabilities as follows. As each storm is effectively an independent observation from a common population, the empirical non-exceedance probabilities follow a Beta distribution (David and Nagaraja, 2003):

$$f(\widetilde{P}_{(k)}) = \frac{n!}{(k-1)!(n-k)!}(\widetilde{P}_{(k)})^{k-1}(1-\widetilde{P}_{(k)})^{n-k} \tag{27}$$

where $f(\widetilde{P}_{(k)})$ is the PDF of $\widetilde{P}_{(k)}$. A 95% confidence interval for the empirical exceedance probabilities is shown on Fig. 9. The fitted GPD falls within the confidence interval for the exceedance probability of the two largest observations, indicating that the deviation is within the range expected from sampling variability.

7.3.2. EPS parameter relations

The relations between the parameters of the fitted EPS models are shown in Figs. 10 and 11 for the two datasets. There is a weak linear relationship is observed between $\ln(a)$ and $\ln(b)$ for Buoy 44025, but with a higher level of scatter than for the ETS model. For Buoy 46014 the mean value of $\ln(b)$ is approximately constant with $\ln(a)$ and there is a higher level of scatter than for buoy 44025. For both datasets the residuals $r = \ln(b) - (A + B\ln(a))$ were found to be better represented by a Student-t distribution than a normal distribution in this case, due to the larger scatter about the regression line. There is a weak relationship between storm shape, k, and storm height, a, although there is a high level of scatter. The storm shape parameter k is negatively

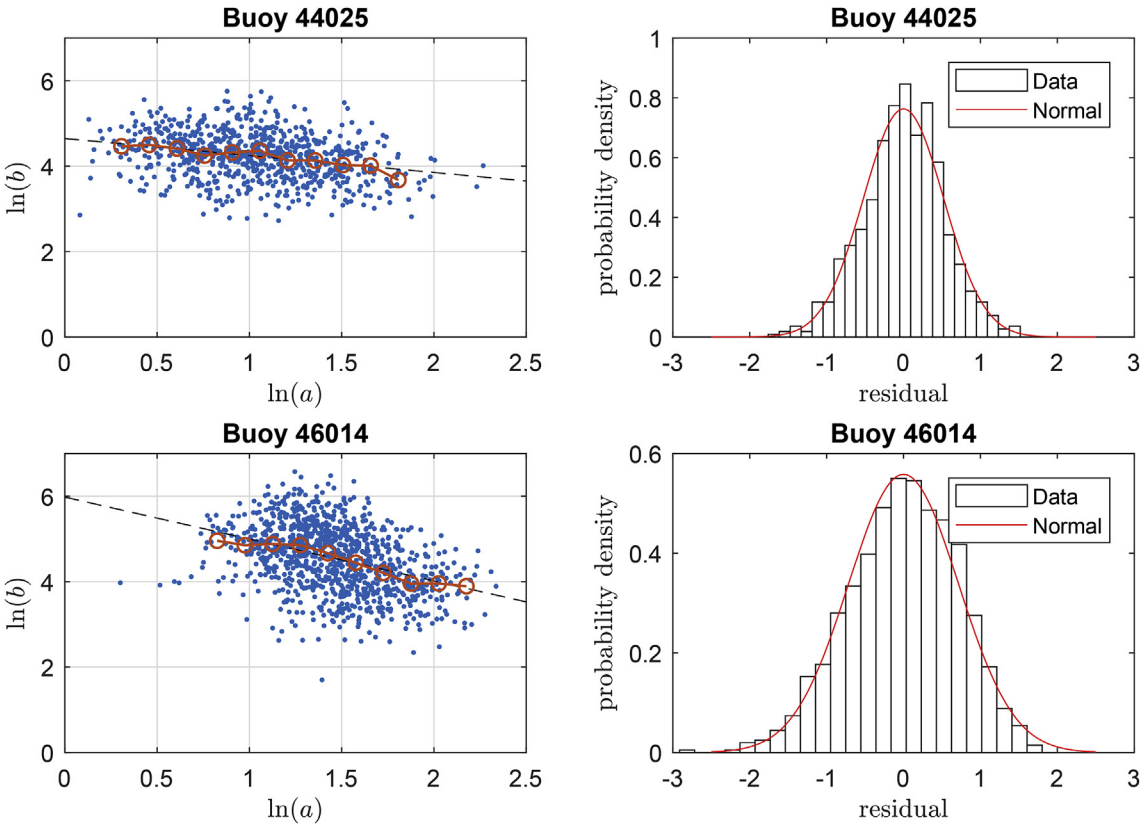

Fig. 8. Left: Regression of storm duration b on storm peak intensity a for the ETS model. Blue dots indicate individual storms, red circles indicate mean values of $\ln(b)$binned by $ln(a)$, black dashed line indicates linear log-log fit. Right: distribution of residuals from linear fit. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

correlated to $\ln(b)$, since lower values of k result in a sub-linear decrease in H_s from the peak value and therefore result in wider bases for the storms.

Due to the strong correlation between k and $\ln(b)$, the conditional joint density function $p(b, k|a)$ will be replaced by a single density function $p(b|a)$, modelled in the same way as for the ETS model, but with a Student-t distribution replacing the normal distribution. The storm shape k is assumed to be linearly related to $\ln(b)$ (i.e. the observed scatter about the regression line is not modelled).

Note that model for the density function $p(a)$ is identical to that for the ETS model, since the storm peak heights, a, are the same for both models.

7.3.3. TV parameter relations

As the TV model fitted numerically performed better than the TV model fitted using the moment estimators, only the numerically fit TV model will be considered in this section. The relationship between the scale and location parameters of the TV model is shown in Fig. 12. As, $a = H_{mp}^2$ and $b = H_{mp}^2/\ln(N)$ by definition, the plot shows $\sqrt{a}$ against $\sqrt{b}$. Both datasets exhibit a very similar relationship, with a strong linear correlation and low levels of scatter. Linear regression for both datasets gives a slight negative intercept for the regression line and very similar slopes.

The relationship between $\ln(N)$ and H_{mp} is shown in Fig. 13. There is a slight negative correlation between $\ln(N)$ and H_{mp}. This is related to the negative intercept in the regression line between $\sqrt{a} = H_{mp}$ and $\sqrt{b} = H_{mp}/\sqrt{\ln(N)}$. Tromans and Vanderschuren (1995) assumed a constant value of $\ln(N) = 8$ in there model. It is evident that the assumption of constant $\ln(N)$ is not appropriate for the datasets examined here, but the mean values are close to 8 for both datasets. It our model it

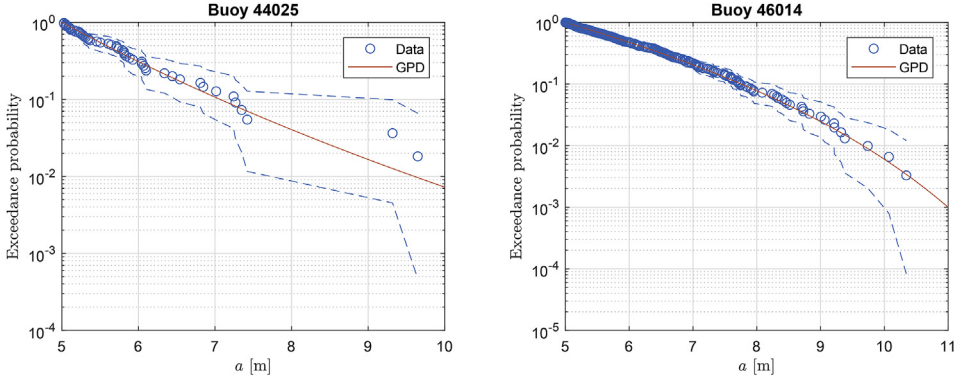

Fig. 9. Empirical exceedance probabilities and fitted GPD for ETS triangle heights, a. Dashed lines indicate 95% CI for the empirical exceedance probabilities.

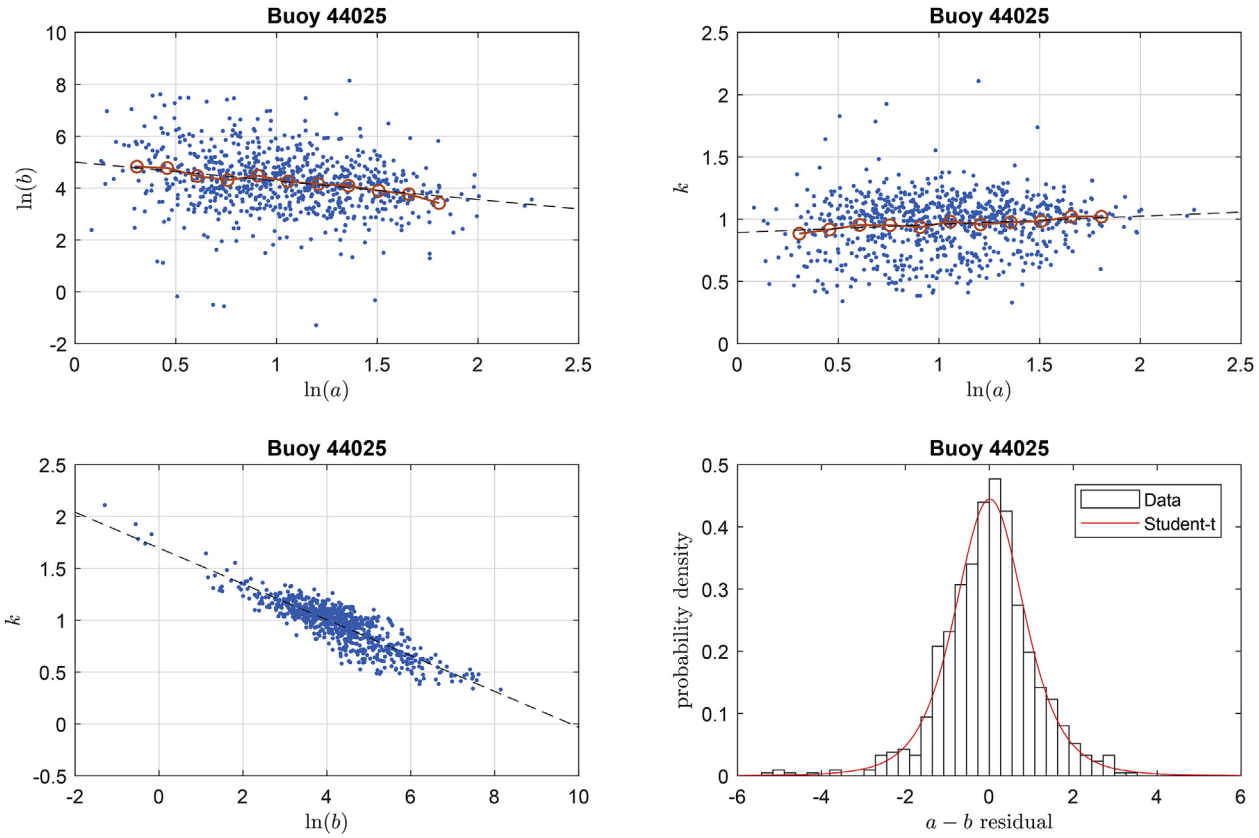

Fig. 10. Scatter plots of relations between EPS model parameters. Blue dots indicate individual storms. Red circles indicate mean values binned by the variable on the x-axis. Black dashed line indicates linear fits. Lower right: distribution of residuals from linear fit between $\ln(a)$ and $\ln(b)$. Data from buoy 44025. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

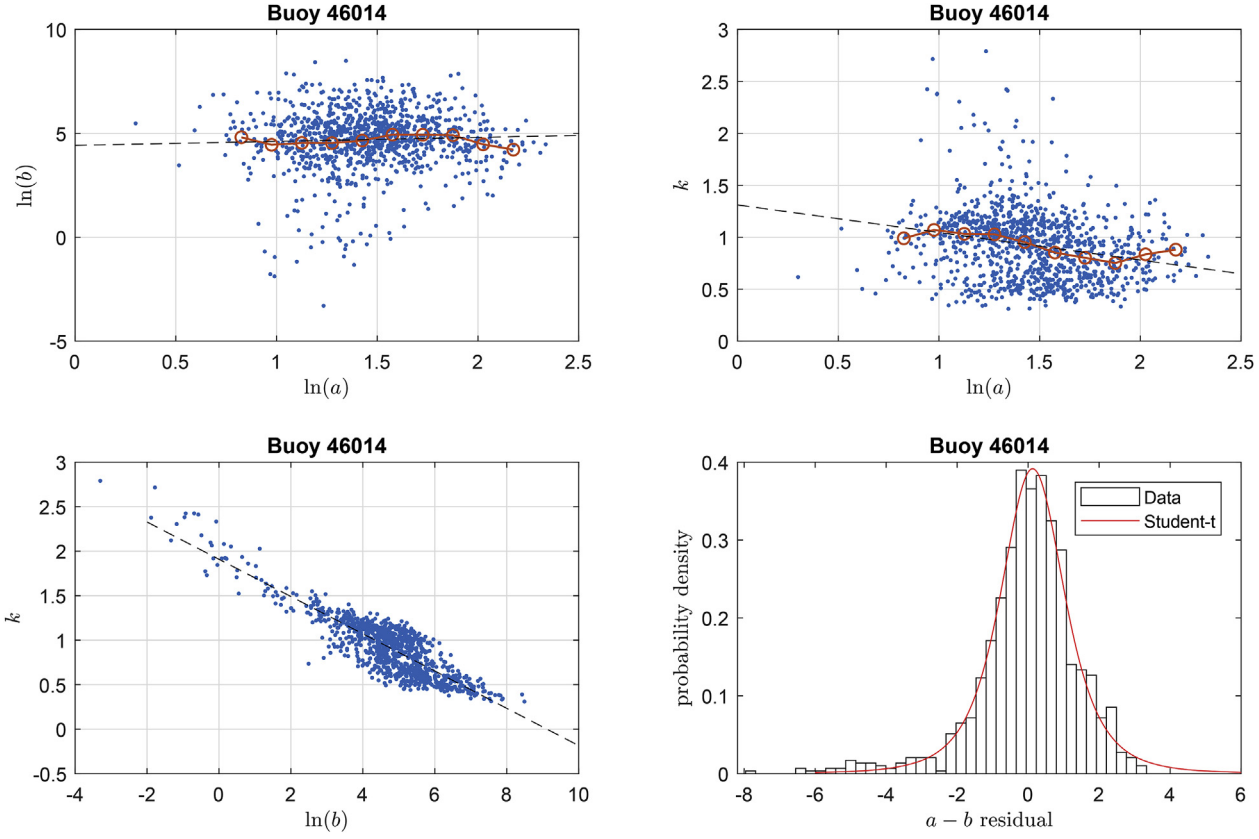

Fig. 11. As Fig. 10, but with data from Buoy 46014.

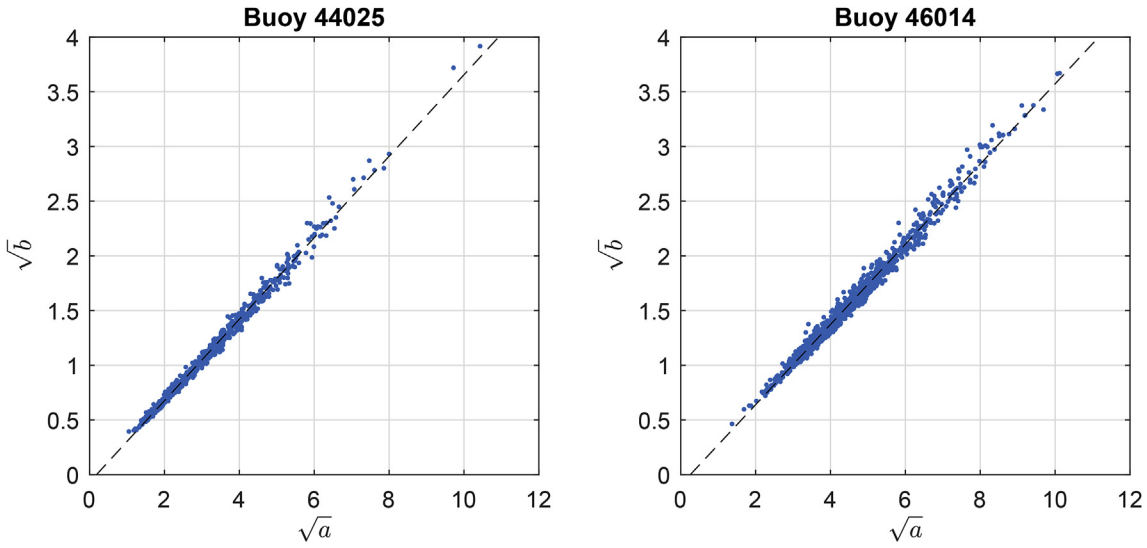

Fig. 12. Scatter plot of square roots of scale and location parameters for the TV model. Black dashed lines are linear fit.

will be assumed that $\ln(N)$ is linearly related to H_{mp}, with the residuals distributed normally about the regression line (see RHS of Fig. 13).

For the TV model the GPD has been fitted to H_{mp} rather than the location parameter $a = H_{mp}^2$, which was also the method adopted by Tromans and Vanderschuren (1995). The integral in (22) over $p(a)$ is replaced by an integral over $p(H_{mp})$. It will be shown in the next section that the GEV parameter a is almost identical to H_{mp}, so the same GPD model can be used for both the TV and GEV methods. As discussed below, there is an almost one-to-one relationship between H_{mp} and storm peak H_s. Therefore the same threshold values have been used for the GPD fit to H_{mp} as for the ETS/EPS models. The fit of the GPD is shown in Fig. 14, which displays very similar behaviour to the fit for the ETS/EPS model, with estimated shape parameters of $\xi = 0.102$ and $\xi = -0.157$ for buoys 44025 and 46014 respectively.

Fig. 15 shows a scatter plot of H_{mp} against storm peak H_s (i.e. ETS/EPS parameter a). The two variables are strongly correlated with an almost 1:1 relationship. The slightly higher values of H_{mp} observed for Buoy 44025 are related to nonlinear effects caused by shallow water. To

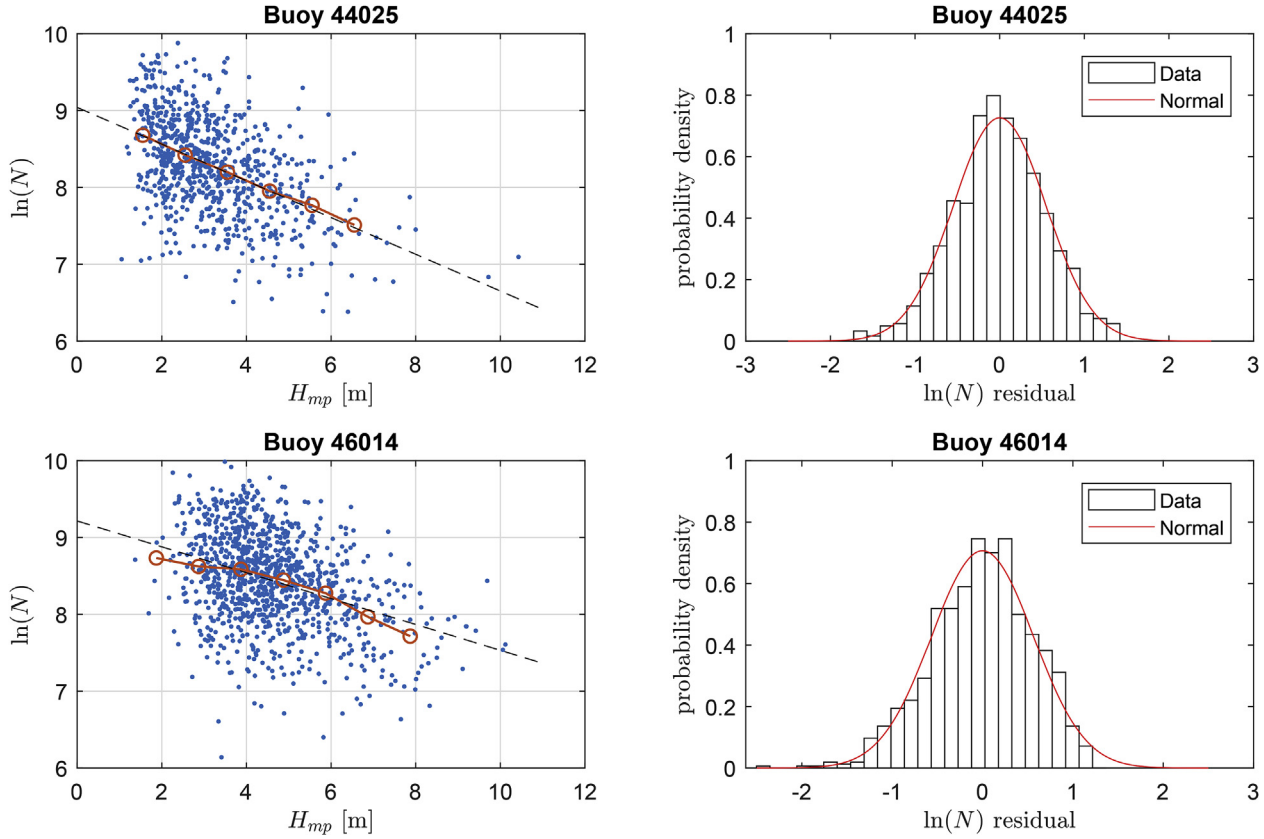

Fig. 13. Left: Scatter plots of $\ln(N)$ against H_{mp} for the TV model. Blue dots indicate relations for individual storms. Red lines are mean values binned by H_{mp}. Black dashed lines are linear fit. Right: Distribution of residuals about regression line and fitted normal distribution. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

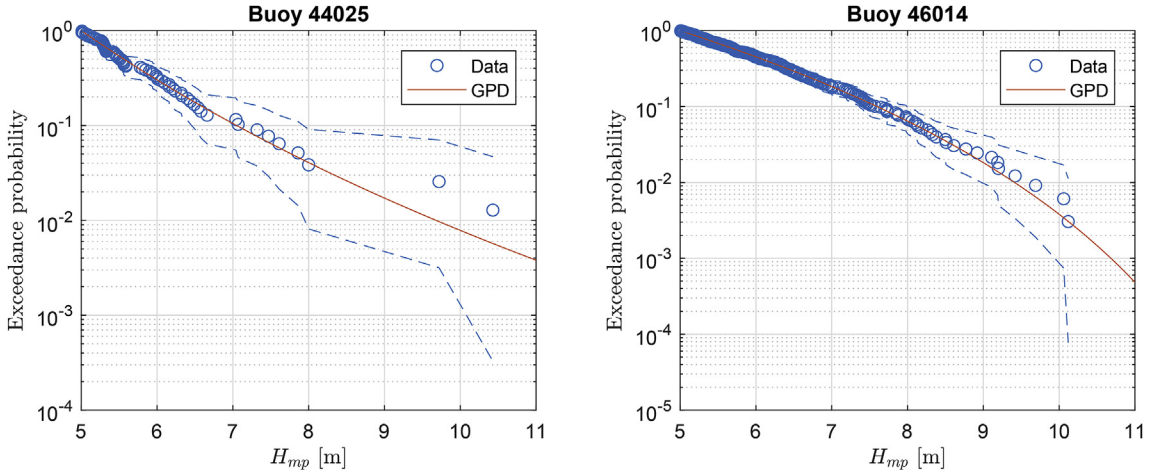

Fig. 14. Empirical exceedance probabilities of H_{mp} and fitted GPD, a. Dashed lines indicate 95% CI for the empirical exceedance probabilities.

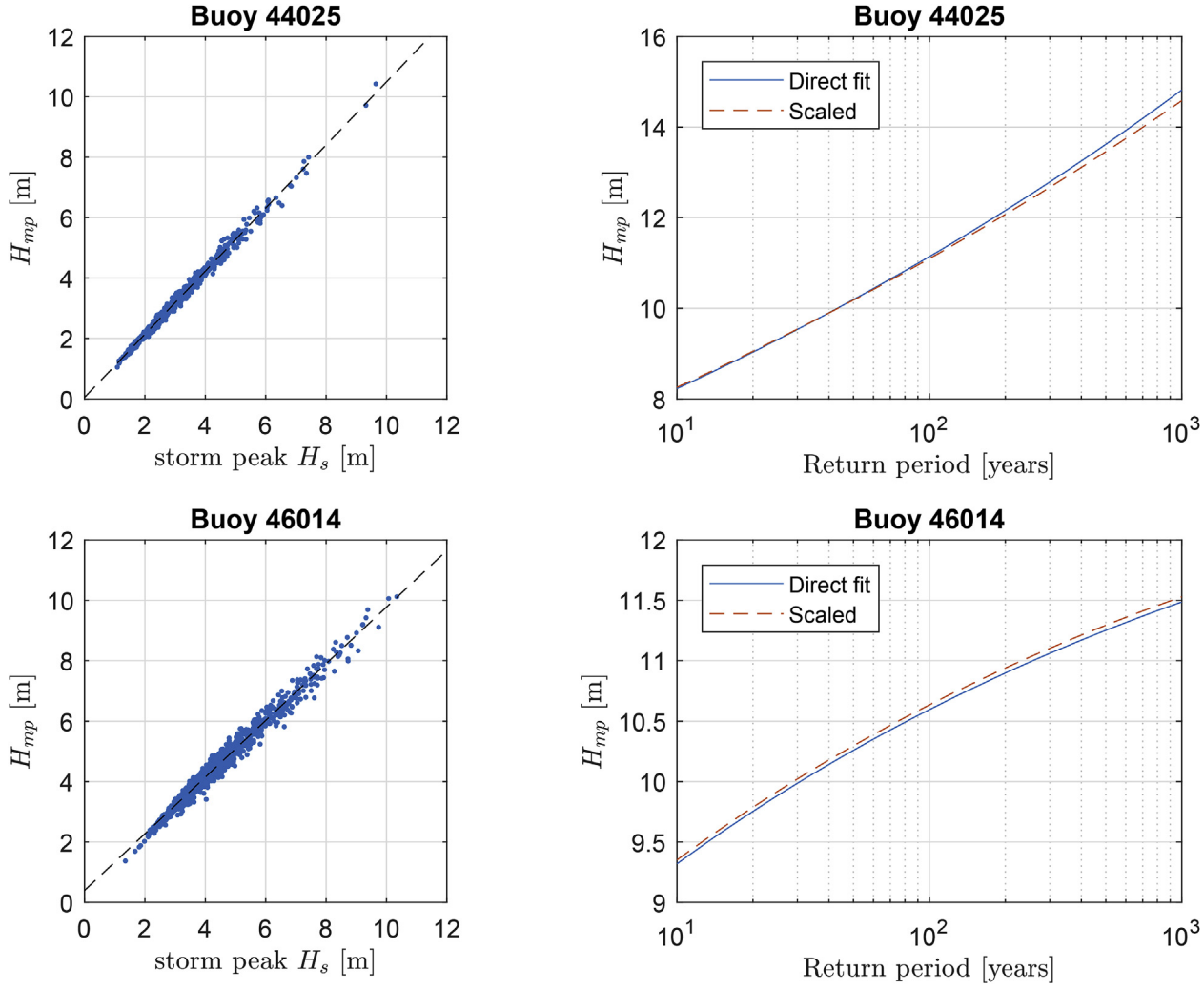

Fig. 15. Left: Linear regression of H_{mp} on storm peak H_s. Right: Return periods of H_{mp} from (a) fit of GPD direct to H_{mp} and (b) fit of GPD to storm peak H_s, scaled by linear relation shown on LHS.

compare the fitted GPD for the ETS/EPS and TV/GEV models, return values of H_{mp} have been calculated from the GPD model for H_{mp} and from the GPD model for storm peak H_s, scaled by the linear relationship shown in Fig. 15. The results shown on the RHS of Fig. 15 indicate that the two models are in good agreement.

7.3.4. GEV parameter relations

The relations between the fitted GEV parameters for each dataset are shown in Figs. 16 and 17. There is a strong positive correlation between the location and scale parameters, a and b, with low levels of scatter. The shape parameter, k, is uncorrelated to the location

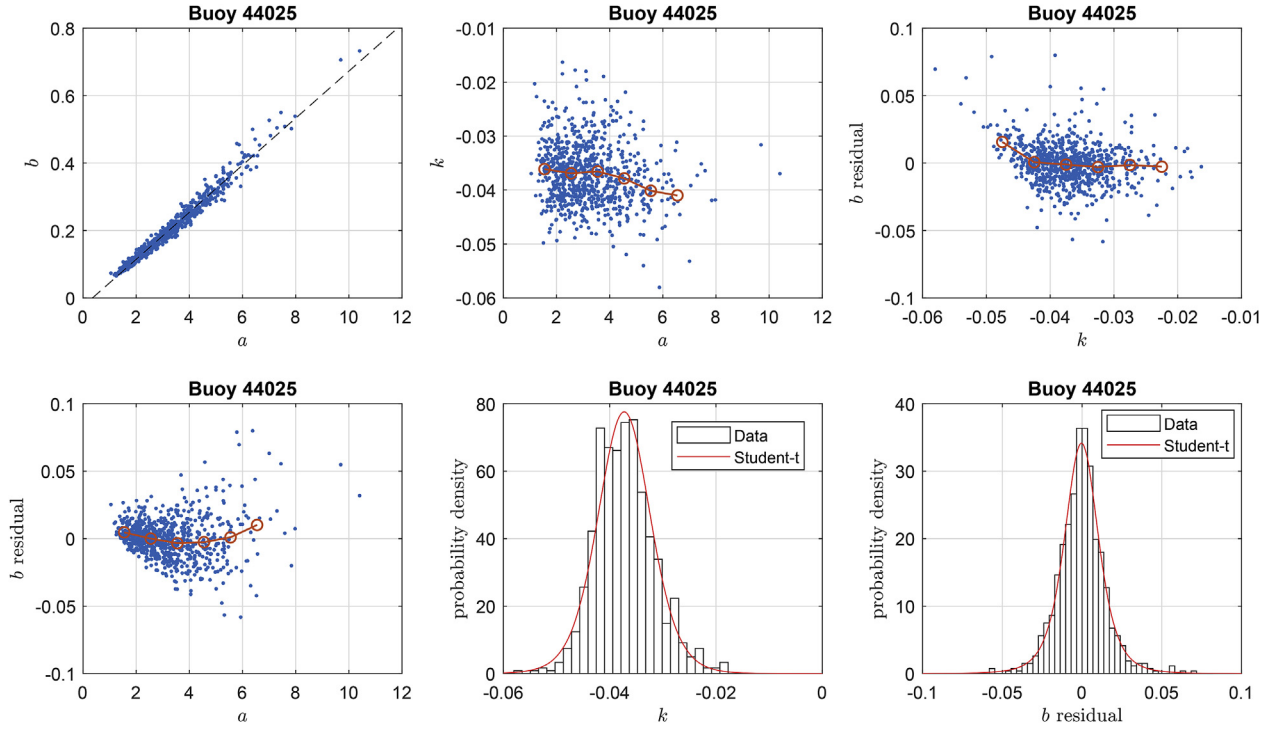

Fig. 16. Scatter plots of relations between GEV parameters and fitted distributions for buoy 44025. Blue dots indicate relations for individual storms. Red lines are mean values binned by the x-axis variable. Black dashed lines are linear fit. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

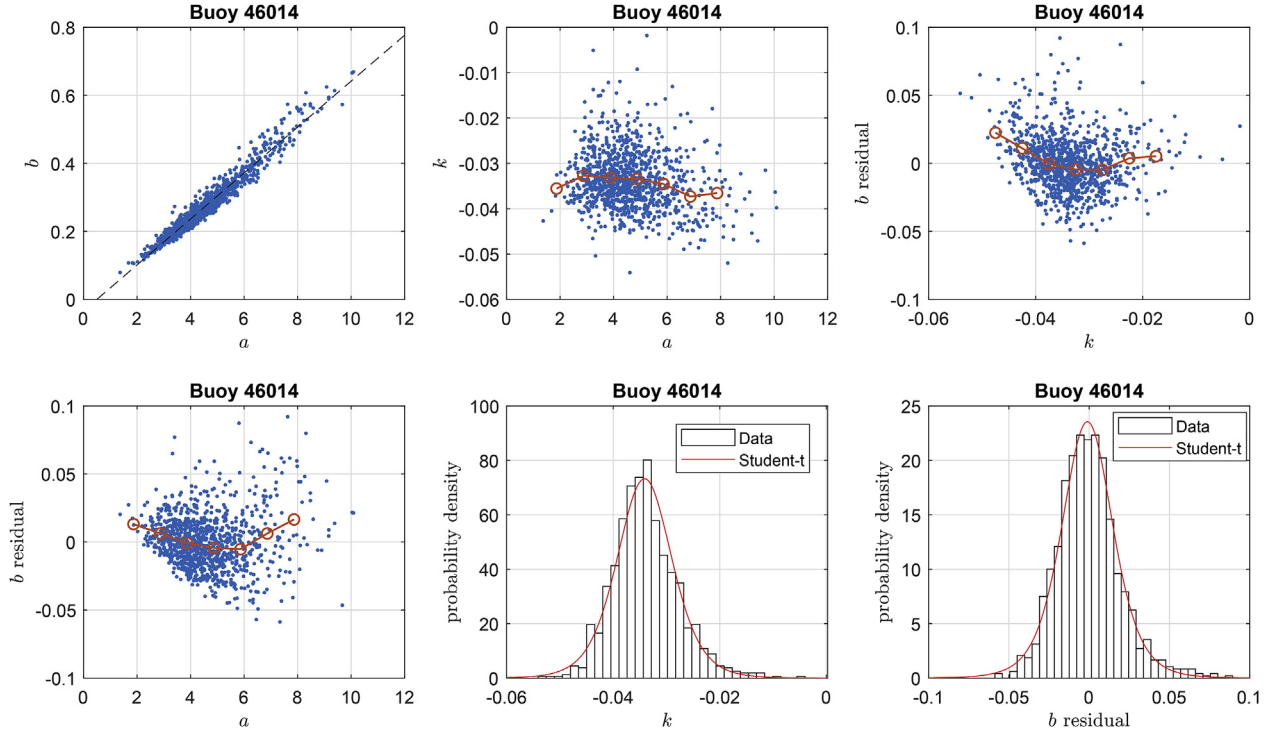

Fig. 17. Scatter plots of relations between GEV parameters and fitted distributions for buoy 46014.

parameter, a, and the residuals of the linear regression of b on a. There is also no strong trend in the regression residuals with a. The distributions of k and the regression residuals are both well fit by a Student-t distribution. As k is approximately independent of a and the residuals of b, the conditional density function $p(b, k|a)$ will be modelled as $p(b, k|a) = p(b|a)p(k)$, where $p(k)$ and $p(b|a)$ are both modelled as a Student-t distribution, with the mean value of $p(b|a)$ varying linearly with a.

It is interesting to note that the distribution of the GEV parameters is similar for both datasets. The shape parameter takes a narrow range of values, with a mean value around -0.035, close to the initial guess for the optimisation of $k = 0$. Given that k is consistently close to zero, it is

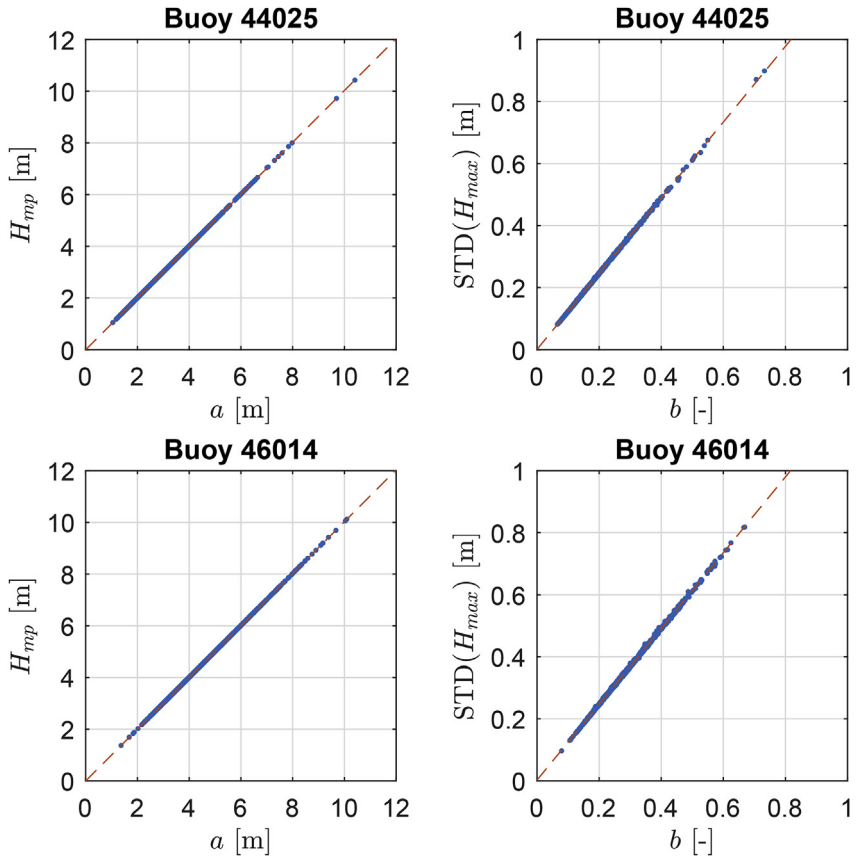

Fig. 18. Left: Scatter plot of H_{mp} against GEV parameter a. Right: Scatter plot of $STD(H_{max})$ against GEV parameter b. Dashed lines are linear regression.

interesting to examine how the parameters a and b relate to the mode and STD of the distribution. Fig. 18 shows scatter plots of H_{mp} against a and $STD(H_{max})$ against b for each dataset. The fitted value of a is almost identical to H_{mp}. From (16), for $k \neq 0$,

$$H_{mp} = a + b\frac{(1+k)^{-k} - 1}{k}. \tag{28}$$

In our datasets, $b \approx 0.07a$ and the mean value of the function of k on the RHS is ~ 0.034, so $H_{mp} \approx 1.0023a$, resulting in the near equality observed in Fig. 18. Similarly, from (17), for $k \neq 0$,

$$\text{std}(H_{max}) = \frac{b}{k}\sqrt{(g_2 - g_1^2)}, \tag{29}$$

where $g_n = \Gamma(1 - nk)$. For our datasets the mean value of the terms involving k on the RHS was ~ 1.22, which was equal to the gradient of the regression line shown in Fig. 18. These observations give an interpretation of the fitted GEV parameters in terms of properties of the distribution of H_{max} in the measured storm.

7.4. Comparison of return periods

Return value of maximum crest height were calculated using each equivalent storm model for both datasets, based on the fitted models for the joint distribution of equivalent storm parameters described in the preceding section. The return values from each model are shown in Fig. 19. The return values for Buoy 44025 are in good agreement between the models, with less than 0.2 m difference at the 100-year level and 0.5 m difference at the 1000-year level. For Buoy 46014, the ETS, TV and GEV models are in close agreement, but the EPS model gives higher return values.

There are three elements that could cause differences between the models, corresponding to the three terms in the integrand of (22): $p(a)$, $p(b, k|a)$ and $\Pr(H_{max} \leq h|ES)$. In Section 7.3.3 it was confirmed that the GPD model for $p(a)$ matches closely between the ETS/EPS and TV/GEV models. Moreover, the same model for $p(a)$ is used for both the ETS and EPS methods. In Section 7.2 it was shown that the EPS model is more accurate than the ETS and TV models at representing $\Pr(H_{max} \leq h|MS)$. Therefore the differences in the return values from the EPS method for Buoy 46014 must be a result of the model for the conditional density function $p(b, k|a)$. The relationships between the EPS parameters displayed a high level of scatter, with the inter-dependence between parameters being unclear and more difficult to model than for the other methods. It is therefore likely that the model assumed for the EPS parameter joint distribution does not correctly capture the correlation between the parameters, causing the observed differences in the return values. For the other models the dependencies between the parameters was clearer, so there is a higher confidence in the model for the joint distribution of storm parameters. The sensitivity to the assumed models for the parameter joint distributions is assessed in the next section.

7.5. Reduction to single integrals

The sensitivity to the assumed form of the model for the conditional density function $p(b, k|a)$ (or $p(b|a)$ in the case of ETS and TV models) can be assessed by comparing to results calculated using the mean values of b and k, conditional on a and replacing (22) by the single integral:

$$\Pr(H_{max} \leq h|RS) = \int_u^\infty p(a)\Pr(H_{max} \leq h|ES)\text{d}a. \tag{30}$$

Figs. 20 and 21 show a comparison return values from the ETS and EPS models calculated using single and double integrals. For Buoy 44025, both methods give the same return values using the single integral with the mean values of b and the linear model $k = A + Bb$ for

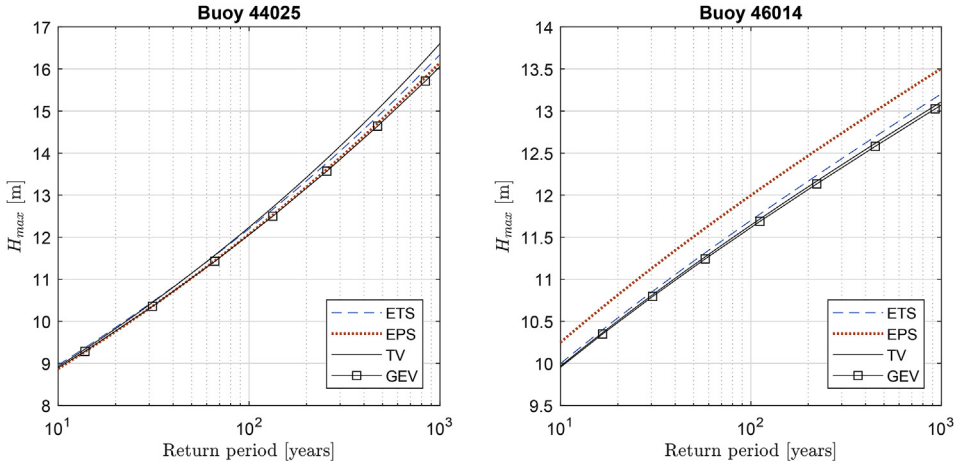

Fig. 19. Return values of maximum crest height calculated using each equivalent storm model.

the EPS method. However, for Buoy 46014, the single integral results in different estimates of return values, indicating a greater sensitivity to the models for $p(b|a)$. This sensitivity is likely related to the greater level of scatter in the parameter relations for Buoy 46014.

Fig. 22 shows that return values from TV model calculated using the double integral are the same as those calculated from a single integral assuming a linear model for $\ln(N)$ in terms of H_{mp}. However, if a constant mean value of $\ln(N)$ is used (as assumed by Tromans and Vanderschuren (1995) and recommended in the DNV GL guidelines, DNV GL, 2017), this results in lower return values for both datasets. Given that the data showed a linear dependence of $\ln(N)$ on H_{mp}, it is likely that the results using the constant mean value of $\ln(N)$ are less accurate.

For the GEV model, the single integral (30) is calculated using a constant mean value of k and a linear model for the dependence of b on a. Return values from the GEV model calculated using the triple and single integrals are shown in Fig. 23. In both cases the single integral gives identical results to the triple integral, indicating that the results are not sensitive to the assumed form of the conditional density function $p(b, k|a)$. This is due to the narrow range of values of k and low level of scatter in the relationship between a and b.

8. Summary

Table 1 presents a comparison of the equivalent storm models and interpretations of the model parameters. The ETS and EPS models both require a model to be established for how T_z and the parameters of the short-term distribution vary with H_s, whereas the TV and GEV models do not. The form of $\Pr(H_{max} \leq h|ES)$ is more complex for the ETS and EPS models than for the TV and GEV models, with the ETS/EPS models involving an integral of the short-term distribution over H_s. The TV and GEV models give simple closed form expressions for $\Pr(H_{max} \leq h|ES)$. The increased complexity in the ETS/EPS models means that the computational time required to fit them and calculate return periods is significantly higher than for the TV and GEV models.

The ETS and EPS models both have the height parameter, a, defined as the peak H_s of the measured storm. However, the duration, b, is not directly related to a well-defined characteristic of the measured storm. The TV and GEV models both have the parameters a and b related to well-defined characteristics of the measured storm. However, the TV model has its parameters defined in terms of the distribution of H_{max}^2, so the interpretation of the GEV parameters is more immediate.

The parameters of the TV and GEV models exhibited clear relations in comparison to the parameters of the ETS and EPS models. This resulted in the TV and GEV models being less sensitive to models for the conditional density function $p(b, k|a)$ than for the ETS and EPS models.

9. Conclusions

A new equivalent storm model has been introduced and compared to existing equivalent models. The new model generalises the method of Tromans and Vanderschuren (1995), by using the GEV to model the distribution of maximum wave height in a storm. It has been demonstrated that the GEV model gives a more accurate representation of the

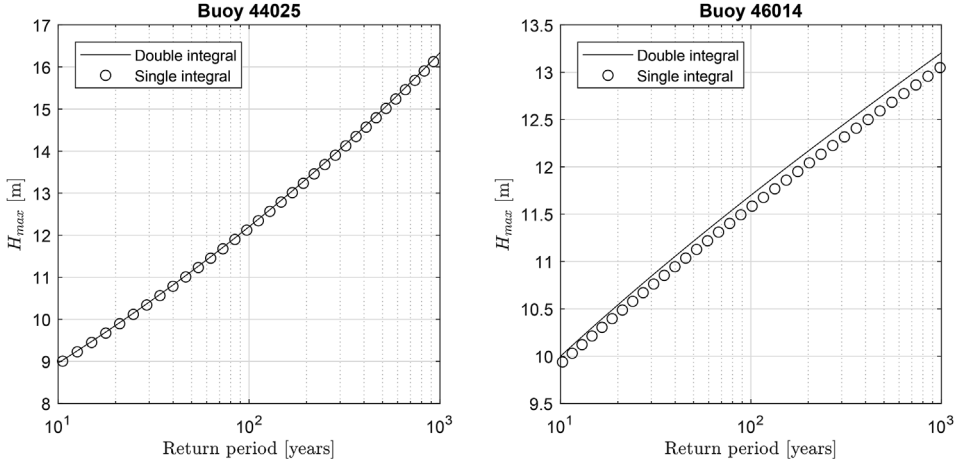

Fig. 20. Comparison of return periods from the ETS model using single and double integrals.

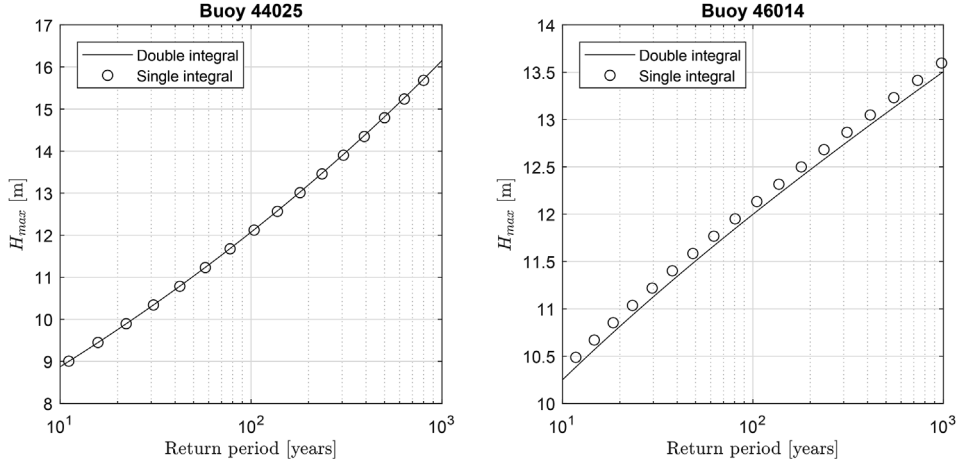

Fig. 21. Comparison of return periods from the EPS model using single and double integrals.

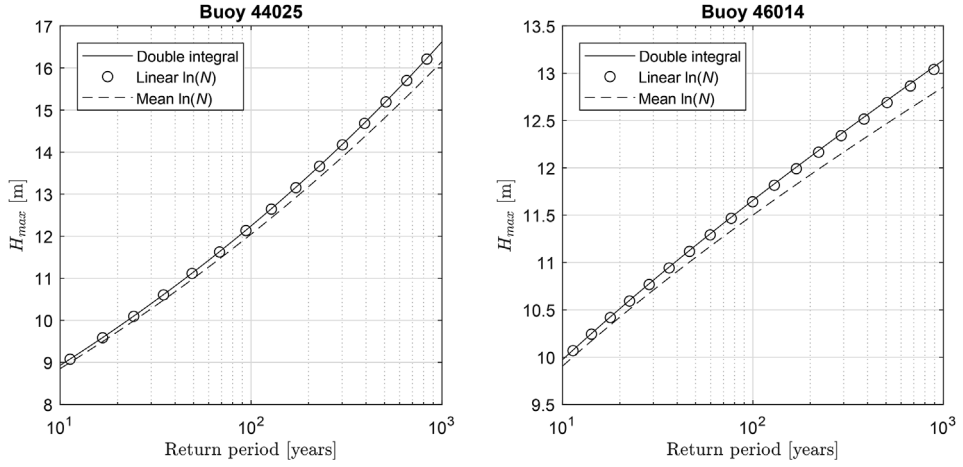

Fig. 22. Comparison of return periods from the ETS model using double integral and single integral using either a linear model or a constant mean value for $\ln(N)$.

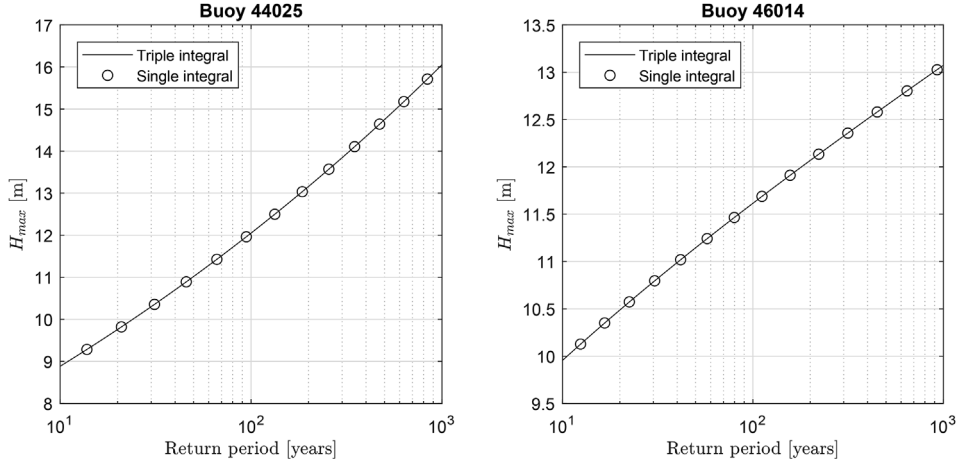

Fig. 23. Comparison of return periods from the GEV model using single and triple integrals.

distribution of maximum wave or crest height in a storm than the TV, ETS and EPS models. The TV model fitted using moment method was shown to have the highest bias of all the models considered. The numerically-fitted TV model had a reduced bias compared to the TV model fitted from the moments of the distribution, but still performed worse than the ETS and EPS models. The GEV model is only marginally more complex to implement than the numerically-fitted TV model, but results in much lower biases and almost two orders of magnitude reduction in the Cramer-von Mises goodness-of-fit parameter.

The ETS and EPS methods provide an indirect way of parameterising the distribution of the maximum wave height in a storm. The GEV method removes the intermediate and unnecessary step of modelling the temporal evolution of sea states in a storm and instead provides a direct representation of the statistical properties of the measured storms. The GEV method also removes the need to assume a model for how other sea state parameters vary with H_s, which is of

Table 1
Comparison of equivalent storm models and interpretation of parameters.

Model	$\mathbf{Pr}(\boldsymbol{H_{max} \leq h \mid ES})$	$\boldsymbol{a}$	$\boldsymbol{b}$	$\boldsymbol{k}$
ETS	$\exp\left(\frac{b}{a}\int_0^a \frac{\ln[\Pr(H \leq h\|\sigma(H_S))]}{T_Z(H_S)} dH_s\right)$	Storm peak H_s	Storm duration	N/A
EPS	$\exp\left(\frac{b}{ka}\int_0^a \frac{\ln[\Pr(H \leq h\|\sigma(H_S))]}{T_Z(H_S)}\left(1-\frac{H_S}{a}\right)^{\frac{1}{k}-1} dH_s\right)$	Storm peak H_s	Storm duration	Storm shape
TV	$\exp\left(-\exp\left(-\left(\frac{h^2-a}{b}\right)\right)\right)$	H_{mp}^2	$\frac{\sqrt{6}}{\pi}\mathrm{std}(H_{max}^2)$	N/A
GEV	$\exp\left(-\left(1+k\left(\frac{h-a}{b}\right)\right)^{-\frac{1}{k}}\right)$, for $k \neq 0$ $\exp\left(-\exp\left(-\left(\frac{h-a}{b}\right)\right)\right)$, for $k = 0$	H_{mp}	$\sim 1.22\ \mathrm{std}(H_{max})$	Distribution shape

benefit if the short-term distribution is dependent on multiple sea state parameters, such as steepness, spectral bandwidth or directional spread.

The GEV model was found to be the most robust in terms of the models for the joint distribution of storm parameters. The relations between the GEV parameters was clear in both datasets considered and it was shown that return values can be calculated from a single integral (30) using the mean value of the shape parameter and a linear model for the relation between the scale and location parameters.

In summary, the GEV method is simpler, more accurate and more robust than the ETS and EPS methods and is therefore recommended for calculating return periods of individual wave and crest heights.

Declarations of interest

None.

Acknowledgement

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

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.coastaleng.2018.06.001.

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.
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.
Boccotti, P., 1986. On coastal and offshore structure risk analysis. Excerpta Ital. Contrib. Field Hydraul. Eng. 1, 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.
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.
DNV, G.L., 2017. Environmental Conditions and Environmental Loads. DNV GL Recommended Practice RP C205, August 2017. .
Dupuis, D.J., 1996. Estimating the probability of obtaining nonfeasible parameter estimates of the generalized Pareto distribution. J. Stat. Comput. Simulat. 54, 197–209.
Dupuis, D.J., Tsao, M., 1998. A hybrid estimator for generalized Pareto and extreme-value distributions. Commun. Stat. Theor. Meth. 27, 925–941.
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., 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.
Forristall, G.Z., 2000. Wave crest distributions: observations and second order theory. J. Phys. Oceanogr. 30 (8), 1931–1943.
Hosking, J.R.M., Wallis, J.R., 1987. Parameter and quantile estimation for the generalized Pareto distribution. Technometrics 29, 339–349.
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 Statist. Soc. 46 (4), 487–501.
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.
Krogstad, H.E., 1985. Height and period distributions of extreme waves. Appl. Ocean Res. 7, 158–165.
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., Malara, G., Romolo, A., Arena, F., 2016. Peak over threshold vis-à-vis equivalent triangular storm: return value sensitivity to storm threshold. Coast. Eng. 116, 220–235.
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.
Latheef, M., Swan, C., 2013. A laboratory study of wave crest statistics and the role of directional spreading. Proc. R. Soc. A 469, 20120696.
Mackay, E.B.L., 2012. Resource assessment for wave energy. In: In: Sayigh, A. (Ed.), Comprehensive Renewable Energy, vol. 8. Elsevier, Oxford, pp. 11–77.
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.
Mathiesen, M., Goda, Y., Hawkes, P., et al., 1994. Recommended practice for extreme wave analysis. J. Hydraul. Res. 32 (6), 803–814.
Martín-Hidalgo, M., Martín-Soldevilla, M.J., Negro, V., Aberturas, P., López-Gutiérrez, J.S., 2014. Storm evolution characterization for analysing stone armour damage progression. Coast. Eng. 85 (2014), 1–11.
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.
Tayfun, M.A., Fedele, F., 2007. Wave-height distributions and nonlinear effects. Ocean Eng. 34, 1631–1649.
Toffoli, A., 2016. Rogue waves in random sea states: an experimental perspective. In: In: Onorato, M., Resitori, S., Baronio, F. (Eds.), Rogue and Shock Waves in Nonlinear Dispersive Media. Lecture Notes in Physics, vol. 926 Springer, Cham.
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 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.
Wu, Y., Randell, D., Christou, M., Ewans, K., Jonathan, P., 2016. On the distribution of wave height in shallow water. Coast. Eng. 111, 39–49.
Zhang, J., 2007. Likelihood moment estimation for the generalised Pareto distribution. Aust. N. Z. J. Stat. 49 (1), 69–77.
Zhang, J., 2010. Improving on estimation for the generalized pareto distribution. Technometrics 52, 335–339.