The assessment of extreme wave analysis methods applied to potential marine energy sites using numerical model data

Citation for published version:
Agarwal, A, Venugopal, V & Harrison, GP 2013, 'The assessment of extreme wave analysis methods applied to potential marine energy sites using numerical model data' Renewable and Sustainable Energy Reviews, vol. 27, pp. 244-257. DOI: 10.1016/j.rser.2013.06.049

Digital Object Identifier (DOI):
10.1016/j.rser.2013.06.049

Published In:
Renewable and Sustainable Energy Reviews

The assessment of extreme wave analysis methods applied to potential marine energy sites using numerical model data

Atul Agarwal , Vengatesan Venugopal , Gareth P. Harrison

Institute for Energy Systems, School of Engineering, University of Edinburgh, Kings Buildings, Mayfield Road, Edinburgh

E-mail addresses: a.agarwal@ed.ac.uk, v.venugopal@ed.ac.uk, gareth.harrison@ed.ac.uk

The accurate estimation of extreme conditions, such as 100-yr return levels of significant wave height is an important aspect in the design of marine energy converters, offshore and coastal structures. This study investigates the different approaches for the estimation of extreme waves that have been applied in the past, and determines the 100-yr return levels using the high resolution ERA-Interim dataset produced by the European Centre for Medium-Range Weather Forecasts (ECMWF). It is demonstrated in the paper that fitting a Generalized Pareto Distribution to all exceedances over a high threshold is the most suitable approach. The estimates thus obtained are compared with previously computed estimates for buoys and offshore platforms. The effect of duration of data on the estimates is also investigated. Finally, a 100-yr return level map for the North Atlantic region is presented.

Keywords: 100-Year wave height, Marine energy, Peak over threshold, ERA-Interim data, North Atlantic, North Sea

Nomenclature

$\hat{G}$	Empirical probability of exceedance (Generalized Extreme Value)
$\tilde{G}$	Model probability of exceedance (Generalized Extreme Value)
$\hat{H}$	Empirical probability of exceedance (Generalized Pareto)
$\tilde{H}$	Model probability of exceedance (Generalized Pareto)
ξ	Shape parameter
μ	Location parameter
σ	Scale parameter
u	Threshold
λ	Average exceedances per year
k	Total number of exceedances
n	No. of years in dataset
m	Return period

1. Introduction

The need to increase energy security and mitigate climate change has caused significant interest in low carbon sustainable sources of energy. Onshore wind and solar energy converters, being the more mature technologies available, have enjoyed greater investment. More recently, increasing attention is being paid to marine sources of energy such as waves and tidal currents, as evidenced by financial support from governments and research organizations for research and development. This impetus is particularly visible in countries with large potential for exploiting these resources, like the UK and other European countries. Many wave energy converters, such as the Aquamarine Oyster, Pelamis, Wave Dragon and Wave star, are in various stages of sea trials. The leasing of wave and tidal energy sites in UK waters by the Crown Estate (http://www.thecrownestate.co.uk/energy/wave-and-tidal/pentland-firth-and-orkney-waters/leasing-round-and-projects) is an indicator of the impending large scale deployment of such devices in arrays.

For the conversion of energy, these devices rely on the marine environment which can be very harsh at times. It is for this reason that before investing in large scale wave and tidal farms, knowledge of the marine environment, particularly the extreme conditions, is required. This is especially important if the prevailing climate is expected to change in coming decades.

The analysis of extreme waves has been an area of scientific interest for many decades because of its importance in the design of marine structures and vessels. Studies in the past have proposed several methods to estimate return values, like

- The 'Initial Distribution method' (e.g. [1,2]).
- Extreme value methods (e.g. [3,4]).
- and Threshold methods.

Data from a variety of sources has been used, such as buoys and ship borne wave recorders (e.g. [3]), models and reanalyses (e.g. [5–7]) and satellite missions (e.g. [8–9]). Popular methods used by the above researchers are described in Section 3. This study proposes to identify an appropriate method from these for the North Atlantic region and UK waters.

Extreme wave heights, also referred to as m-year return values, are wave heights that are likely to be exceeded, on average, once in every m years. In the context of survivability and economic viability of marine energy devices, the accurate estimation of extreme wave conditions for their design is paramount. An underestimation of these extremes could adversely affect the survivability the device leading to catastrophic failure, while a high, but safe estimate would inevitably lead to over design, resulting in needlessly high capital cost, making the return on investment financially unattractive.

Failure of devices operating in marine environments can occur suddenly, during extreme events, or over a period of time, for example by the action of corrosion, fatigue, wear, etc. During storms and other extreme events, the stresses induced on the foundations, moorings, pylons, sub-structures, etc. can exceed the design stress causing failure of the device. To minimize the chances of such a failure to acceptable levels, these extreme conditions need to be estimated with a high degree of confidence.

Although wave energy converters (WECs) are designed for a service life of between 20 and 30 years, return values of longer periods need to be considered when estimating extreme wave conditions. In the selection of an extreme wave height when designing for survivability, 100-yr return values are often used because of the low probability of occurrence associated with them.

As the methods and distributions popular in the estimation of extreme wave heights are essentially statistical extrapolations, if the data does not span over a period that is significant in comparison with the return period, it becomes extremely difficult to say with certainty that one particular method or distribution is more accurate or closer to the real 100-yr extreme than others. Moreover, these methods are based on assumptions that the data is statistically independent of one another and are identically distributed (i.i.d.). The observation of the Poisson property in a time series of significant wave height (H_s) makes the use of these methods awkward because of the non-independence of data. Given the conditions of non-independence and non-stationarity, this study investigates different approaches and models and identifies the most suitable approach for estimating long range return values.

An accurate method for estimating extreme wave conditions has applications in the design of the structures, foundations, pylons and mooring systems of not only WECs, but also offshore wind turbines, tidal turbines, and other marine installations. The results of this study will aid in the design of floating components of marine energy systems, as well as submerged substructures.

For structures operating in the offshore environment, design conditions include extreme wave heights along with an associated peak period. The focus of this study is the estimation of extreme wave heights only, and a similar study for the associated wave period may be undertaken in future. In the meantime, to estimate the associated wave period, either joint probability models (bivariate distributions of extreme wave height and period), as suggested by Goda [10] and Wolfram et al. [11], or an empirical relationship with wave height, as recommended by DNV, 2010 [12] may be used.

The objective of this study is to identify a robust method for accurately estimating extreme wave heights. The identified method will then be demonstrated by preparing a 100-yr wave map for the North Atlantic and North Sea, defined as the region bounded by the latitudes 10°N and 80°N, and the longitudes 70°W and 20°E. In principle these methods are generic and the calculations can be applied to any region. This study is unique in that a high reanalysis dataset (0.75° resolution) is used. Previous studies reported return values based on lower resolutions (e.g. ERA-40 data at 1.5° resolution by [5]) which may be deemed inadequate considering the scale of marine energy sites.

The authors feel that despite many previous studies in the estimation of extrema, this study is warranted because of its uniqueness in terms of the dataset used. The relative high resolution, along with high quality data on decadal scales is likely to produce estimates with a higher degree of confidence associated with them. The study applies the findings to prepare estimates on key marine energy sites around the UK, for which no prior work has been undertaken. Moreover, this study compares a previously recommended method (i.e. fitting a Weibull or Gumbel distribution to a population of storm maxima) with the approach of fitting a Generalized Pareto distribution to a population of all excesses above a threshold. It also assesses the effect of the duration of data used. These, to the best of the authors' knowledge, have not been previously studied.

Section 2 describes the datasets used to perform the various analyses, and for the comparison of results, Section 3 reviews methodologies available for estimating extreme conditions, Section 4 describes the methodology followed by the study and the analyses of data conducted, and Section 5 explains the preparation of the 100-yr extreme wave map for the region. Discussion of the results and the conclusions are in Sections 6 and 7.

2. Data description

In this study, ERA-Interim data produced by the European Center for Medium-Range Weather Forecasts (ECMWF) was used for the calculation of extremes. The obtained estimates were compared against those obtained from data from various buoys and offshore platforms. Although satellite data from several missions was available at the time of the study, and has been used in the past for extreme wave estimation, e.g. Wimmer et al. [9], these were not used because an assumption about the type of distribution is required, e.g. Fisher–Tippet (FT-1) as used by Carter, [8] and any inferences made are based on the assumption being correct.

2.1. *ERA-Interim*

ERA-Interim is the most recent reanalysis produced by ECMWF. It contains several gridded datasets describing ocean-wave conditions in addition to land-surface conditions. Of interest is the hindcast of significant wave height, H_s, sampled at 6-hourly intervals. The hindcast spans over three decades, extending from January 1, 1979 to February 29, 2012, and is continuously extended forward in near-real time. It can be publicly accessed online at full spatial resolution (http://data-portal.ecmwf.int/data/d/interim_full_daily) and for the present study model wave data at a resolution of 0.75° was used.

In comparison with previous reanalyses, the ERA-Interim dataset is considered superior not only on account of its high resolution but also because of its superior data assimilation system. In the preparation of the ocean-wave analysis, reprocessed altimeter wave-height data from satellite missions ERS- and ERS-2, as well as near-real-time data from ENVISAT, JASON-1 and JASON-2 were used, which were not used previously. A comparison of ERA-40 and ERA-Interim data for the overlapping period January 1989 to May 2010 shows that the ERA-Interim data assimilation system is able to capture future observations better, resulting in improved temporal consistency [13]. To the best of the authors' knowledge, no work has been carried out to determine the accuracy and of ERA-Interim data and suggest a calibration function based on other data sources.

For the estimation of extreme waves and the preparation of the 100-yr return level contour map for the North Atlantic, H_s hindcast data from the ERA-Interim reanalysis was used, as is, for the period spanning from January 1, 1979, to December 31, 2011.

2.2. *Estimates based on buoy/platform measurements*

Buoy observations are considered to be the most reliable observations of wave height. However, these are limited to only a few locations along the North American and west European coasts, and few buoys exist in the region with observations extending on a decadal scale. The available data requires significant quality control on account of large gaps and out-of-range measurements.

Long term data from buoys in UK waters are unfortunately not available in the public domain. In the absence of this data, pre-calculated estimates from other studies, which use data from buoys and offshore platforms, will be used to assess the ERA-Interim return value estimates. These estimates from literature [14] would be compared with the return values of the nearest intersection of the $0.75^\circ \times 0.75^\circ$ data-grid obtained from this study.

3. Review of methods of extreme value estimation

3.1. *Initial-distribution method*

In the traditional method of estimating return values all the available H_s data is gathered in a single sample and a suitable parametric model is fitted through the data. As the extreme conditions fall outside the observed range, the curve is then extrapolated to the desired low probability of occurrence and the corresponding H_s value is taken as the extreme value [15,16]. The selection of a suitable distribution is empirical, and there is little scientific justification to use one distribution over another [5,16]. To overcome this, several possible distributions are fitted to the data and the one that fits best is extrapolated to obtain the return value. The best fitting curve can be identified by visual inspection, or by goodness-of-fit tests, e.g. χ^2- test, the Kolmogorov–Smirnov test and the Anderson–Darling test. It can be observed from previous work in extreme wave estimation that the Weibull and log–normal distributions are popular models when this approach is used, e.g. [1,17].

There are three main problems associated with this method. One, as suggested by Ferreira and Soares [15], is that with measurements sampled at 3-hourly or 6-hourly intervals, it is difficult to identify the data to a single statistical population, because measurements from the same reference period can be significantly different from each other. Also, as data are not independent and non-stationary, common statistical methods based on i.i.d. conditions cannot be used and consequently invalidates the definition of return value [5,15]. Finally, goodness-of-fit tests may not be reliable in the selection of a distribution because, given the size of data analyzed, they may not be able to distinguish the tail type. Extrapolating outside the sample range by ignoring the tail type may be incorrect.

3.2. *Block maxima method*

In this method the time period over which data is collected is divided into blocks, and the maximum wave height in each block is used in the analysis. The division of the period can be done on the basis of periods of fixed length, e.g. daily or monthly, or on the basis of storms [18]. When the size of blocks is one year, the block maxima method is also called 'Annual Maximum method'. Sometimes, a high threshold is used to identify storms, and if a process of declustering (a process of identifying and separating individual storms) is applied to identify individual storm maxima for the estimation of extremes, the method simplifies to a variation of the block maxima method, where the size of each block may vary, and the blocks will most likely not be adjacent in the time series.

Prior to the grouping of the Gumbel, Fréchet andWeibull distributions into a single family of models, known as the Generalized Extreme Value (GEV) model, the block maxima method involved fitting a one of the three distributions to the data, and extrapolating to the desired probability levels to obtain a return value. Further details on the above distributions may be found in [10].

The drawback of this method is that one of the three must be selected based on some criteria, and once a distribution is selected, subsequent inferences presume the selection to be correct. The asymptotic generalized extreme value (GEV) theory provides a method for the analysis of block maxima by fitting a GEV model to the data to remedy this [19]. Another disadvantage of the block maxima method is that if monthly or seasonal maxima are used, seasonal variation in these maxima can result in a poor fit. To overcome this, either some kind of scaling process needs to be applied to remove the seasonal variation, or the annual maxima should be used. The Annual maxima method, on the other hand, has the drawback that the estimation of 50-yr and 100-yr return levels requires the data to span for several decades to provide enough points for curve fitting.

3.3. Peaks-over-threshold method

In the peaks-over-threshold (POT) method, all the values of H_s that exceed a threshold value are considered in the calculation of a return value. If the threshold (u) is sufficiently high, the exceedances can be modeled using the Generalized Pareto Distribution (GPD) [19]. Other models, such as the Gumbel (FT-1) and Weibull distributions have also been used to describe the exceedances, as recommended by Mathiesen et al. [20] and demonstrated by Haver and Nyhus [21], Carter [8], and Neelamani et al. [7] to name a few. In modeling exceedances and maxima, the Gumbel distribution is considered an important distribution because it is the domain of attraction of the Weibull and log–normal distributions. In other words, if the parent distribution is either a Weibull or a log–normal distribution (both of which are popular models used in the initial distribution method), the GEV distribution of its maxima reduces to a Gumbel distribution [19].

Fitting a GPD or similar distribution to exceedances is a valid approach when the data is independent and identically distributed. However, the assumptions of i.i.d. may not be accurate, as any hourly wave height may bear some relationship with the wave heights of previous hours on account of the Poisson property. Under these conditions of non-independence, modeling strategies include: (a) identifying clusters, such as storms, and modeling cluster maxima only, and (b) ignoring the dependence, but inflating the standard errors to take into account the limited independence of data. The former approach involves a process of declustering to obtain the maxima of individual storms. As storms are separated by some, non-constant period of time, the set of storm maxima can be treated as being statistically independent.

The second strategy is simpler, and can be justified on the basis that the marginal model is valid [19].

The block maxima and POT methods have been applied to representative data and its suitability is explored in detail before the preparation of the extreme wave map of the region.

4. Preliminary analyses

In order to investigate the suitability of methods discussed previously and compare parametric models for estimating extreme wave heights, five different locations around the UK and Republic of Ireland were identified (see Table 1). Analyses were conducted for these locations for identifying a suitable method to apply across the North Atlantic and North Sea.

The selection was done on the basis of importance with regard to marine energy while trying to keep them as far away from each other as possible in order to make general inferences. The five sites for the investigation of methods and models include test sites for wave energy converters (WECs) at the European Marine Energy Center (EMEC) at Orkney Islands and Wavehub off the coast of Cornwall, a proposed wave energy test site in Irish waters west of Belmullet, a prospective wave energy site near the Outer Hebrides expected to be leased in the Further Scottish leasing Round, and a location in the North Sea near a proposed offshore wind energy site in the Crown Estate Round 3 auction zone.

Where H_s data is not available for the exact location because of the $0.75^\circ \times 0.75^\circ$ resolution of the dataset, data from the nearest grid intersection is used for analyses.

4.1. Selection of method and model

For the estimation of return values across the North Atlantic and North Sea region, a suitable approach needed to be selected between the block maximum and POT methods. The initial distribution method was eliminated from the pool because of the high uncertainties and the challenges associated with fitting multiple distributions and testing goodness of fit. As the data spans over only 33 years (1979–2011), the annual maxima method appears unattractive because of the low confidence levels associated with fitting a curve to as few as 33 data points and extrapolating it. Using monthly maxima would provide a larger dataset to fit a curve to, but scaling would be required to remove seasonal variations.

The simpler options available were to either consider only storms (defined as periods during which H_s exceeds a high threshold) and fit a GEV distribution to the storm maxima, or to use the POT method and fit a GPD to all exceedances above a high threshold.

The selection of a suitable threshold is key in both approaches. There are several difficulties associated with threshold selection. A single ‘high’ value (e.g. 5 m for the North Sea), cannot be used across the region because the lower latitudes experience less energetic storms than the higher latitudes. Thus, a floating threshold is required which varies from cell to cell, depending on the H_s data, such that it is sufficiently high in the high latitude cells, and low in the low latitude cells. The threshold needs to be sufficiently large to permit a good fit of an extreme value distribution, while at the same time being low enough to obtain a sufficiently sized sample.

Table 1
Marine energy locations for preliminary analyses along with the obtained model parameters.

Site name	Nearest grid point	GEV			GPD		
		ξ	σ	μ	ξ	σ	μ^*
EMEC	59.25°N, 3°W	0.295	0.649	5.365	−0.079	0.966	4.675
Wavehub	50.25°N, 6°W	0.334	0.689	5.323	−0.068	1.002	4.587
Outer Hebrides	58.5°N, 6.75°W	0.333	0.659	5.677	−0.091	1.025	4.987
Belmullet	54°N, 10.5°W	0.285	0.886	6.711	−0.052	1.222	5.782
North Sea	58.5°N, 2.25°W	0.340	0.530	4.615	−0.092	0.866	4.065

Several methods have been proposed in the past for selecting a suitable threshold. Of these, visual inspection methods are ruled out because of the impracticality of inspecting threshold choice plots for each cell in the region. Dupuis [22] presents a method based on a robustness estimator, while Tancredi et al. [23] use a Bayesian approach. Both of these are computationally demanding, and more so when applied cell-by-cell over as large a region as the North Atlantic. Thompson et al. [18] present a quicker and computationally less demanding method of selecting an automated threshold based on the distribution of differences of parameter estimates, and Wimmer et al. [9] select a floating threshold determined by a predefined minimum sample size.

As the automated threshold selection method will need to be applied cell-by-cell across the region, it needs to be as simple as possible, while being effective. A threshold for each cell was selected as the 95% quantile of the H_s data for that cell. This is a higher quantile than the 93% quantile used by Caires and Sterl, 2005 [5] which yielded a GPD that fit most of their data well. A higher threshold, i.e. 95% quantile was selected as it was likelier to yield samples that could be well described by a GPD. For each of the test locations, the threshold thus selected was found to satisfy the two main criteria—the sample of exceedances was of adequate size ($n > 400$), and the extreme value models fit the data well, as will be demonstrated later.

With the selected threshold, two data sets were prepared—the first containing only the maxima of storms identified by a process of declustering; and the second containing all exceedances over the threshold (without declustering). In order to achieve a reasonable degree of independence in the first dataset, declustering of storms was achieved by only considering peaks that exceeded the threshold and were separated by three days or more. A GEV model was fit to the first sample, and a GPD was fit to the second. The parameters were estimated using the Maximum Likelihood method and are presented in Table 1 for the five representative locations. (For more details on the Generalized extreme value and Generalized Pareto models and their applications in extreme value analysis refer [19])

The GEV model has a distribution function of the form Eq. (1) [19]

$$G(z) = \exp\left\{-\left[1+\xi\left(\frac{Z-\mu}{\sigma}\right)\right]^{-1/\xi}\right\} \tag{1}$$

where z is the ordered block maxima, such that $z(1), z(2)\ldots, z(k)$; μ is the location parameter, σ is the scale parameter, and ξ is the shape parameter.

The distribution function, Eq. (1) can be re-written as [19]

$$\hat{G}(z_i) = \begin{cases} \exp\left[-\exp\left\{-\left(\frac{z_i-\mu}{\sigma}\right)\right\}\right], & \xi = 0 \\ \exp\left[-\left[1+\xi\left(\frac{z_i-\mu}{\sigma}\right)\right]^{-1/\xi}\right], & \xi \neq 0 \end{cases} \tag{2}$$

The empirical distribution at z_i can be evaluated by Eq. (3) [19]

$$\tilde{G}(z_i) = \frac{i}{k+1} \tag{3}$$

If the GEV model is working well, $\hat{G}(z) \approx \tilde{G}(z)$. A probability plot consisting of the points $\{(\tilde{G}(z_i), \hat{G}z_i),\ i = 1,\ 2, \ldots k\}$ should be linear, lying close to the unit diagonal. Substantial deviation from linearity would indicate the failure of the GEV model in describing the data. In addition to the probability plot, a quantile plot (QQ-plot), consisting of the points $\{(\hat{G}^{-1}(i/k+1),\ z_i),$ $i = 1,\ 2,\ \ldots k\}$ can be used to check the model, where [19]

$$\hat{G}^{-1}\left(\frac{i}{k+1}\right) = \begin{cases} \mu - \sigma\ln\left[-\ln\left(\frac{i}{k+1}\right)\right], & \xi = 0 \\ \mu - \frac{\sigma}{\xi}\left[1-\left\{-\ln\left(\frac{i}{k+1}\right)\right\}^{-\xi}\right], & \xi \neq 0 \end{cases} \tag{4}$$

Departure from linearity in the QQ-plot would also indicate the inability of the GEV model to describe the data.

For a high threshold u, the ordered set of threshold excesses can be described by a GPD, where, $y = H_s - u$, and $y(1) \leq y(2)\ldots \leq y(k)$. The GPD function takes the form [19]

$$\hat{H}(y) = 1 - \left(1+\frac{\xi y}{\sigma}\right)^{-1/\xi} \tag{5}$$

provided that $\xi \neq 0$.

Similar to the GEV distribution, the empirical distribution of the GPD a y_i can be found by

$$\tilde{H}(y_i) = \frac{i}{k+1} \tag{6}$$

A probability plot consisting of the pairs $\{(\tilde{H}(z_i), \hat{H}z_i),\ i = 1,\ 2, \ldots k\}$ can be plotted and inspected. If the GPD describes the data well, $\hat{H}(z) \approx \tilde{H}(z)$ and the plot should be approximately linear, lying close to the unit diagonal. The QQ-plot can also be plotted consisting of the points $\{(\hat{H}^{-1}(i/k+1),\ z_i),\ i = 1,\ 2,\ \ldots k\}$, where,

$$\hat{H}^{-1}\left(\frac{i}{i+1}\right) = u + \frac{\sigma}{\xi}\left[\left(\frac{i}{k+1}\right)^{-\xi} - 1\right] \tag{7}$$

The QQ-plot thus obtained should also be linear.

The procedure described here was applied to H_s data for the five test locations, and diagnostic plots for the GEV and GPD models were prepared by applying Eqs. (3) and (4) to the set of ordered storm maxima, and (6) and (7) to the set of ordered threshold excesses.

Diagnostic plots for the GEV and GPD models for the five sites are shown in Figs. 1 and 2. From the visual inspection and comparison of these diagnostic plots, it can be seen that fitting a generalized Pareto distribution to all excesses, rather than fitting a GEV model to storm maxima, describes the data better. Although the probability plot for the GEV model for each of the five locations is almost linear, it is evident that the probability plot for the GPD exhibits a closer match between the model and empirical probabilities. Similarly, a comparison of the QQ-plots also reveals that the GPD is a better model, in comparison with a GEV. The QQ-plots for the GEV model show a reasonable correlation between empirical and model data for low values of H_s, but deviate substantially from the line-of-best-fit for high values; whereas the QQ-plots for the GPD model show a greater correlation for most of the data with the exception of a few very high wave heights. As our interest lies in extrapolating the curve to obtain an extremal point which would lie in the high-value region, this would imply that the risk associated with such a projection would be lower if the GPD was used, on account of the better ability to describe the data.

The empirical and model data obtained above was further subjected to the Kolmogorov Smirnov (KS) and χ^2 tests to assess the goodness of fit of the GPD and GEV distributions. The obtained p-values are tabulated in Tables 2 and 3. Examination of these also shows that on the collective basis of the two tests, the GEV model can be rejected at a 10% significance level in most cases.

In addition, the GEV model was applied to the population of all points above the threshold, and goodness-of-fit tests were conducted. Table 3 shows that the influence of the method of selection of the statistical population is not significant and does not improve the fit of the GEV model. The p-values indicate that the GEV model can be rejected for the POT sample at the 1% significance level.

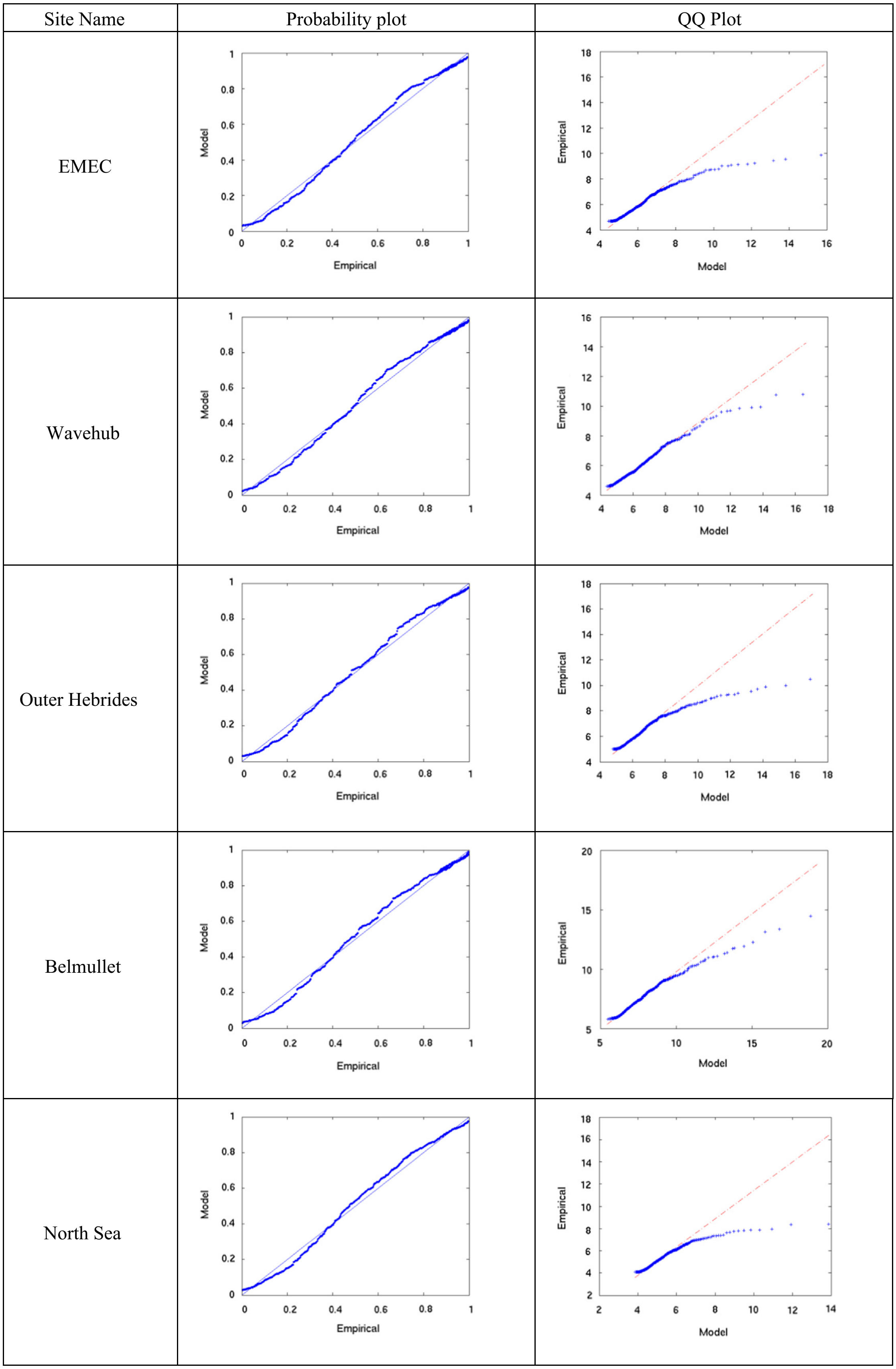


Fig. 1. Diagnostic plots for the GEV model.

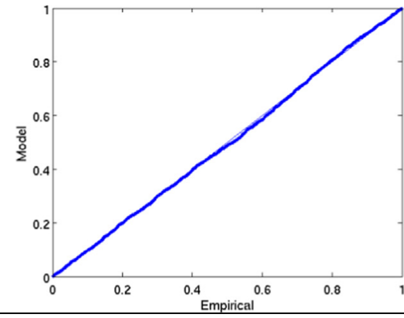
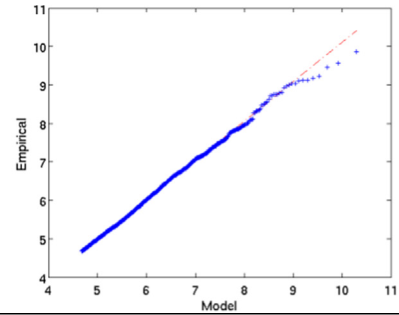
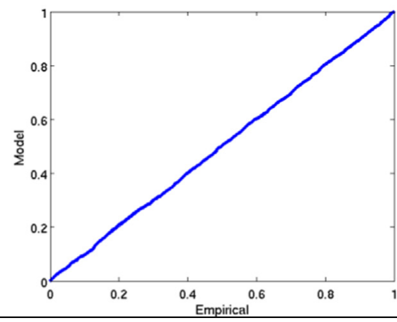
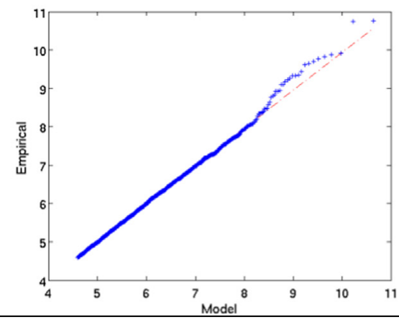
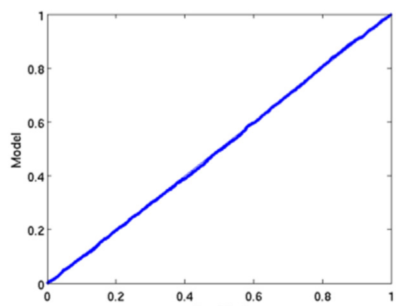
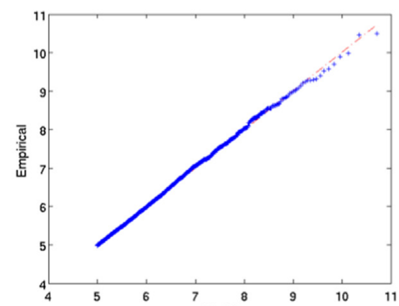
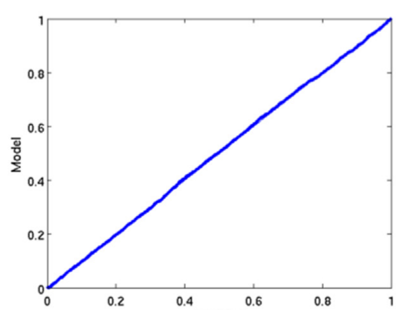
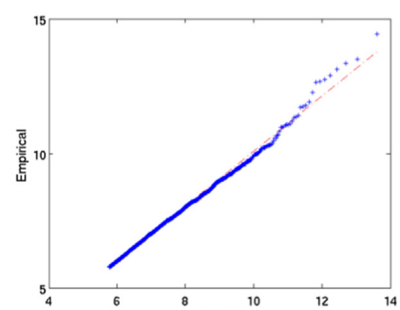
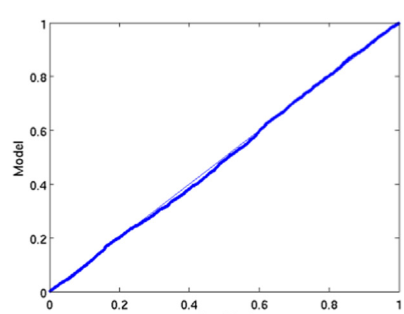
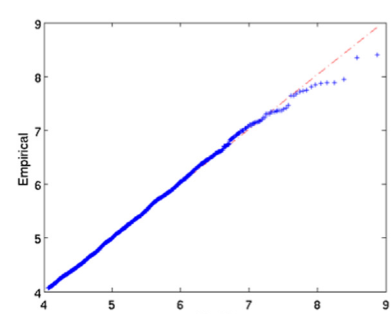
Site Name	Probability plot	QQ plot
EMEC		
Wavehub		
Outer Hebrides		
Belmullet		
North Sea		

Fig. 2. Diagnostic plots for the Generalized Pareto Distribution.

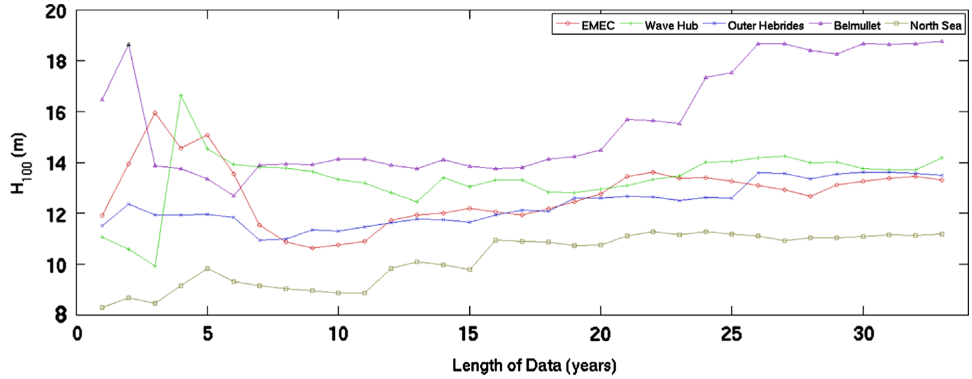


Fig. 3. Variation of 100-yr return values with period of data considered.

4.2. Return values

Having identified the POT approach of fitting a GPD to the set of ordered excesses above a high threshold as a suitable method, the *m*-year return value can be found by using the parameters of the model, provided that m is large. If k is the size of the set of excesses, and n is the number of years for which data is available, the average number of exceedances per annum, λ, is calculated as,

$$\lambda = \frac{k}{n} \tag{8}$$

Knowing the scale and shape parameters, σ and ξ, and the threshold u, the return value, H_m, for the GPD can be estimated by [19]

$$H_m = u + \frac{\sigma}{\xi}\left[(m\lambda)^{\xi} - 1\right] \tag{9}$$

4.3. Minimum data length

In the estimation of extreme wave heights for return periods as long as 50 or 100 years, the period of the data used in the calculation should be sufficiently long to obtain reliable estimates.

Table 2
p-Values from goodness-of-fit tests applied to the GEV models describing the population of storm maxima. The bold font indicates that the hypothesis cannot be rejected.

Site name	KS test (*p*-value)	χ^2 test (*p*-value)
EMEC	0.0645	4.93×10^{-3}
Wavehub	0.0944	0.0034
Outer Hebrides	0.0942	1.44×10^{-4}
Belmullet	**0.1086**	0.0499
North Sea	**0.1124**	9.92×10^{-7}

Table 3
p-Values from goodness-of-fit tests applied to the GPD and GEV models describing the population of all excesses above the threshold. The bold font indicates that the hypothesis cannot be rejected.

Site name	GPD (*p*-values)		GEV (*p*-values)	
	KS test	χ^2 test	KS test	χ^2 test
EMEC	**0.4014**	**0.3118**	2.26×10^{-4}	8.03×10^{-21}
Wavehub	**0.9509**	**0.3158**	2.56×10^{-5}	2.69×10^{-18}
Outer Hebrides	**0.6983**	**0.2106**	2.85×10^{-5}	1.88×10^{-23}
Belmullet	**0.8821**	**0.2859**	5.06×10^{-6}	6.75×10^{-18}
North Sea	**0.2106**	0.0203	1.41×10^{-6}	9.98×10^{-28}

It would be folly to estimate a 100-year return period based on one year's data, while at the same time, thirty or forty years' data might yield estimates that differ very slightly from estimates based on shorter periods. For instance, EquiMar protocols recommend using a dataset that has a duration of at least 20% of the return period (e.g. 10 years data for 50-yr return levels) [24], however this is merely a guideline. Thus, having little idea of what would constitute a sufficiently long period, one would tend to use all the available data and calculations are carried out with the hope that datasets which span over three to four decades, such as the ERA-40 and ERA-Interim reanalyses, would be sufficiently long and yield reasonably accurate results.

To ascertain the minimum period of data required to make estimates, the reanalysis data was treated in the reverse chronological order, i.e. most recent data first. The following iterative procedure was then applied, with the sample consisting of one year's data (i.e. 2011).

1. A threshold equal to the 95% quantile of the sample was used, and the POT approach was used, fitting a GPD to the ordered set of excesses.
2. The 100-yr return level was estimated using Eqs. (8) and (9), and recorded.

The sample period was increased by one year and the above steps repeated. Thus, the second iteration would use data from 2010 to 2011, the third would use data from 2009 to 2011, and so on. The final iteration would use the entire dataset, spanning from 1979 to 2011. If $H_{100}(1)$ is the 100-yr return value obtained from the first iteration, $H_{100}(2)$ from the 2nd iteration (thus based on 2 years' data), and so on, the set of estimates thus obtained would be $\{H_{100}(1), H_{100}(2)... H_{100}(33)\}$. These are plotted in Fig. 3.

It can be seen from the plot that for each of the locations, the difference between consecutive estimates of H_{100} is large when the period of data used is small; for instance, when one considers a duration less than 6 years the variation in extreme wave heights is large. For most of the locations, the graph stabilizes, or the variation seems to be less significant when the data-length is between 10 and 20 years, and shows little variation after flattening. This would indicate that using short periods of data, (for example, <7 years), would yield inaccurate estimates of long range return levels. An exception to this is the graph for Belmullet and the reasons for this behavior are discussed in Section 6.

To demonstrate this further, the set of estimates obtained was divided into 2 blocks, the first with estimates using 1–10 years' data, and the second using 10–20 years' data. Thus, the subsets obtained would be $\{H_{100}(1), H_{100}(2)\ldots H_{100}(10)\}$ and $\{H_{100}(11), H_{100}(12)\ldots H_{100}(20)\}$. The means and standard deviations for these subsets were calculated for the five sites and are tabulated in

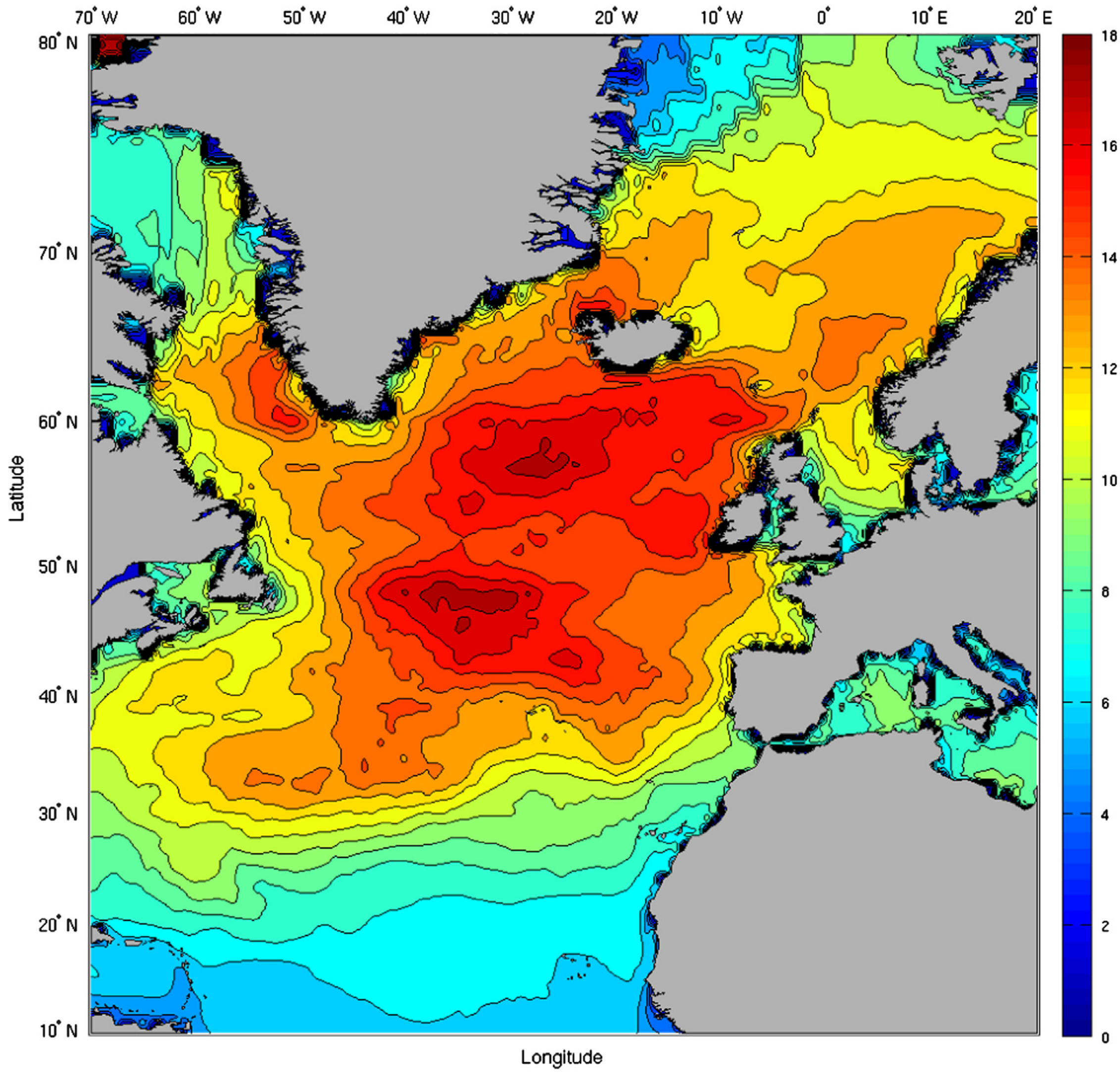


Fig. 4. Map showing the 100-yr return levels for the North Atlantic Ocean and North Sea.

Table 4. A comparison of the standard deviations for the five representative sites shows that in most cases estimates from the second subset lie closer to the mean (i.e. they are less spread) than estimates from the first subset. These results indicate that the minimum duration of data to be considered in extreme wave analysis may be as little as 10 years to obtain a reliable estimate.

5. Preparation of the 100-yr extreme wave map

The POT method, as described, was applied to the ERA-Interim data for the entire region, and the 100-yr return values (denoted as H_{100}) were calculated for each cell. Fig. 4 presents the contour map for the 100-yr return values for the North Atlantic region.

Table 4
Means and standard deviations for test locations.

Site name	First 10-year block		Second 10-year block	
	Mean (m)	Standard deviation (m)	Mean (m)	Standard deviation (m)
EMEC	12.88	1.87	12.01	0.47
Wavehub	13.12	1.92	13.01	0.28
Outer Hebrides	11.61	0.45	11.60	0.37
Belmullet	14.46	1.66	14.01	0.23
North Sea	8.97	0.41	10.27	0.65

The 100-yr return value estimates obtained were found to be consistently lower than those reported by previous studies [5,9,14], especially near the middle of the North Atlantic and the implications are discussed in Section 6.

5.1. Comparison with past results

The most desirable estimates for comparing the 100-yr return values obtained by using the POT-GPD method (described previously) would be from a similar process applied to buoy data. However, buoys in the region under study are too few, and are located too close to the coasts. Moreover, the available buoy data does not extend over a period sufficiently long to be able to reliably estimate the return values ($>$ 10 years, as demonstrated in Section 4.3).

To comparatively assess the performance of the Generalized Pareto Distribution in conjunction with the POT method, 100-yr extreme wave heights from the HSE report [14] were used in the absence of buoy data. The return values in the report [14] were calculated by fitting a 3-parameter Weibull distribution to the uppermost 5% of the wave data. The locations and spread of the

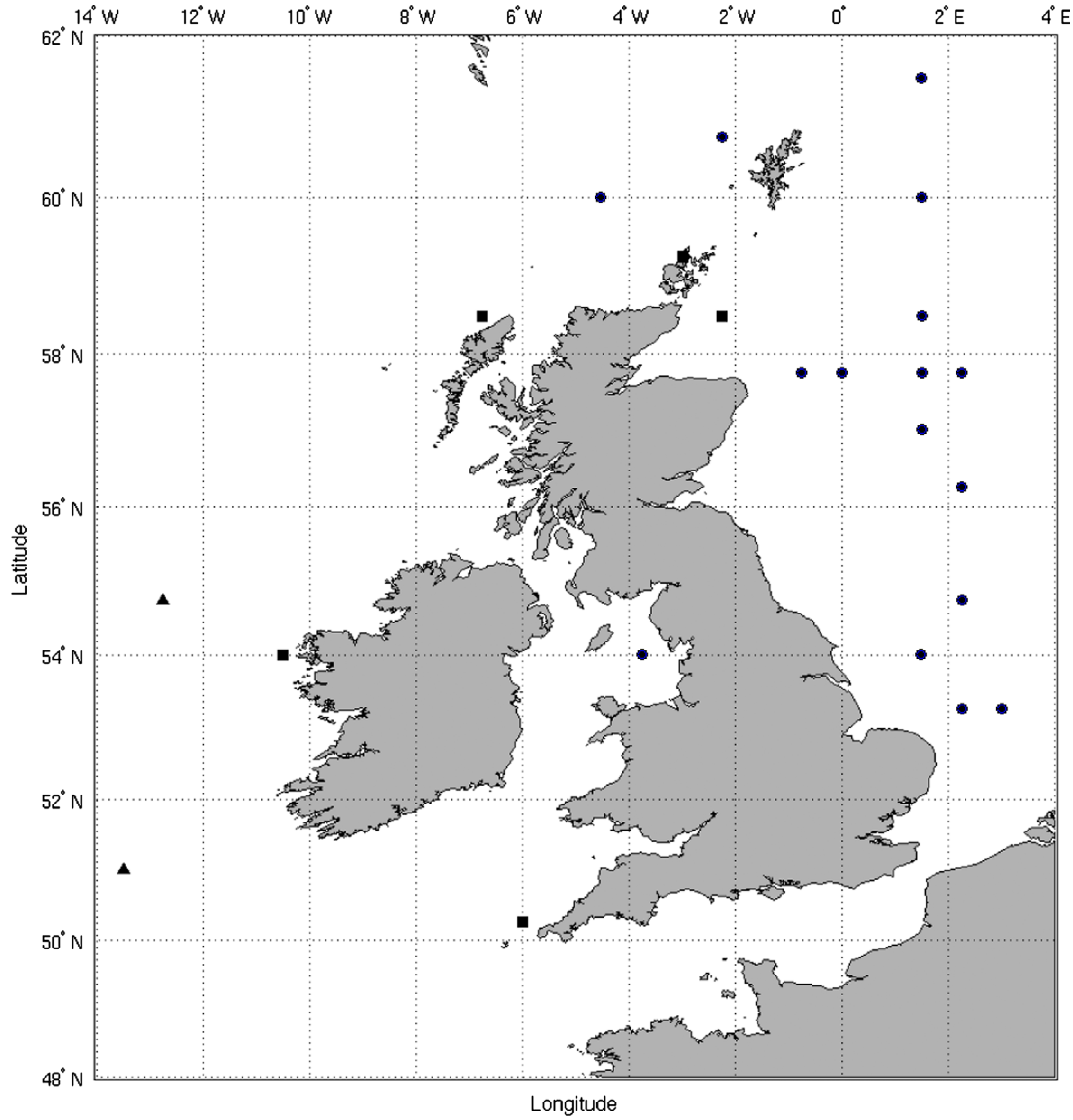


Fig. 5. Locations of test sites (square markers), UKMO buoys (triangular markers), and offshore platforms (circular markers).

platforms and buoys in UK waters are shown in Fig. 5 and the H_{100} estimates for these are tabulated in Table 5.

This comparison is also presented visually in Fig. 6. An examination of the scatter plot reveals that the ERA-Interim estimates are well correlated with buoy estimates and the Pearson's correlation coefficient was found to 0.976, indicating a strong linear relationship between the two. It can also be observed that the estimates obtained from the study are lower than the previously published estimates. The implications of this are discussed further in the following sections.

5.2. *Calibration*

The relatively lower estimates of the 100-yr return values obtained from the study would suggest, among other things, the possible need for calibration of the data source. As discussed in Section 2.1, the reanalysis data was used as is.

An alternate approach would be to calibrate the return values, as demonstrated by Caires and Sterl, 2005 [5]. However, in this case, it would be illogical to calibrate the POT/GPD return values with the results of other studies because the methods of population selection and parametric models used are different (e.g. Weibull or Gumbel distribution fit to a sample of storm maxima).

6. Discussion

The results obtained from the analyses conducted in the study stimulate discussion in several areas.

Figs. 1 and 2 present the findings from Section 4.1, using probability and quantile plots to assess the suitability of the GEV model applied to storm maxima compared to the GPD model applied to all excesses above the threshold. A visual comparison of these diagnostic plots shows that the Generalized Pareto Distribution was superior to the GEV model in describing the data, especially in the high-value region. This suggests that the treatment of all excesses over a high threshold might yield more robust return level estimates than data containing storm maxima obtained after a process of declustering. This is further corroborated by the p-values obtained from the goodness-of-fit tests (Tables 2 and 3).

Table 5
Comparison of POT-GPD estimates against previously published estimates from [14].

Buoy/ Platform	Nearest data gridpoint	H_{100} (m) from [14]	H_{100} (m) ERA-Int	Difference (%)
K2	51°N, 13.5°W	18.7	15.21	18.66
K4	54.75°N, 12.75°W	19.2	15.11	21.33
Thistle	61.5°N, 1.5°E	15.9	12.26	22.87
Rhum	60°N, 1.5°E	14.2	11.22	20.96
Miller	58.5°N, 1.5°E	14.2	11.19	21.18
Andrew	57.75°N, 1.5°E	13.5	10.92	19.10
Goldeneye	57.75°N, 0.75°W	13.3	9.36	29.60
Buchan	57.75°N, 0°E	13.2	10.25	22.38
Fulmar	57.75°N, 2.25°E	13.3	11.08	16.67
Curlew	57°N, 1.5°E	13.3	10.55	20.69
Auk	56.25°N, 2.25°E	13.3	10.56	20.57
Tyne	54.75°N, 2.25°E	9.3	9.16	1.51
Cleeton	54°N, 1.5°E	10.3	8.19	20.44
Carrack	53.25°N, 3°E	9.2	7.57	17.71
Leman	53.25°N, 2.25°E	7.9	7.27	7.97
Clair	60.75°N, 2.25°W	16.6	12.97	21.89
Foinaven	60°N, 4.5°W	18.0	14.10	21.66
Morecambe N.	54°N, 3.75°W	8.7	7.66	11.94

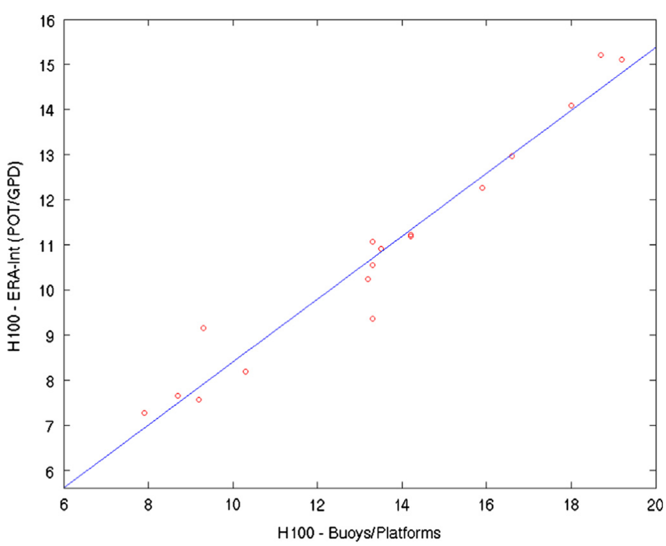


Fig. 6. Scatter plot of 100-yr return value estimates against previously published estimates from buoys/offshore platforms from [14].

An interesting observation can be made from the parameter estimates for the GEV models used to describe the sample of storm maxima listed in Table 1. For each of the test sites the shape parameter (ξ) is positive ($\xi > 0$) implying that the Fréchet (FT-II) type of extreme value distribution best describes the samples. In the traditional POT approach of fitting a Gumbel or Weibull model to the storm maxima, any inferences made rely on the assumption that the correct model was selected. The observation that an FT-II distribution (and not a Gumbel or Weibull distribution) best fits the data raises questions about the suitability of estimating extremes using Gumbel or Weibull distributions in UK waters, as well as the estimates thus obtained.

It is known that the extreme value distribution of maxima reduces to a Gumbel distribution when the parent distribution is either Weibull or log–normal. As mentioned earlier, for the five locations, it was an FT-II model that best described the data, suggesting that for these locations, and perhaps the rest of the region, the parent distribution might not be Weibull or log–normal.

The results obtained in Section 4.3 on the variability in the 100-yr return levels with the length of data considered, shows that for all locations (with the exception of Belmullet) the variation in return levels decreased, as the duration of data is increased, with the difference between consecutive estimates becoming very small when at least 10 years of data is used as can be seen in Fig. 3.

For any sample, it is known that a large standard deviation would imply that the values are spread away from the mean, and a small standard deviation would imply that the values are clustered close to the mean. It is interesting to note from the statistics presented in Table 4, that for most of the sites, the difference between the means of the two subsets is not very large, but the difference in the standard deviations is appreciable, from which it can be inferred that the second subset is less spread than the first subset. It is obvious from this and the evidence from Fig. 3 that using data of short durations (e.g. 7 years) would yield less reliable estimates than data of longer periods. These appear to suggest that data of a minimum length of 10 years may be sufficient to make reliable estimates of extreme conditions. This is in accordance with the observation made by Goda [10] that the minimum duration of a data record should be 10 years.

The Belmullet site also seems to follow this trend (of high initial variations in estimates which then reduce significantly) until data from the year 1988 was included in the analysis (which treated data in the reverse chronological order). Unusually high waves in 1986, 1988, and in 1991 (shown in Fig. 7) distort the curve, and cause the sudden rise in the return level observed around the 24 year mark. The data used for the analyses corresponds to a location approximately 25 km off the coast of Ireland and for a depth of approximately 130 m. The authors carried out the analyses on the basis that these large waves are genuine and not a result of errors in measurement. This, however, is not verifiable with the information available. If these values are indeed a result of errors, it is possible that they may have falsely influenced the return level estimates. In such a case, the authors expect that this site too will exhibit a trend similar to the other sites when these waves are excluded.

The study shows that ERA-Interim data tends to underestimate the return values when compared with estimates obtained from buoy data. Because of this, a correction or calibration of the data may be required when this hindcast is used. To the best of the authors' knowledge, no such evaluation has been done for the ERA-interim reanalysis, and should this be the case, the results obtained from the method assessed in this study are likely to be different.

Another possible reason for the lower estimates is that the ERA-Interim data contains H_s data from simulations sampled at 6-hourly intervals. It is possible that the waves of maximum height in storms occur between observations, and are thus not recorded due to the low sampling rate. It might be possible to get better estimates from data with a higher sampling rate, and the need for calibration and correction may be eliminated.

Without data extending over the entire duration of the return period, i.e. 100 years, it is impossible to ascertain which method – fitting a GEV model to storm maxima or fitting the GPD model to all exceedances – yields more accurate estimates. However, should the POT/GPD approach yielding lower return values be more accurate, it would imply a considerable savings in the capital costs of offshore installations like wave and tidal energy converters. It would directly translate to a savings in material costs, among other things.

On inspecting the contour map presented in Fig. 7, it can be seen that the significant height of extreme waves varies from one location to another, i.e. it is location specific. A comparison of the contour maps for extreme waves in Scottish Waters (Fig. 8) and south west UK (Fig. 9) shows that return values can vary quite significantly, from a maximum of about 13 m in the exploitable areas around the South West coast of UK, to a maximum of approximately 16 m in the exploitable areas off the coast of Scotland. This would imply the need for WECs to be designed keeping in mind the extreme wave height of the region of intended deployment. However, as full customization of devices for each wave energy site is not a financially viable option in the long run, the design process is further complicated by the need to optimize between robustness, cost and universality in site selection. Note that Figs. 8 and 9 were prepared by interpolating the H_{100} values with resolution of $0.75^{\circ} \times 0.75^{\circ}$ by the cubic interpolation method provided by Matlab. Caution must be exercised when interpreting these figures because near-shore data have not been verified with independent sources.

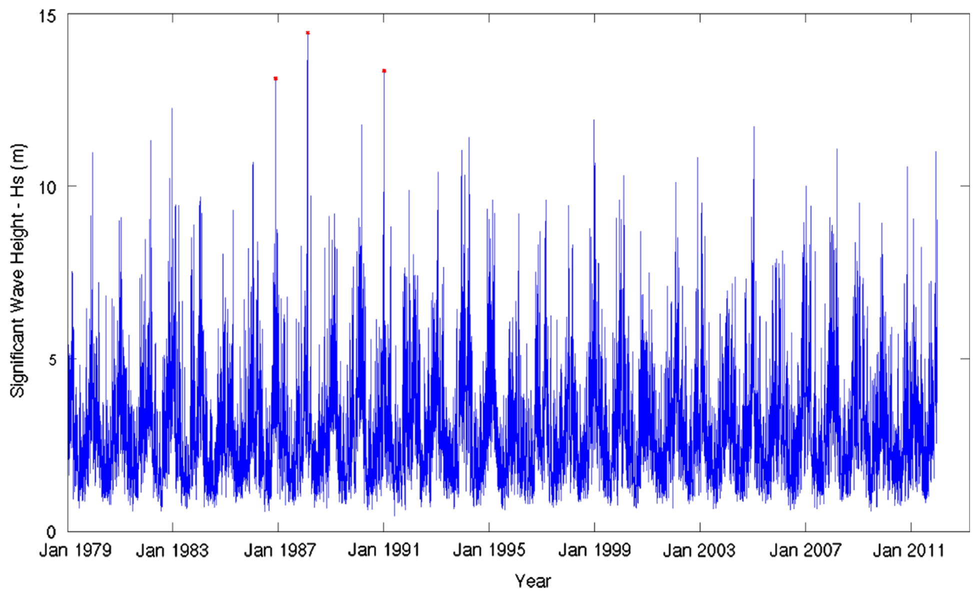


Fig. 7. Time series plot for the Belmullet location showing unusually high waves.

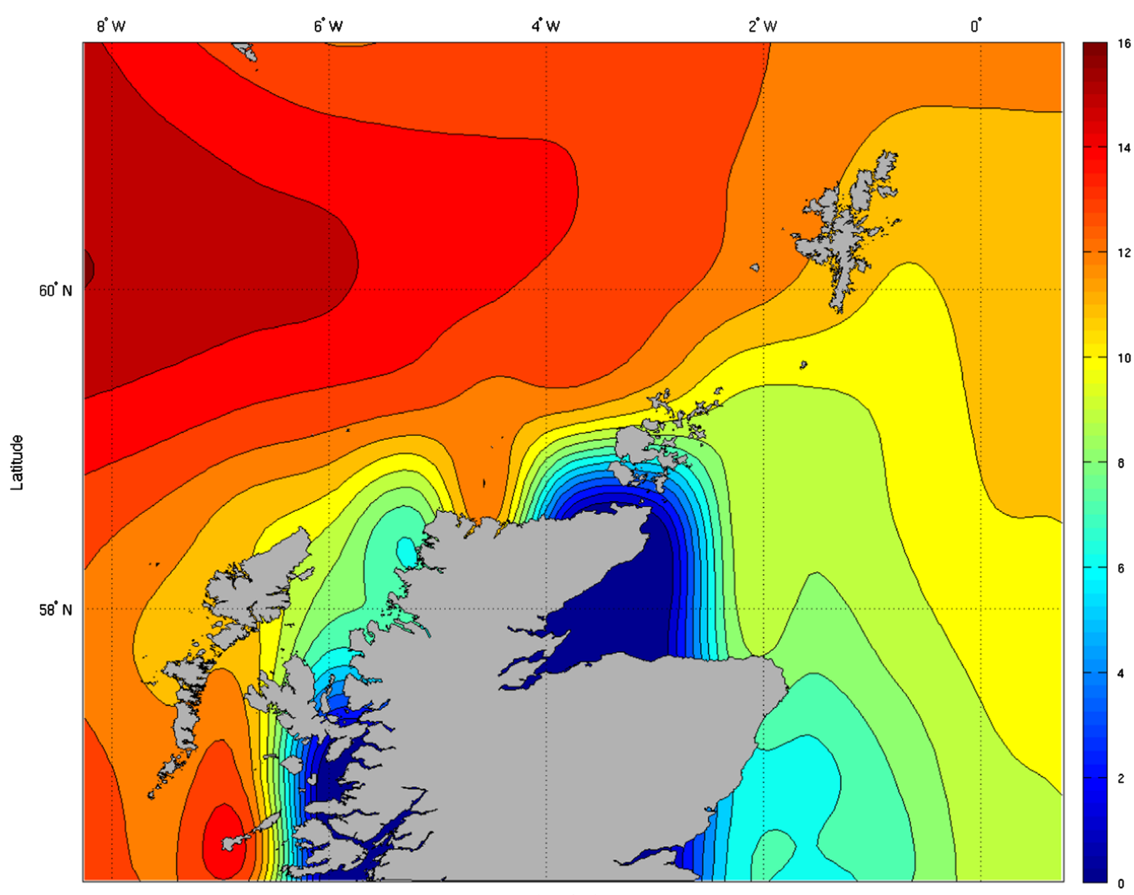


Fig. 8. 100-yr extreme waves in Scottish waters (m).

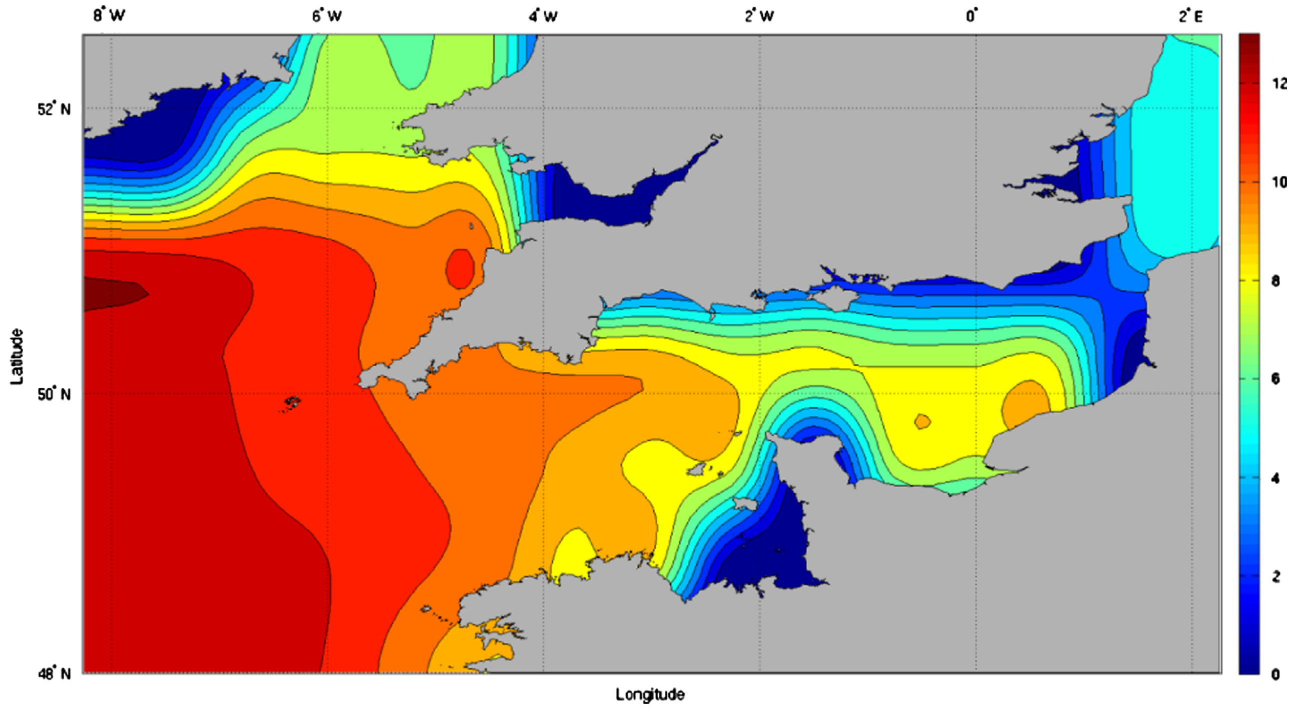


Fig. 9. 100-yr extreme waves in south UK (m).

There are also some limitations to the approach followed in the study, such as the selection of threshold as the 95% quantile, which is, perhaps, simplistic. While this method of selecting a threshold yielded good fits, there is little scientific justification to selecting the 95% quantile, and not any other quantile as the threshold. Also, seasonal variations in the data were not taken into account. It must also be kept in mind that the standard errors associated with the results will need to be suitably inflated to account for the non-independence of the data. A method for evaluating the new error bounds and confidence intervals will need to be arrived at. This is in accordance with the findings of Anderson et al. [25].

The limitations of wave data from buoys—namely the lack of availability of data on decadal scales, their concentration near coasts and sparseness in deeper waters make it difficult to assess the robustness of estimates obtained from reanalysis data (ERA-Interim data). The use of satellite altimeter data in conjunction with buoy data might be a possible solution to this. However, since satellite data is irregularly sampled, parameters of a suitable model to fit the data may be estimated, but the difficulty of obtaining meaningful return levels from these still exists.

The description of extreme waves for the design of offshore structures and vessels consists of an extreme wave height along with an associated extreme wave period. This study focused on the estimation of the extreme wave height only, and the preparation of a 100-yr return level map for the North Atlantic region.

Having said that, the analyses carried out and results produced would be useful when looking at potential sites for marine energy development, both near-shore and in deep waters, in the North Atlantic region and North Sea. For such sites, previously obtained results might prove inadequate because of the coarseness of data used and their geographic coverage. While the results of the study are generic and have potential application in general marine and offshore engineering, the 100-yr return value map of the North Atlantic Ocean and North Sea presented in the study is particularly useful for marine energy developers during the site selection stage. It may serve as a quick guide to identify regions where extremes lie within the design criteria of the devices to be deployed. Conversely, once a site has been identified, it may serve as a guide to the selection of devices based on the extreme conditions anticipated.

It must be kept in mind when using these results that the data used in the study is the product of a model and data from each grid-point is not individually verified by using independent measurements, e.g. from buoys.

7. Conclusions

This study used the ERA-Interim dataset produced by ECMWF, spanning over 33 years from January 1979 to December 2011, to investigate different methods of estimating extreme wave conditions, and produced a 100-yr return level map for the North Atlantic region and North Sea. The different approaches to estimating 100-yr return values were reviewed and data analyses were conducted. The results of the study are summarized below.

The block maxima approach, by dividing clusters based on storms, and the POT method, treating all excesses above a threshold, were compared. Using a Generalized Pareto Distribution to fit all excesses above a high threshold yielded more reliable estimates of return values than the approach of fitting a GEV model to the storm maxima.

Among the extreme value distributions, the FT-II type of distribution best fit storm maxima. This raises doubts about the suitability of the traditional method of fitting a Gumbel or Weibull distribution to the sample of maxima, especially for the region studied. This also suggests that the parent distribution of wave height data from the region might not be either Weibull or log–normal.

The effect of the duration of data on the return value estimates was investigated and the results suggest that using short periods of data (e.g. less than 7 years) may yield inaccurate results. Using data with a duration of more than 10 years is likely to produce more reliable return value estimates.

100-yr return values were estimated by fitting a GPD to all excesses above a high threshold, selected as the 95% quantile. These values were compared with previously published estimates from buoys and offshore platforms found to be lower in comparison. This may be because the raw data used (ERA-Interim) may require calibration, or may need to be sampled at a higher rate (<6 hourly intervals). It may also be that these estimates are more accurate, which would have a direct consequence on the economics of the structures and devices.

A 100-yr return level map was produced for the North Atlantic Ocean and North Sea, demonstrating the POT method applied to all excesses. By interpolation, maps of areas of marine energy interest around the UK were also produced.

Acknowledgment

The authors are grateful for the support of the Engineering and Physical Sciences Research Council (EPSRC) through the Adaptation and Resilience In Energy Systems consortium.

References

[1] Ochi MK, Whalen JE. Prediction of the severest significant wave height. Coastal Engineering 1980:587–600.
[2] Ochi M. New approach for estimating the severest sea state from statistical data, Coastal Engineering; 1992.
[3] Carter DJT, Challenor PG. Return wave heights at seven stones and famita estimated from monthly maxima; 1978.
[4] Castillo E, Sarabia JM. Engineering analysis of extreme value data: selection of models. Journal of Waterway, Port, Coastal Ocean Engineering 1992;118(no. 2):129–46.
[5] Caires S, Sterl A. 100-Year return value estimates for ocean wind speed and significant wave height from the ERA-40 data. Journal of Climate 2005;18: 1032–48.
[6] Caires S. Projection and analysis of extreme wave climate. Journal of Climate 2006;19:5581–605.
[7] Neelamani S, Al-Salem K, Rakha K. Extreme water waves in the UAE territorial waters. Emirates Journal for Engineering Research 2006;11(no. 2):37–46.
[8] Carter DJT. Estimating extreme wave heights in the NE Atlantic from geosat data, London: Oceanographic Sciences Deacon Laboratory for the Health and Safety Executive 1993.
[9] Wimmer W, Challenor P, Retzler C. Extreme wave heights in the North Atlantic from altimeter data. Renewable Energy 2006;31:241–8.
[10] Goda Y. Random seas and design of maritime structures. 3rd ed. Singapore: World Scientific Publishing Co Pte Ltd; 2010.
[11] Wolfram J, Linfoot B, Venugopal V. Some results from the analysis of metocean data collected during storms in the Northern North Sea. Underwater Technology 2001;24(no. 2):153–63.
[12] Det Norske Veritas, Recommended Practice DNV-RP-C205: Environmental conditions and environmental loads; 2010.
[13] Dee DP, Uppala SM, Simmons AJ, Berrisford P, Poli P, Kobayashi S, et al. The ERA-Interim reanalysis: configuration and performance of the data assimilation system. Quarterly Journal of the Royal Meteorological Society 2011;137 (no. 656):553–97 April.
[14] Williams MO. Wave mapping in UK waters. RR621 ed.HSE Books; 2008.
[15] Ferreira JA, Soares CGuedes. An application of the peaks over threshold method to predict extremes of significant wave height. Journal of Offshore Mechanics and Arctic Engineering 1998;120:165–76.
[16] Holthuijsen LH. Waves in oceanic and coastal waters. Cambridge: Cambridge University Press; 2007.
[17] Soares CGuedes, Henriques AC. Long term predictions of significant wave heights at Sines and Faro. Littoral 1994;94:343–56.
[18] Thompson P, Cai Y, Reeve D, Stander J. Automated threshold selection methods for extreme wave analysis. Coastal Engineering 2009;56:1013–21.
[19] Coles S. An introduction to statistical modeling of extreme values. Springer Series in Statistics; 2001.
[20] Mathiesen M, Goda Y, Hawkes PJ, Mansard E, Martin MJ, Peltier E, et al. Recommended practice for extreme wave analysis. Journal of Hydraulic Research 1994;32(no. 6):803–14.
[21] Haver S, Nyhus KA. A wave climate description for Long term response calculations, In: Proceedings of the fifth international offshore mechanics and arctic engineering symposium; 1986.
[22] Dupuis DJ. Exceedances over high thresholds: a guide to threshold selection. Extremes 1998;1(no. 3):251–61.
[23] Tancredi A, Anderson C, O'Hagen A. Accounting for threshold uncertainty in extreme value estimation. Extremes 2006;9:87–106.
[24] EquiMar, Protocols for the equitable assessment of marine energy converters, 1st ed., Institute for Energy Systems, School of Engineering, Edinburgh: University of Edinburgh; 2011.
[25] Anderson CW, Carter DJT, Cotton PD. Wave climate variability and impact on offshore design extremes, 2001.