Extreme Value Analysis of Wave Energy Converters

K. Doherty[1], M. Folley[2], R. Doherty[3], T. Whittaker[2]

[1]Aquamarine Power Ltd.
Edinburgh, Midlothian, UK

[2] Wave Energy Research Group
Queens University, Belfast, UK

[3] Formerly of Aquamarine Power Ltd.
Edinburgh, Midlothian, UK

ABSTRACT

This paper presents a statistical Extreme Value Analysis (EVA) methodology to evaluate the design survivability condition of a Wave Energy Converter (WEC). The technique is applied to the Oyster® WEC design which is being developed by Aquamarine Power Ltd. (APL) but can be easily modified to investigate other technologies that have different operational philosophies. The approach presented considers the extreme statistics of both the incident wave climate at the device location and the complex dynamic response of the WEC in such conditions. This give a more robust evaluation of the WECs survivability condition than a more traditional single design wave approach.

KEY WORDS: survivability; wave energy converter; Oyster; extreme value analysis; peak-over-threshold.

INTRODUCTION

An Extreme Value Analysis (EVA) technique is presented in this paper to evaluate the loading survivability condition of a Wave Energy Converter WEC. Determining the maximum expected load experienced by a WEC during its operational lifetime is of paramount importance to the design process, in particular to the design of the foundation and/or mooring system. Thorough and consistent EVA can reduce/remove the implementation of large safety factors to a WEC design load condition which can have significant cost benefits. Extreme loads experienced by a WEC are not only dependent on the nature of the wave climate within which it is operating but also on the phase relationship between each incident wave and the WECs dynamic response. These unique features make it necessary to develop a custom EVA technique with specific application to WECs. A possible hazard of developing an EVA methodology is that it relies heavily on novel analysis techniques. Despite its mathematical elegance, industry confidence in such a technique is often low because of the novel approaches employed. Consequently large safety factors are often applied to the design load undermining the goal of using EVA. The ethos adopted in this paper is to apply well accepted techniques to minimise uncertainty and hence the subsequent use of excessive safety factors connected with the modelling approach itself. The methodology to determine a WECs design survivability condition can be split into two sections. The first determines the extreme sea states, ranked in terms of their return period, that the WEC is expected to encounter during its lifetime. These sea states are then used in an extensive series of wave tank tests to measure the loads experienced by WEC under such extreme conditions. An EVA technique is then applied to this set of load test data to evaluate the WECs expected survivability load condition.
As a test case the EVA methodology developed is applied to the Oyster® WEC design which is being developed by Aquamarine Power Ltd. Oyster® is a buoyant, seabed mounted, oscillating flap which pierces the water surface and is located in the near-shore region of 10m–15m water depth. In 2009 Aquamarine Power Ltd. successfully deployed a first generation 315kW full scale Oyster® prototype at the European Marine Energy Centre (EMEC) site in Orkney, which has operated it for over 6000 hours. The second generation Oyster® project, a 2.4MW array of three devices is scheduled for deployment at EMEC in the summer of 2011. The survivability strategy of Oyster® has some advantages over other WECs. It is located in the near-shore environment where the incident extreme waves are subject to depth-limited wave breaking and have a significantly reduced directional spread due to wave refraction. In addition to this Oyster® has an inherent safety mechanism in that it decouples from larger waves by rotating to greater angles. Despite these advantages the survivability load conditions must still be evaluated to inform the design of Oysters foundation solution which anchors it to the seabed.

Before the methodology is applied to the Oyster® WEC and results displayed an introduction to the mathematics of the EVA techniques employed is first presented.

EXTREME VALUE METHODS

In recent years the significance of extreme value analysis has become ever more important to a wide variety of disciplines. This has lead to a large development of eloquent and novel statistical techniques, which unfortunately have been accompanied by a gross misuse in their application. Other simpler approaches such as inferring extreme values from an empirically-approximated parent distribution also leads to large uncertainties due to the fact that the extreme values are very often driven by a different physical process than the parent population on which the approximate distribution is based. In addition to this, basic empirical fitting is subject to interpretation as very often there is no basis for choosing one type of distribution over another.

There are however 2 extreme value techniques which have a sound mathematical origin for the prediction of extreme events. These are:

i. the Peak-Over-Threshold (POT) method
ii. the Block Maxima (BM) method

The reader is referred to (Cole, 2001) for a comprehensive description of the theory behind each of the methods. The application of either method to a given data set is strictly only valid if the data set is an *independent and identically distributed* (i.i.d) sequence. It is possible to take into account the effect of dependent sequences, the most commonly accepted approach is with the use of Markov chains. However, this is beyond the scope of the techniques employed here as it can very often complicate the methodology, introducing significant uncertainty in the results which is contradictory to the ethos adopted.
Assuring that a data set is i.i.d. is a very important step to avoid distortion of results from the extreme value technique. Assessing if a data set is an independent sequence can be achieved with use of the correlation coefficient defined as

$$r(w) = \frac{1}{\sigma_d^2 (N-w)} \cdot \sum_{i=1}^{N-w} \left(L_i - \overline{L}\right)\left(L_{i+w} - \overline{L}\right) \quad (1)$$

where L_i is the i[th] event from a set of N data points. $\overline{L}$ and σ_d are the mean and standard deviation of the data set respectively and w is the lag number. For example a lag of $w = 1$ computes the correlation coefficient of adjacent events in the load data set. A zero correlation coefficient indicates that events are statistically independent.
Assessing if a data set is identically distributed is not as straight forward as checking for independence. The first step to ensure an identically distributed set is a good understanding of the underlying physics. This enables unrelated events to be filtered out of the data set. The second step is intrinsically executed by application of the POT and to a less extent the BM method. This point will be highlighted in each of the next two sections.

The Peak-Over-Threshold Method

The POT approach generates a subset of data points from a parent set by only considering those events (data peaks) above a defined threshold. By only considering peaks above a threshold the data is more than likely to be from the same distribution. This intrinsically assists with obtaining an identically distributed data set. In addition to this, provided the data peaks can be considered statistically independent, thus the i.i.d. condition is satisfied, the distribution of the peak events should have a Generalised Pareto distribution. The Generalised Pareto Cumulative Distribution Function (CDF) is given by

$$F_{GP}(x) = 1 - \left(1 + k\left(\frac{x-u}{\sigma^*}\right)\right)^{-\frac{1}{k}} \quad (2)$$

where $k$is the shape parameter, σ^* is the Generalised Pareto scale parameter and u is the threshold level. The distribution has an infinite "upper-bound" if k>0 or $k = 0$. However, if $k < 0$ the distribution is bounded and so it describes the extreme values of a finite tailed parent distribution indicating that the underlying physical process will reach a maximum. The upper bound of the Generalised Pareto distribution in this case is given by

$$U_{GP}^b = u - \frac{\sigma^*}{k} \quad (3)$$

The specification of the Generalised Pareto distribution requires the definition of a threshold value. It has been recognised that the choice of threshold can have a significant influence on the shape of the distribution. A low threshold is likely to violate the asymptotic basis of the model, whilst a high threshold may produce too few data points to produce a good estimate. A number of methods for identification of a suitable threshold have been proposed (Cole, 2001; Thompson, et al. 2009) which are assisted by some attractive properties of the Generalised Pareto distribution. These are

a) The shape parameter k should be constant above a valid threshold
b) The re-parameterised scale parameter $\hat{\sigma} = (\sigma^* - ku)$ is constant above a valid threshold
c) The mean of the excess above the threshold varies linearly with threshold level

The mean of the excess is given by

$$\overline{(x-u)} = \frac{1}{n}\sum_{i=1}^{n}(x_i - u) \cong \frac{\sigma^*}{1-k} \quad (4)$$

Properties (a) and (b) indicate a suitable threshold level above which the data is identically distributed. If (c) is satisfied it is an indication that the Generalised Pareto distribution is valid. This is usually displayed as a *mean residual life plot* which plots the mean of the excess over the threshold against the threshold values.

The Block Maxima Method

This method essentially describes the distribution of maxima of a large set of data. Historically this technique has been applied to large scale physical processes such as tidal levels or global temperatures where the effect of seasonality is accounted for by choosing a data set to be 1 year. Thus this method is often referred to as the "annual-maximum method", however, there is no strict rule on choosing 1 year as the standard data set size.
In a similar fashion to the POT method, the limiting distribution of the maxima of a large data set is the Generalised Extreme Value distribution, provided the data is i.i.d.. The Generalised Extreme Value CDF is given by

$$F_{GEV}(x) = \exp\left(-\left(1 + k\left(\frac{x-\mu}{\sigma}\right)\right)^{-\frac{1}{k}}\right) \quad (5)$$

where k , σ and μ are called the shape, scale and location parameters of the Generalised Extreme Value distribution respectively. Interestingly k is the same shape parameter as in a corresponding Generalised Pareto distribution and there is a relationship between all the parameters of the Generalised Extreme Value and Generalised Pareto distributions which is given by

$$\sigma^* = \sigma + k(u - \mu) \quad (6)$$

The Generalised Extreme Value formula given in Eq. 5 encapsulates three families of distributions labelled Type I, Type II and Type III (or Gumbel, Frechet and Weibull) depending on whether the shape parameter is $k = 0$, $k > 0$ or $k < 0$ respectively. Type I describes the extreme values of an exponentially-tailed parent distribution such as a Gaussian. Type II describes the extreme values of a polynomial tailed parent distribution such as a Student's t. Both Type I and Type II have an infinite "upper-bound". A type III however describes the extreme values of a finite tailed parent distribution. Similar to Eq. 3, the upper bound of a Type III distribution is given by

$$U_{GEV}^{b} = \mu - \frac{\sigma}{k} \quad (7)$$

Applying a Generalised Extreme Value Distribution to an i.i.d. data set is done as follows: The data is blocked into a sequence of observations of length *n*. The maximum value of each block is found. Supposing there are *m* blocks, then the Generalised Extreme Value distribution is applied to this set of *m* maxima. The fact that only the maximum of each block is selected assist further in obtaining an identically distributed extreme value set.
The underlying theory states that Eq. 5 will be the governing distribution of the *m* extremes provided $n \rightarrow \infty$. Having an infinite block size is obviously impossible in practice but provided *n* is sufficiently large Eq. 5 should be a valid approximation to the distribution of the extreme values. Thus the choice of block size is very important to the implementation of the Generalised Extreme Values distribution. If the block size is too small the $n \rightarrow \infty$ condition is strongly violated. Alternatively, large block sizes will dramatically reduce the amount of data utilised. In contrast to the POT method there are no guidelines or mathematical properties which assist in the appropriate selection of block size.

Extreme Value Model Verification

It can be seen that there are some formal threshold selection rules when applying the POT method to a data set. This has some advantage over a more arbitrary block size selection in the BM method. However, the BM technique is less sensitive to the independence constraint, i.e. even if the data consists of a dependent sequence the distribution of block maxima can still be of a Generalised Extreme Value type. The advantage/disadvantage of each method makes it difficult to select on a general basis one method over the other. Where possible both methods should be employed. One benefit of applying both techniques is that it provides a cross check mechanism on the parameters and results obtained due to the relationship between the Generalised Extreme Value and Generalised Pareto distributions.
For both extreme value methods, model verification is essential. Graphical inspection of the fitted model to the empirical data is one of the most useful verification techniques. Four different plots can inform the quality of the statistical model to the data set. These are

- Probability Plot: plots the model CDF against the empirical CDF
- Quantile Plot: plots the modelled inverse CDF against the ordered data
- Return period plot
- Probability Density Function (PDF) plot

The reader is referred to (Cole, 2001) for further details on each of these tools. The return period is defined as the time during which an event is exceeded on average once, it is calculated using

$$R = \frac{1}{A\left(1 - \hat{F}\right)} \quad (8)$$

Where A is the number of events per unit time (e.g. per year) and $\hat{F}$ denotes the model CDF e.g. Generalised Pareto or Generalised Extreme Value. Both the Probability and Quantile plots are the most informative to the model diagnostic process as they should be linear if the modelling technique is valid. It is more difficult to identify model quality from the return period and PDF plots but they do provide a secondary verification check.

DETERMINATION OF EXTREME SEA STATES

This section describes the methodology used to derive the extreme sea states that a WEC is likely to encounter in its lifetime. In particular the study focuses on the Oyster® WEC located at the EMEC site in Orkney.
The ultimate scope of this entire piece of work is to evaluate Oysters survivability load conditions, it is expected that wave height will have the dominant effect on the extreme loads, potentially modified by the wave period due to the dynamics of the device and the effect of wave steepness on wave breaking and water particle acceleration. Thus, the significant wave height (H_s) and spectral mean wave period ($T_m = m_0/m_1$) are analysed for specification of the extreme sea-states. Also at this stage of the analysis the POT method is selected over the BM method to conduct the EVA on the wave climate conditions as seasonal variations in significant wave height are only accounted for by selecting block sizes of 1 year. This has severe restrictions on the amount of data utilised in the BM method which would significantly increase the uncertainty of the estimate of the extreme values.

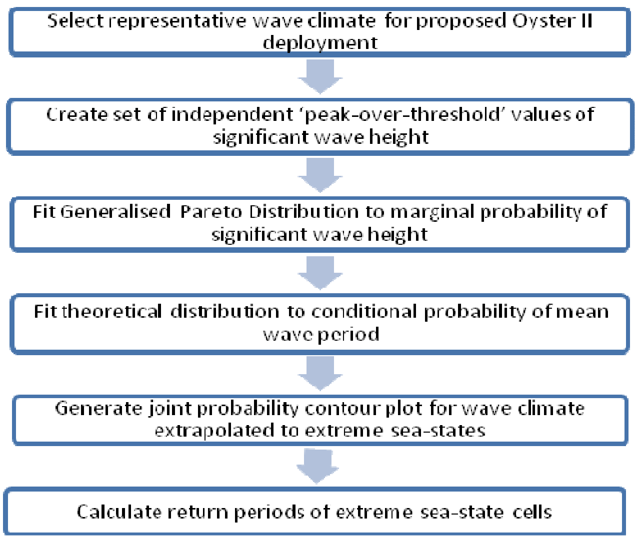

Fig. 1 Outline of methodology for specification of extreme sea state

Fig. 1 shows a flow chart of the methodology to define the extreme sea states. Each of these steps will be discussed below.

Selection of Representative Wave Climate

Twelve years of wind and wave hind-cast data from the NOAA Wave

Watcher III model has been used to force a Mike21 spectral wave model of the approaches to EMEC to generate a representative wave climate for the Oyster® site. Where possible the modelled wave climate was matched to direct ADCP measurements recorded at the EMEC site and was found to compare well. The data set contains a total of approximately 35,000 records representing the 3-hour averages of sea-states. Data is available for the significant wave height, mean wave period, peak wave period, peak wave direction, directional spread and the water depth. As stated earlier significant wave height and spectral mean period are chosen as the primary metric used to define the sea states. The mean wave period is preferred over the peak wave period as it is less sensitive to the exact shape of the wave spectrum, which may not always be known.

Significant Wave Height POT Data Set

The selection of a suitable H_s threshold level is key in achieving a robust data set. This is done by inspecting the behaviour of the fitted Generalised Pareto distribution parameters and mean of the excess as outline in the Peak-Over-Threshold section.

Although the POT method intrinsically assists with selecting an identically distributed data set, it is good practice to actively ensure that this condition is satisfied. Different methods have been proposed to impose statistical independence of events. (Morton and Bowers 1996) proposed the assumption that meteorological events are independent over a period of 30 hours. As the wave climate is closely correlated to meteorological events this assumption was supported by analysis off wave heights of off the coast of Lowestoft. The methodology adopted here however defines the temporal extent of each event to last until the significant wave height has dropped to 1.0 metres less than the threshold, thereby ensuring that there is a period of relative calm between events. It is suggested that these periods of calm provide delimiters to meteorological events to ensure that they are statistically independent. The statistical independence of the peak events selected using this method is tested by calculating the correlation coefficient given in Eq. 1. Analysis of the current data set results in a correlation coefficient of -0.18 which suggests that the peak events are indeed independent.

Significant Wave Height Marginal Probability Distribution

A Generalised Pareto distribution is used to fit to the extreme wave height data. The significant wave height data utilised is over all wave periods and so is referred to as a marginal probability distribution.

Fig. 2 contains four graphs that provide the necessary information to select a suitable threshold level. It should be noted that all fitting procedures used to determine the distribution parameters are carried out using the Maximum Likelihood Estimator (MLE) technique. Although there are several alternative methods the MLE technique was found to be adequate. The top left graph of Fig. 2 is the *mean residual life plot* and it can be seen that the graph appears to have two slopes with the transition at a threshold of about 5.5 metres. However, the Generalised Pareto distribution coefficients, $\boldsymbol{k}$ and $\boldsymbol{\sigma}^*$ vary dramatically above a threshold of 5.5 metres indicating that this may be due to a reduced sample size resulting in greater scatter. In addition, the distribution parameters are approximately constant above a threshold of 4.0 metres (ignoring the data above 5.5 metres) and so this is the threshold used for analysis. Fig. 3 shows the four plots used to check the validity of the applied model as explained in an earlier section. For the Generalised Pareto distribution to be a reasonable approximation of the actual wave height distribution then the probability plot and the quantile plot should both be approximately straight lines. In addition, the return level plot should demonstrate a reasonable extrapolation of the data to extreme values and the probability density plot should show that the model (line) is a reasonable approximation of the actual data distribution (bars). It can be seen in Fig. 3 that these conditions are all satisfied reasonably well, implying that the model is a reasonable approximation of the actual distribution. The return level plot in Fig. 3 also shows the 95% confidence levels (dotted lines) for the significant wave height. These confidence levels have been generated using a case re-sampling bootstrap method.

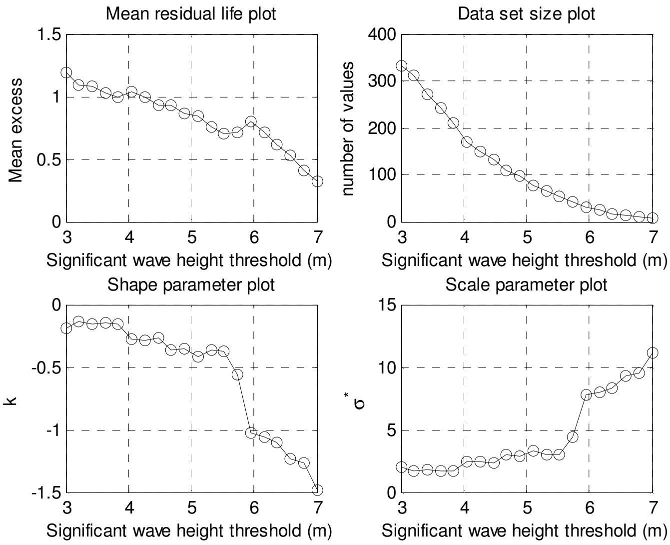

Fig. 2 Graphs use to determine a suitable H_s threshold value

This method re-samples the data to generate statistics of the variability of the data by assuming that the actual data is a reasonable representation of the underlying data distribution (and that this data all comes from the same distribution - identically distributed). It can be seen that the significant wave height with a 10,000 year return period is expected to be 7.8 -10.6 metres (95% confidence levels) with an expected mean value of 8.8 metres. It is important to note that the uncertainty associated with the estimate of the maximum significant wave height with a particular return period is separate from the probability of the significant wave height actually occurring as the former is the uncertainty associated with the estimation.

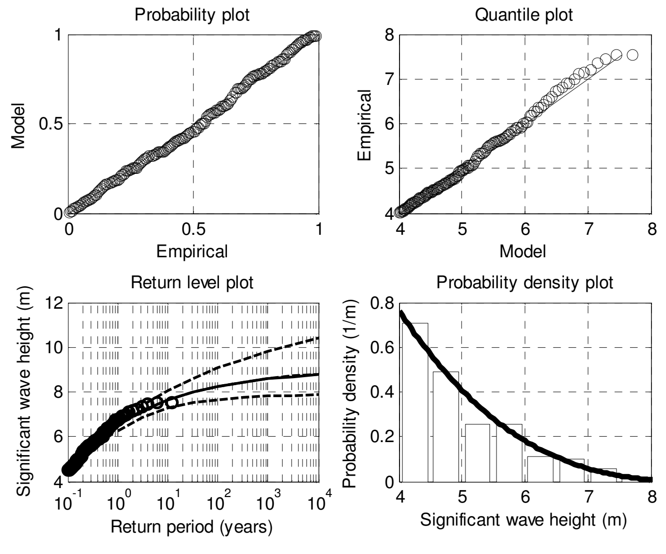

Fig 3. Model verification

Mean Wave Period Conditional Probability Distribution

Extreme sea states are defined in terms of both significant wave height and mean period. The extreme significant wave height distribution was determined in the previous section and so a complementary methodology is required to establish the distribution of the mean wave period.

A Log-normal distribution is used to fit the probability distribution of the mean wave period by binning the significant wave height data into bins with a width of 0.5 metres. This is referred to as a conditional probability distribution.

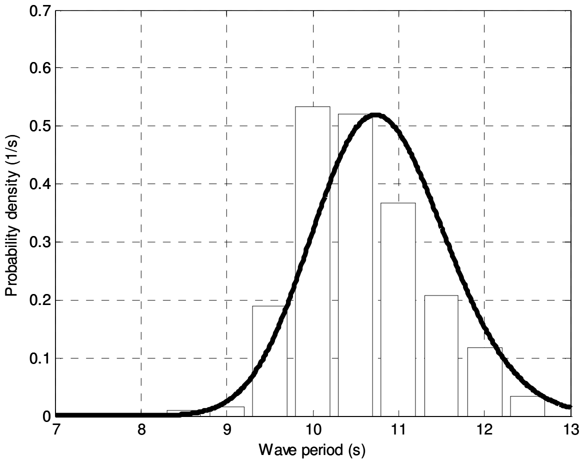

Fig. 4 Probability distribution of mean wave period at 4m<H_s<4.5m

A Log-normal distribution has been shown before to successfully describe the distribution of wave period (Burrows and Salih, 1986; Mathisen and Bitner-Gregersen, 1990) and adopting tried and tested techniques is in keeping with the ethos of this paper. The validity of a Log-normal distribution is also demonstrated in Fig. 4, which shows the conditional probability distribution of wave period for peak significant wave heights between 4.0 and 4.5 metres. An MLE is used to produce estimates of the Log-normal distribution parameters for each significant wave height bin. To allow extrapolation of the conditional probability distribution for extreme sea-states empirical regression functions are used for the two Log-normal distribution parameters; mean, μ_{LN} and standard deviation, σ_{LN}.

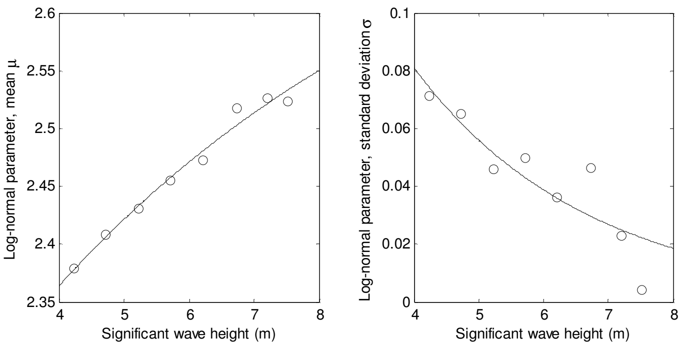

Fig. 5 Regression functions for the Log-normal distribution function

From observation the following two empirical regression functions were identified. These are given by

$$\hat{\mu} = \alpha\left(1 - \beta \exp\left(-\gamma H_s\right)\right) \quad (9a)$$

$$\hat{\sigma} = \delta \exp\left(-\varepsilon H_s\right) \quad (9b)$$

where α, β, γ, δ and ε are the regression coefficients. Fig. 5 shows the fit of the empirical regression functions, which shows that there the functions are a reasonable fit to the data.

Joint Probability Distribution Function

The joint probability density function of significant wave height and mean period is required in order to calculate the return period of the extreme sea states. The joint probability density function is essentially a theoretical reconstruction of the wave resource scatter table. The joint probability density function can be calculated in a variety of ways depending on how the marginal and conditional density functions are constructed. In this case it is simply a multiplication of the marginal probability density function of the significant wave height with the conditional probability density function of the wave period. However, because the marginal and conditional probability distributions calculated refer to the probability of a peak event occurring consideration must be given to large non-peak events as outlined below.

Sea-states with large significant wave heights will also occur during the peak events selected that will influence the calculation of the probability of sub-peak significant wave height. Fig. 6 uses the current data to calculate the relationship between the ratio of the total number of events with a significant wave height to the number of peak significant wave heights and the peak significant wave height. For example, there are approximately 2.0 sea-states with significant wave heights of 7.0 metres for every peak significant wave height of 7.0 metres. To support the analysis of the data an empirical regression function shown on the graph is used to represent the data.

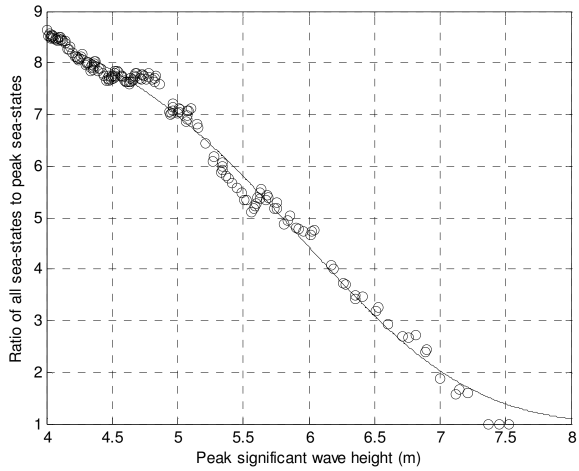

Fig. 6 Ratio of sea-state events to peak significant wave height events

Thus the marginal probability distribution of significant wave height is modified by the ratio of events to peak events before being used to construct the joint probability distribution of the extreme wave climate. Fig. 7 shows the joint probability density function for the wave climate where the probability density contours are shown in 'parts-per-million', e.g. the contour labelled '1' is associated with a sea-state that has a 1 / 1,000,000 probability of occurring. The crosses indicate sea-states from the full data set and help to verify that the joint probability density function is reasonable because the crosses are denser where the value of the probability density function is highest; on the 1 / 10,000 (100) probability contour the average expected number of crosses is approximately 1.75 per grid box.

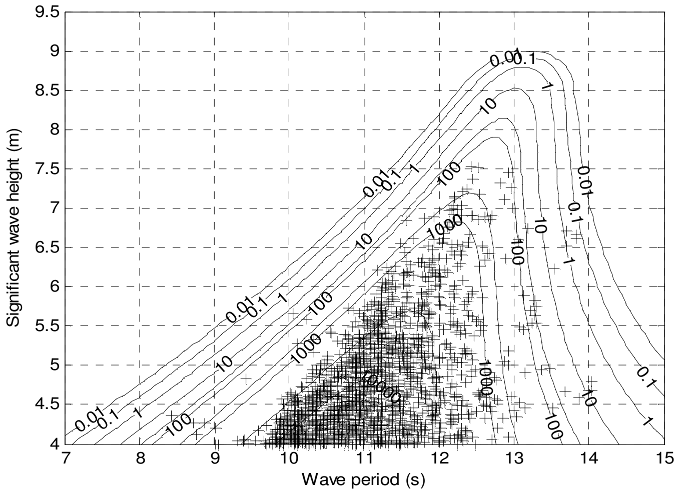

Fig. 7 Joint probability density function contour plot

Return Period of Extreme Sea States

Definition of an extreme sea state in terms of its return period can be calculated in two ways. One approach defines a region on the contour plot shown in Fig. 7 intended to represent the sea state. The joint probability density function is then integrated across this region to determine the probability of the sea state occurring. Alternatively the desired return period can be fixed and the sea state parameters (i.e. H_s and T_m) iteratively calculated using the joint probability density function. The latter approach is favoured here. Because of the narrowness of the joint distribution at the extreme significant wave heights a single representative sea-state is used for all the wave periods. Where appropriate these could be separated into two or more wave periods with a modification of the return period.

Table 1. Definition of extreme sea states

Representative sea-state		Region of Representation	Return period (yrs)
Hs (m)	Tm (s)		
8.8	13	Hs > 8.8 m, all T_m	10,000
8.3	12.8	Hs > 8.1 m, all T_m	50
8	12.5	Hs > 7.7 m, all T_m	10
7.1	12	Hs > 6.7 m, T_m < 12.5 s	1
6.6	13	Hs > 6.2 m, T_m > 12.5 s	1

Table 1 contains the extreme sea states with a particular return period that the Oyster device is expected to encounter at the EMEC site in Orkney. The representative significant wave height is calculated using the first moment of the probability density function *P*

$$\tilde{H}_s = \frac{\int_{H_s'}^{\infty} P(H_s) H_s dH_s}{P(H_s > H_s')} \quad (10)$$

where the tilde indicates the representative significant wave height and the prime indicates the significant wave height for the return period.

EXTREME FOUNDATION LOAD

Extensive experimental testing of the Oyster 2 device was conducted in the wave tank facilities at Queens University, Belfast. In particular, Oyster was tested in the 1, 50 and 10,000 year return period extreme sea states defined in the previous section and all foundation loads experienced by the device recorded. As stated earlier a single event in a wave resource scatter table is the average over a 3 hour period. Thus each sea state tested in the wave tank is run for the equivalent of over 3 hours assuming a Bretschneider wave spectrum. However, because the short term statistics of the foundation loads are strongly dependent on wave-device interactions, it is important to capture as many different wave-device phasing events as possible. Thus all sea states were repeated (with different wave phasing induced each time) so that at least 4000 wave cycles were achieved. This also has the added advantage of increasing the size of the data set which is beneficial for any statistical analysis. However, there is a trade off between quality or robustness of the data set and the time taken to execute the experimental tests.
At this stage of the analysis there is no evidence to suggest that one extreme value method should be preferred over the other, thus both the POT and BM methods are employed. To demonstrate each EVA method the surge foundation load is used as a representative metric for Oysters survivability condition.

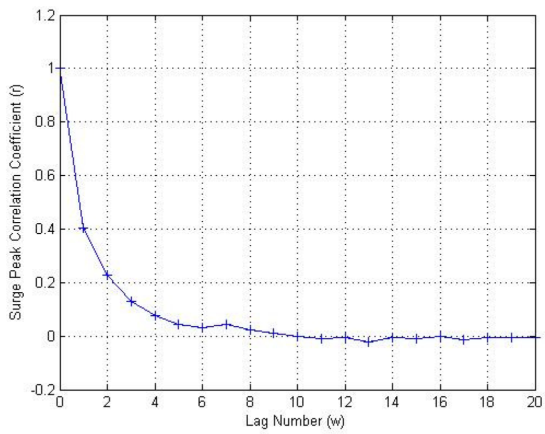

Fig. 8 Correlation coefficient for a set of positive surge load peaks.

The first step to the EVA is to determine an i.i.d. data set. We can ensure that the set is identically distributed by only considering for example positive (forcing in the landward direction) surge events. The underlying physics that excite the device are repeated every wave cycle, therefore it can be inferred that the extreme surge events in a particular direction are governed by the same dynamics. Independence is once again assessed making use of the correlation coefficient given in Eq. 1. The variation of the correlation coefficient with lag number is shown in Fig. 8 for a set of positive surge-load peaks. It is clear in this case that load events that are separated by w=10 peaks or more are completely uncorrelated and so can be considered as independent. It is found however that this independence condition can be relaxed as both the POT and BM methods have a so called “built in” independence filter when applied to the load data. Extreme surge load peaks, i.e. above a threshold or maxima of a data block, of an Oyster 2 device do not regularly occur close together. Thus the peak selection criterion of both methods proves to be sufficient to achieve an independent data set. To highlight this feature the strict independence condition (w=10) is imposed in conjunction with the POT method the correlation coefficient of the resulting data set is -0.005 indicating that it is indeed and independent set. If the condition is then relaxed (w=0) and the POT method again applied the correlation coefficient is still small and is typically less than 0.1. Similarly the correlation coefficient for the BM method is -0.02 and 0.006 respectively. The ability to relax this independence condition has the advantage of not reducing the size of the data set and so the distribution parameters estimated from the data have less variance associated with them.

POT Prediction of Extreme Surge Load

Fig. 9 shows the how the Generalised Pareto distribution parameters and mean of the excess vary with surge threshold levels when tested in the 50 year wave climate condition. These relationships are used to select an appropriate threshold value. For reasons of commercial sensitivity the surge threshold scale has been retrospectively normalised by the predicted 10,000 return period surge load. This however does not limit our ability to interpret the results and draw conclusions.

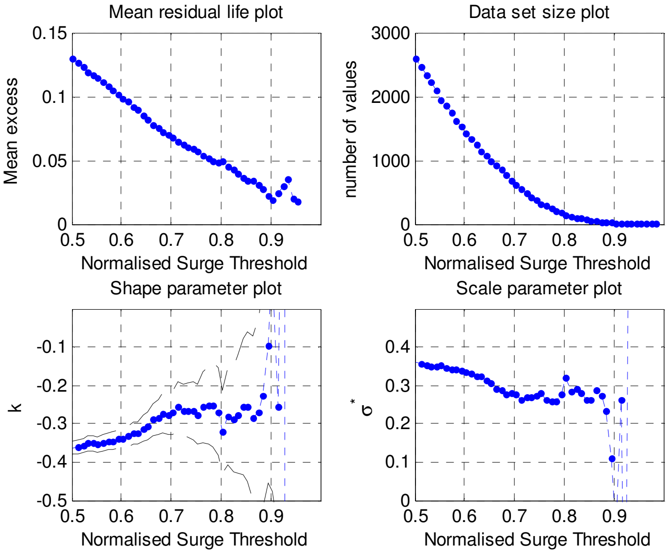

Fig. 9. Generalised Pareto distribution parameters in a 50 year extreme sea state

It can be seen that the mean of the excess varies linearly and the Generalised Pareto distribution parameters are approximately constant above a normalised threshold value of 0.68. This indicates that the Generalised Pareto distribution should be a good approximation to the distribution of the extreme surge load events above this threshold. Erratic behaviour of the distribution parameters is observed above a normalised threshold of 0.85 which is a result of having too small a data set. It is important to note that the shape parameter k (along with the 95% confidence intervals shown as dash lines) remain negative over the threshold range. This indicates that the extreme surge load distribution will have a finite upper bound and shows that the surge load generation mechanism of the Oyster® device is a self limiting process. Fig. 10 shows the POT model diagnostic plots, the Generalised Pareto distribution is fitted to the data assuming a normalised surge threshold value of 0.68. As the independence condition was relaxed in this case there is no need to account for the ratio of all events to peak events. The linear behaviour of the Probability and Quantile plots indicate good agreement between the theoretical Generalised Pareto distribution and the empirical data. The return period graph shows the extrapolation of the theoretical surge load distribution up to a return period of 10,000 years. The dashed lines shown are the 95% confidence

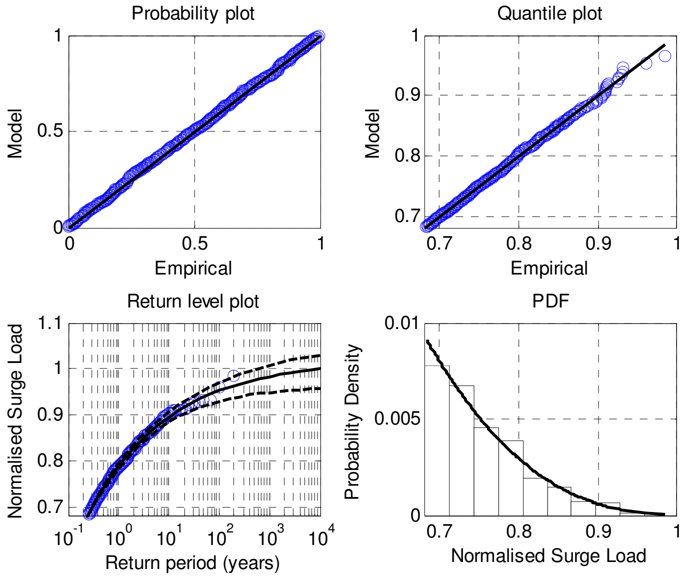

Fig. 10. Model verification of the POT method

intervals again calculated using a resample bootstrap method. The confidence interval at the 10,000 year point is ±3.5% of the most probable surge load indicating that this Generalised Pareto distribution is an excellent approximation of the extreme surge load distribution. The finite limiting behaviour of the distribution is once again highlighted by the return period graph which shows that the 1, 50 and 100 year surge load events are 79%, 94% and 95% of the 10,000 year event respectively.

BM Prediction of Extreme Surge Load

The BM method is applied to the load data in a very similar manner to the POT method just described. In contrast to the POT method where there are guidelines for threshold selection there is no analogous rules for selecting the appropriate block size in the BM method. Nevertheless it is informative to plot the variation of the GEV parameters with block size to perhaps determine a region where the parameters are “invariant” to block size.

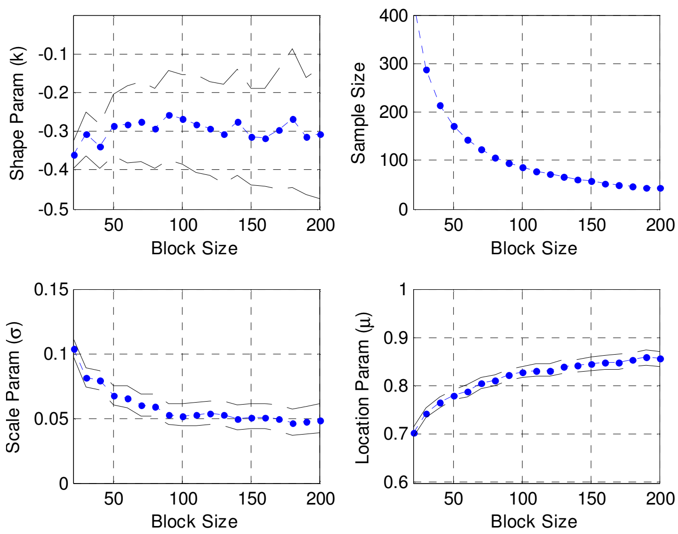

Fig. 11 Generalised Extreme Value distribution parameters in a 50 year extreme sea state

In particular it is the shape parameter k which is of most importance as the scale σ and location parameters μ are often self compensating. Fig. 11 shows the variation of the Generalised Extreme Value distribution parameters with block size. It is observed that irrespective of the block size the shape parameter along with its 95% confidence interval remains negative. This indicates that the block size selection does not distort the domain of attraction and that the physical load generation process has a finite maximum upper bound. Confirming the conclusion reached via the POT method. Fig. 11 also highlights that the BM method only utilises a small amount of the available data and a careful balance between parameter behaviour and data sample size must be achieved. In this case a block size of 110 was selected as an appropriate compromise. Fig. 12 shows the BM model diagnostic plots which indicate that the Generalised Extreme Value distribution with a block size of 110 is a fairly good approximation of the extreme surge load distribution. Again the load axes have been retrospectively normalised by the predicted 10,000 year load event. The variance of the 10,000 year load prediction is ±6% and similar the POT method the finite limiting behaviour of the distribution is highlighted as the 1, 50 and 100 year surge load events are 88%, 94% and 95% of the 10,000 year event respectively.

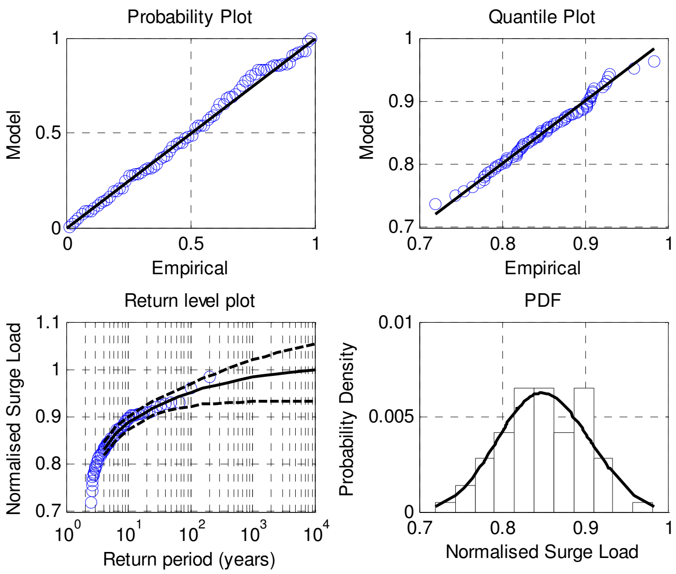

Fig. 12. Model verification of the BM method

Comparison of BM and POT Methods

It is important to compare the results obtained from either method to verify that each technique has been correctly executed. A comparison of the parameters of the Generalised Pareto and Generalised Extreme Value distribution shows that the shape parameter is -0.275 and -0.28 respectively and the scale parameters are within 3% of each other, making use of Eq. 6. In addition to this it is found that the standard deviation of both distributions are almost identical and that there is less than a 0.5% difference in the theoretical finite upper bound surge load calculated using Eq. 3 and Eq. 7 respectively. The excellent correlation of the results from both extreme value methods indicates that either technique could be employed for the analysis of other data sets. It is recommended here however that the Peak-Over-Threshold method is adopted in favour of the Block-Maxima method for the main reasons that it is less wasteful of the available data and so the variance of the predictions is minimised and that there are mathematical guideline which can be used to determine and appropriate threshold value making the implementation of the technique easier.

Load Predictions

Interpretation of the predicted loads in terms of return period can often be misinterpreted. For example the 10,000 year surge load predicted in the previous sections (based on load data in a 50 year sea state) is actually the most probable surge load the Oyster® device would encounter if the 50 year sea state occurred 200 times. There is no guarantee that this is the same load expected if the 10,000 year sea state occurred once. Thus it is necessary to compare the predicted return period loads from the different sea state data sets and assess the variance of the predictions. This was achieved by applying the POT method (with no strict independence condition imposed) to the 1, 50 and 10,000 year sea state data sets. It was found that the maximum difference between the surge loads predicted at the 10,000 year return period is 0.5MN and that the 95% confidence intervals of each set encompass the predicted values for the most probable load of the other data sets. This once again highlights the fact that load generation mechanism of the Oyster® device is governed by a limiting process.

CONCLUSIONS

An Extreme Value Analysis technique has been developed to assess the extreme foundation loads of a Wave Energy Converter. In particular the Oyster® device was examined in a wave climate condition typical of the EMEC test site in Orkney. However the technique could be applied to any other type of Wave Energy Converter. The long term extreme wave statistics were evaluated using a Peak-Over-Threshold method to determine the sea states which Oyster® may encounter at this site. The short term extreme load statistics induced by wave-device interactions were explored using both a Peak-Over-Threshold method and a Block Maxima method. Although the results from both methods were in excellent agreement, the Peak-Over-Threshold method is favoured as it is less wasteful of the available data and its implementation is assisted by some attractive mathematical features.

A significant result of this study is that the extreme foundation load mechanism of the Oyster® device is a self limiting process. This is a result of the depth induced wave breaking of large waves in the near-shore environment, where Oyster® is located along with Oysters ability to decouple from large waves. This phenomenon is identified in the analysis by the fact that the shape parameter of the fitted distributions are negative, indicating that Oysters extreme foundation load has a finite upper bound which can be calculated analytically. This is a key result in determining the survivability design condition of the Oyster® device.

REFERENCES

Burrows, R. And Salih, B.A. (1986). "Statistical modelling of Long-Term Wave Climates". *In Proc. 20th Int. Conf. Coastal Engng*., Taipei, Taiwan, **1**, 42-56.

Cole, S. (2001). "*An Introduction to Statistical Modelling of Extreme Values.*" London, Springer-Verlag

Mathiesen, J. And Bitner-Gregersen, E. (1990). "Joint Distribution for Significant Wave Height and Zero-Crossing Period". *Appl. Ocean Res.*, **12**, 2, 93-103.

Morton, I.D. and Bowers, J. (1996). "Extreme Value Analysis in a Multivariate Offshore Enviroment". *Applied Ocean Research*, **18** (6), 303-317.

Thompson, P. Cai, Y. Reeve, D. And Stander, J. (2009). "Automated Threshold Selection Method for Extreme Wave Analysis." *Coastal Engineering* **56** (10), 1013-1021.