Reducing the uncertainty of ULS load estimates in offshore structural design

João Cruz, Mairéad Atcheson Cruz, Michele Martini

Abstract—**This paper presents a novel post-processing methodology to assess the uncertainty when estimating long-term extreme loads acting on an offshore structure, which can be associated with the capacity of a structure to withstand Ultimate Limit State (ULS) conditions. To assess which statistical distributions are best suited to estimate long-term extreme loads, goodness-of-fit tests were performed using a series of input load time-series (Kolmogorov – Smirnov, Kuiper, Cramer – Von Mises, Anderson – Darling). As an example, the methodology was initially applied to the evaluation of 50-year return load estimates acting on the foundation of a generic Submerged Pressure Differential (SPD) Wave Energy Converter (WEC). Results suggest that current conventional practices based on visual inspection may lead to the selection of non-representative fitting functions, which in turn may lead to inaccurate extreme load estimations.**

Ultimately, the methodology described in this paper aims to contribute to a probabilistic approach to the definition of suitable safety factors, which is expected to reduce the uncertainty in key design metrics and consequently mitigate the risk for either under- or over-designing an offshore structure.

Keywords—**concept design, extreme load, goodness-of-fit tests, load analysis, uncertainty quantification, Ultimate Limit State (ULS), Wave Energy Converter (WEC).**

I. Introduction

THE design process of any offshore structure, including Wave Energy Converters (WECs), aims to assess a series of critical conditions, often encapsulated in the concept of 'limit states', that define the 'design borders' across which a structure is unlikely to respond satisfactorily, potentially leading to functional failure(s). Limit state design, also referred to as the Load and Resistance Factor Design (LRFD) method, aims to quantify such 'design borders' with adequate safety margins on both the loading and the resistance sides of the design inequality. An inherent risk of the LRFD method is that the definition of 'adequate' may be somewhat clouded by multiple sources of uncertainty, making it challenging to assess if a system is likely to be under- or over-designed from the inception of the design process.

At a high-level, structural failure is likely to occur when the stress at a specific structural location exceeds the local material resistance capacity. Noting that stress can be associated with a loading regime, the strength of a structural component can be related to the Ultimate Limit State (ULS) loading conditions that lead to structural failure. For systems such as WECs, which must withstand combinations of environmental conditions and / or machine states, the ULS load at a specific structural location may be defined as the highest of the extreme loads associated with each of these combinations. Additionally, and as WECs are likely to be subjected to stochastic loading environments, extreme loads are typically evaluated by adopting reliability-based approaches. For example, the guidance issued in [1] requires a WEC to achieve a Safety Level (SL) of 2, which corresponds to being designed to withstand a load that may statistically occur once every 50 years.

Although overall guidance for the design process of a WEC can be sought from related industries such as offshore wind – see e.g. [2] to [6], and while noting that generic guidance for WEC design is also available in a shortlist of references – see e.g. [1], there is at present limited practical guidance on how to derive estimates of long-term return period loads. In particular, and while the use of extreme value distributions is suggested in multiple standards, only generic guidance is issued for the calculation of extreme loads, i.e. no specific, quantifiable method on how to select a specific distribution is discussed. For WEC design, recent work has highlighted the importance of addressing such gap in the design process: for example, a comparative assessment of different load estimation methods is documented in [7], illustrating the fundamental dependence of the output load estimates on the underlying method followed to derive them; additionally, and in connection with the assessment of extreme loads acting on WECs, the load

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreements No 101006927 and No 10100707.

J. Cruz, M. Atcheson Cruz and M. Martini are with Yavin Four Consultants Lda., Rua do Montepio Geral, N.4-B, 1500-465 Lisboa, Portugal (e-mail: joao.cruz@yavinfourconsultants.com).

Digital Object Identifier: https://doi.org/10.36688/ewtec-2023-147

post-processing methodology was identified in [8] as a major contributor to the uncertainty in ULS load estimates.

This paper aims to present a novel methodology that extends existing best practices by providing a practical method to assess the suitability of specific extreme value distributions when estimating extreme loads, in the context of the design of offshore structures. The novel methodology uses multiple goodness-of-fit tests to derive quantifiable metrics that support the assessment of the 'accuracy' of extreme load estimates related to different extreme value distributions, eliminating the need to rely (solely) on methods based on visual inspection. The authors acknowledge the preliminary nature of the findings – noting further verification and validation are required to ensure that it can be adopted with confidence.

Owing to the importance of the post-processing methodology in the overall uncertainty associated with ULS load estimates [8], the methodology presented in this paper solely addresses the influence of different extreme value distributions in the resulting load estimate; future work may address additional uncertainty sources, at e.g. pre-processing and processing levels.

II. METHODOLOGY

A. Extreme Load Analysis (ULS)

There are multiple methods to estimate an extreme load in offshore structural design. These may include approaches based on e.g. short-term analysis, long-term analysis, contour-based methods, full-environmental characterisation, etc. In this paper, a short-term extreme response analysis method was followed, as illustrated in Fig. 1. To apply the method, load time-series $L(t)$ were derived via a time-domain numerical model, using an Extreme Sea State (ESS) with 1-year return period (T_R) as the main environmental input. The load peaks were extracted from $L(t)$ and sorted in ascending order, to populate the dataset $\boldsymbol{x}$ with the values x_i:

$$\boldsymbol{x} = \{x_1, x_2, \dots, x_N\}, \qquad i = 1,2, \dots, N \tag{1}$$

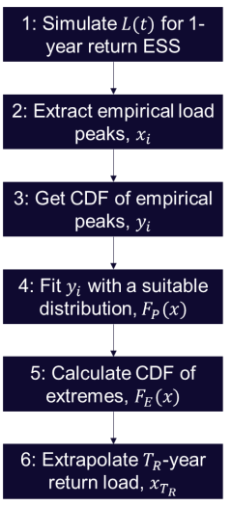

Fig. 1. High-level breakdown of the original methodology applied to derive extreme load estimates.

In (1), N is the total number of load peaks. An empirical Cumulative Distribution Function (CDF), $\boldsymbol{y}$, was then calculated and its corresponding estimates, $y_i = y(x_i) = \frac{i}{N}$, were fitted with a statistical distribution, $F_P(x)$, which represents the probability of one load peak being below a value of x for the considered ESS. The probability that all load peaks are below a value x during the same sea state can be estimated as:

$$F_E(x) = F_P(x)^N \tag{2}$$

In (2), and acknowledging that the selected input ESS has a 1-year return period, $F_E(x)$ can also be referred to as the 'annual maximum distribution'. Finally, the extreme load associated to a return period T_R, x_{T_R}, which physically represents a load threshold which has the probability of being exceeded once every T_R years, can be estimated as:

$$x_{T_R} = x \left| \left\{ F_E(x) = 1 - \frac{1}{T_R} \right\} \right. \tag{3}$$

If e.g. $T_R = 50$ years, and following (3), the 50-year return load estimate, x_{50}, is given by:

$$x_{T_R} = x | \{F_E(x) = 0.98\} \tag{4}$$

B. Selection of a 'Suitable' Distribution

As illustrated in Fig. 1, the selection of a 'suitable' statistical distribution, $F_P(x)$, is a critical step in the proposed methodology. The criticality of such step is discussed in e.g. [8], where the ULS load post-processing methodology was identified as the largest contributor to the load estimate uncertainty following a Probabilistic Variation Mode and Effects Analysis (P-VMEA) approach, accounting for over 84% of the total distribution of variance associated with all uncertainty sources.

Despite the recognition of its importance in the overall uncertainty, existing standards provide only generic guidance on how to select a 'suitable' distribution. For example:

- DNV-ST-0126 [2] and DNV-ST-0119 [3] refer to the consideration of "*all (…) states that contribute to the upper-tail of the distribution of the annual maximum load*", i.e. $F_E(x)$ in (2), without specifying how such distribution may be selected.
- IEC 61400-1 [4] states that "*the distribution type selected should be checked to see if the fit to data is acceptable*", without suggesting a methodology to define / quantify 'acceptability'.
- DNV-RP-C205 [5] suggests "*to use extreme estimates based on fitting of the tail of the peak distribution*", but no details on how to perform fitting are provided.
- DNV-ST-0437 [6] makes reference to "*load extrapolation methods*", without providing any further guidance.

In summary, and although relevant, available standards do not tend to specify which distribution / fitting function should be used when performing extreme load analysis, nor do these suggest quantifiable metrics to assess if a distribution is able to accurately represent the empirical data.

A subjective evaluation approach based on visual inspection is often the result of such lack of practical guidance. To overcome such shortcomings, this paper investigates the use of goodness-of-fit tests to reduce the uncertainty in extreme load estimates. Specifically, goodness-of-fit tests are proposed as a means to assist in the selection of a 'suitable' distribution to fit the empirical load peaks, i.e. Step 4 in Fig.1 and Fig. 2. Although previous work, e.g. [7] and [9], investigated the suitability of different distributions by comparing extreme load estimates to those calculated via a block-maxima method, to the authors' knowledge the present approach is the first to propose the evaluation of the accuracy of the fit via quantifiable metrics, eliminating, even if partially, a purely subjective evaluation criterion when defining and selecting a 'suitable' distribution.

C. *Goodness-of-Fit Tests*

Goodness-of-fit tests can be used to test the (null) hypothesis, referred to as H_0, that a data sample x can be associated with a specific distribution $F_P(x)$. For functions that are continuous and fully specified, Empirical Distribution Function (EDF) tests are a class of goodness-of-fit statistics that may be considered – see e.g. [10].

Four EDF tests were assessed in the present study. These are briefly described below, and include:

- Kolmogorov-Smirnov.
- Kuiper.
- Cramer – Von Mises.
- Anderson-Darling.

The Kolmogorov-Smirnov test is based on the test statistic D:

$$D = \max(D^+, D^-), \text{ with:} \tag{5}$$

$$\begin{cases} D^+ = \sup_x \{y - F_P(x)\} \\ D^- = \sup_x \{F_P(x) - y\} \end{cases} \tag{6}$$

The metric D therefore measures the maximum absolute distance between the empirical and the fit distributions.

The Kuiper test is a modification of the Kolmogorov-Smirnov test, and is based on the test statistic V:

$$V = D^+ + D^- \tag{7}$$

The metric V therefore accounts for the largest difference between the empirical data and the fit distribution, both when y is larger than $F_P(x)$ and vice-versa.

TABLE I
MAIN FEATURES OF THE SPD WEC – ADAPTED FROM [12]

Parameter	Unit	Value
Float Mass	ton	35
Float Diameter	m	7
Float Height	m	5
Nominal Float Submersion	m	1.5

The Cramer – Von Mises test statistic, W^2, can be defined as:

$$W^2 = \int_{-\infty}^{+\infty} [F_P(x) - y]^2 dF_P(x) \tag{8}$$

The test statistic W^2 differs from D and V as it is based on an aggregate measure (i.e. integral) of a quadratic, rather than linear, distance between the empirical and the fit distributions.

Finally, the Anderson – Darling test statistic, A^2, can be defined as:

$$A^2 = N \int_{-\infty}^{+\infty} \frac{[F_P(x) - y]^2}{F_P(x)\,[1 - F_P(x)]} dF_P(x) \tag{9}$$

Similarly to W^2, the test statistic A^2 is based on an aggregate measure of a quadratic distance between the empirical data and the fitting function. Additionally, it gives more importance to the tails of the distribution, i.e. when the value of $F_P(x)$ is close to either 0 or 1, the distance measurement is amplified.

For practical applications, and following [10], (6), (8) and (9) can be rearranged via a Probability Integral Transformation (PIT) to obtain suitable computing formulas, as defined by (10), (11) and (12), respectively:

$$\begin{cases} D^+ = \max\limits_{1 \le i \le N} \left[\frac{i}{N} - F_P(x_i)\right] \\ D^- = \max\limits_{1 \le i \le N} \left[F_P(x_i) - \frac{i-1}{N}\right] \end{cases} \tag{10}$$

$$W^2 = \frac{1}{12N} + \sum_{i=1}^{N} \left[\frac{2i-1}{2N} - F_P(x_i)\right]^2 \tag{11}$$

$$A^2 = -N - \sum_{i=1}^{N} \frac{(2i-1)}{N} [\ln F_P(x_i) + \ln(1 - F_P(x_{N+1-i}))] \tag{12}$$

In this study, goodness-of-fit-tests were incorporated into the methodology followed to derive extreme load estimates.

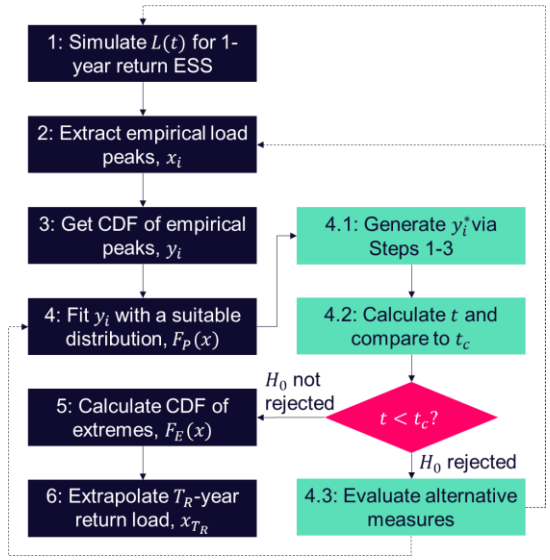

Fig. 2. Revised methodology applied to derive extreme load estimates, including goodness-of-fit tests.

The combined process is illustrated in Fig. 2, where it is proposed to evaluate the suitability of a specific distribution $F_P(x)$ to fit the empirical peaks via goodness-of-fit tests in Steps 4.1 to 4.3.

As one of the hypotheses of goodness-of-fit testing is that distributions have to be fully specified, the evaluation of the test statistics detailed in (5) to (9) was based on a second empirical peak distribution, $y_i^* = y(x_i^*)$ (Step 4.1 in Fig. 2). The load peaks x_i^*, which are statistically independent of the x_i series, were derived via additional simulations, where the same input sea state was considered with a different seed.

The test statistic t was then calculated and compared to a critical value, t_c, for a chosen level of significance α (Step 4.2 in Fig. 2). If $t < t_c$, the test fails to reject the hypothesis H_0 and the extreme load may be estimated by following Steps 5 and 6 in Fig. 2; otherwise, additional measures shall be taken by e.g. selecting another distribution $F_p(x)$, extracting an alternative subset of peaks, performing additional simulations to collect a higher number of peaks, etc. (Step 4.3 in Fig. 2).

As an illustrative example, the methodology summarised in Fig. 2 was first applied to a dataset of load peaks generated from a reference model of a WEC. The model and the corresponding input settings are detailed in III.

III. Reference Model Setup and Main Input Data

A SPD WEC type device was considered in this study as a reference offshore structure. The design was inspired by the Bref-SHB device of the NumWEC project [11] - [12], and consists of a positively buoyant floater, submerged below the free-surface, tethered to the seabed (rigid link).

In static conditions, the floater buoyancy is reacted by a pre-tensioning system. Under the action of waves, the floater moves in the water column and its motion provides input to a linear Power Take-Off (PTO) system – see Fig. 3. The main geometric features of the SPD WEC are summarised in Table I.

The SPD WEC was modelled in WEC-Sim (Wave Energy Converter SIMulator), an open-source WEC simulation tool developed in MATLAB/Simulink using the multi-body dynamics solver SimMechanics. A schematic of the SPD WEC-Sim model is presented in Fig. 4. Hydrodynamic coefficients were calculated in Nemoh, a first-order Boundary Element Method (BEM) solver. Additionally, a Morison model was added to the formulation, based on the drag coefficients suggested in [11]. The simulations included the evaluation of nonlinear hydrostatic and nonlinear Froude-Krylov forces.

The environmental inputs were limited to wave climate data, extracted from a 30-year hindcast compiled using the third-generation spectral wave model, WaveWatch III (WWIII). For illustrative purposes, the SPD WEC was assumed to be deployed off the coast of France (3.75W, 48.83N), in a moderate wave energy climate around the 20m water depth bathymetric contour.

The response of the SPD WEC was simulated in a parked design situation, as defined in [1]. In this state, the PTO was assumed to be locked via the action of e.g. a mechanical brake, and the buoy lowered by 3m in the water column, bringing the submersion level to 4.5m.

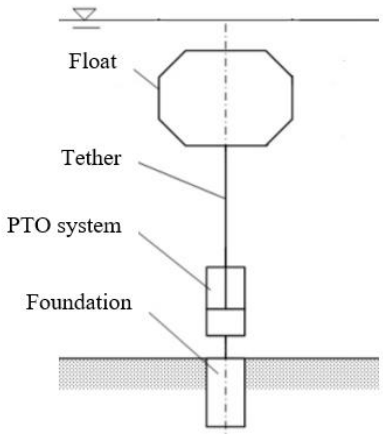

Fig. 3. Schematic of the SPD WEC, adapted from [12].

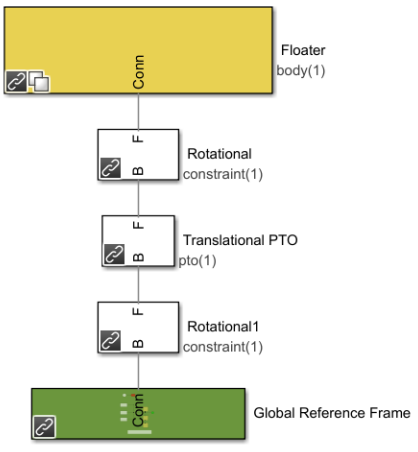

Fig. 4. Schematic of the WEC-Sim model of the SPD WEC.

TABLE II
SELECTED ESS CONDITIONS (1-YEAR RETURN PERIOD)

Parameter	Unit	Value
H_s	m	7.67
T_p	s	15.96
Spectral Shape	-	Pierson-Moskowitz

The 1-year return period ESS conditions used in the simulations are summarised in Table II in terms of significant wave height (H_s), peak wave period (T_p) and spectral shape. To define the ESS conditions from the 30-year dataset, a 1-year return contour was derived following Principal Component Analysis (PCA) and the Inverse First-Order Reliability Method (I-FORM), as detailed in [13], using a modified version of the Extreme Sea State Contour module of the WEC Design Response Toolbox (WDRT).

The sea state defined in Table II corresponds to that with the highest H_s on the estimated 1-year return contour.

IV. RESULTS

A. Sample Load Time-Series

For the model setup summarised in III, and as illustrated in Step 1 of Fig. 2, load time-series were obtained in WEC-Sim for specific locations across the WEC structure, namely for key bodies and structural joints. Following [1], 3-hour sea state representations were used as the main environmental input.

As an illustrative example, the horizontal load at the foundation was analysed, i.e. the horizontal reaction force at 'constraint(1)' in Fig. 4. An example of the resulting load time-series and of the corresponding peaks is given in Fig. 5. Both positive and negative load peaks were selected for the initial analysis.

B. Extreme Value Analysis

Having identified all the peaks from a given load time-series, and following the guidance from e.g. [4], a threshold, x_T, was set as a cut-off to exclude lower amplitude events. Such approach is based on the assumption that the physical processes that lead to long-term extremes are related to those that produce maxima on the tail of the (original) short-term distribution – see e.g. [14]. The approach is also aligned with applicable recommended practices such as [5], which suggest focusing on the upper-tail of the peak distribution. Following [9], a threshold of 1.4 times the standard deviation above the mean of the force time-series was used, leading to $x_T = 294$ kN. The influence of other threshold values may be investigated in future work.

The CDF of the load peaks was then estimated, following (3) - Step 3 in Fig. 2. The values y_i were fitted with a range of extreme value distributions, using the WDRT and the ScyPy Python library (version 1.7.3). The following distributions were considered:

- Weibull (WEI).
- Weibull Tail-Fit (WTF) [7].
- Gumbel (GUM).
- Generalised Pareto (GP).
- Generalised Extreme Value (GEV).

Fig. 6 exemplifies the resulting distribution fits, with the y_i data represented by red crosses. To facilitate the assessment of the upper-tail fits, the y-axis is presented in logarithmic scale and the Complementary CDF (CCDF) values are reported, i.e. the probability of exceedance. Based solely on visual inspection, and following the current state-of-the-art approach illustrated in Fig. 1, the WEI, WTF and GP distributions would appear as good candidates, while the GUM and GEV distributions would be considered the least likely to be suitable.

The impact of the different distribution fits is clearer when assessing the load estimates associated with the long-term return period, following (2) – see also Steps 5 and 6 of Fig.2. For representative purposes, these steps were followed for the five tested distributions, and Fig. 7

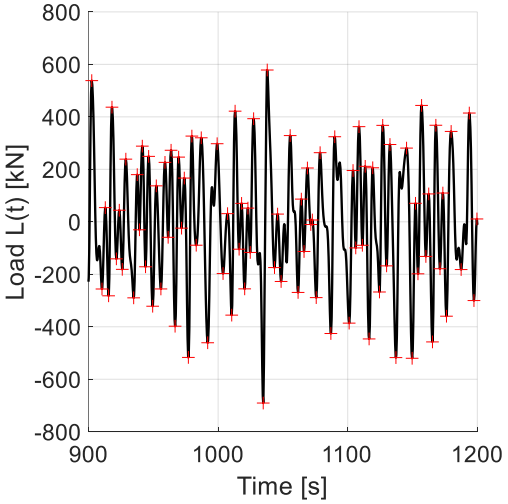

Fig. 5. Extract of a sample load time-series (black line) and peaks (red crosses): horizontal load at the foundation.

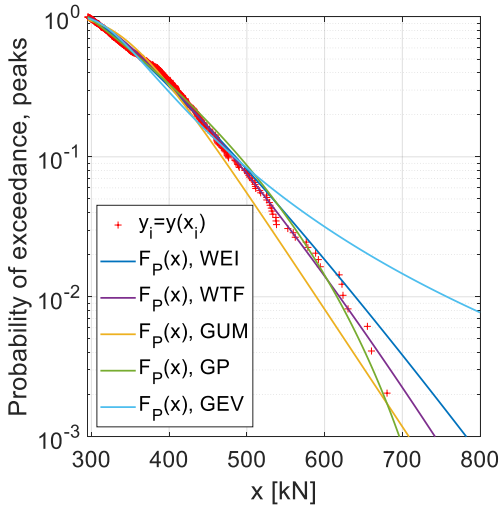

Fig. 6. Empirical CCDF of peaks for horizontal load at the foundation, y_i, and associated fitting distributions, $F_P(x_i)$.

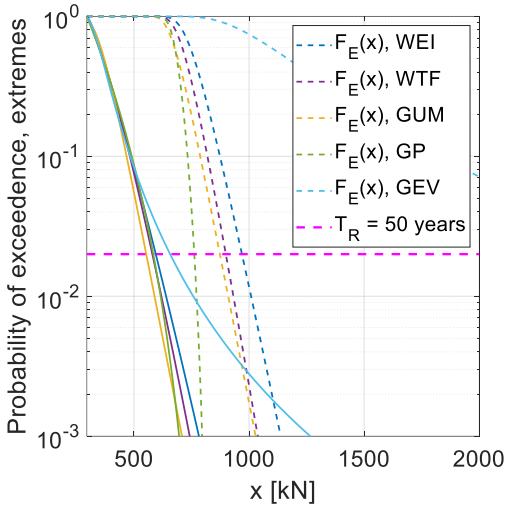

Fig. 7. CCDF of extremes for horizontal load at the foundation, $F_E(x)$. The associated peak distributions, $F_P(x)$, are presented as continuous lines (same colour coding).

TABLE III
ESTIMATES OF THE 50-YEAR HORIZONTAL LOAD AT THE FOUNDATION

x_{50} (kN)				
WEI	WTF	GUM	GP	GEV
969	899	873	762	2,779

illustrates the extreme distributions, $F_E(x)$ (dashed lines), alongside the peak distributions, $F_P(x)$ (continuous lines); the horizontal magenta line represents the 50-year return probability value.

As expectable, it can be seen in Fig. 7 that extreme load estimates may significantly exceed (one-off, isolated) peak load estimates. Also, Fig. 7 illustrates that differences in peak distributions that appear to be minor (visually) may result in large differences between extreme distributions. Ultimately, such differences can be quantified via the evaluation of the 50-year return load estimates, x_{50}, which are reported in Table III.

As visible in Table III, the five tested distributions provide a wide range of estimates of the 50-year return load. Significant variations are present even for the WEI, WTF and GP related estimates, which emerged as 'good candidates' when following the visual inspection approach. Such variations emphasise the need of additional checks to provide robustness in the extreme load calculation process, to mitigate the resulting uncertainty and, where possible, avoid unnecessary under- or overdesign.

C. *Goodness-of-Fit Tests*

In order to perform the goodness-of-fit tests, a supplementary peaks dataset y_i^* was generated via additional simulation(s), with a different realisation (seed) for the same input sea state (as per Step 4.1 in Fig. 2). An example of the resulting estimates of y_i^* is illustrated in Fig. 8, where comparisons with the original load peaks dataset y_i can be made. For clarity, the values y_i (red crosses in Fig. 8) were used to define the parameters for the fitting distributions, while the values y_i^* (black crosses in Fig. 8) were used solely for goodness-of-fit testing (Step 4.2 in Fig. 2).

TABLE IV
CRITICAL VALUES FOR DIFFERENT TEST STATISTICS, AT 5% SIGNIFICANCE LEVEL [9]

D	V	W^2	A^2
1.358	1.747	0.461	2.492

TABLE V
ESTIMATED VALUES OF THE TEST STATISTICS ASSOCIATED WITH $F_p(x)$

Distribution	D	V	W^2	A^2
WEI	0.822	1.476	0.136	1.480
WTF	1.323	1.802	0.510	4.439
GUM	1.951	3.103	0.983	6.220
GP	0.509	0.836	0.034	0.396
GEV	1.337	2.572	0.487	3.580

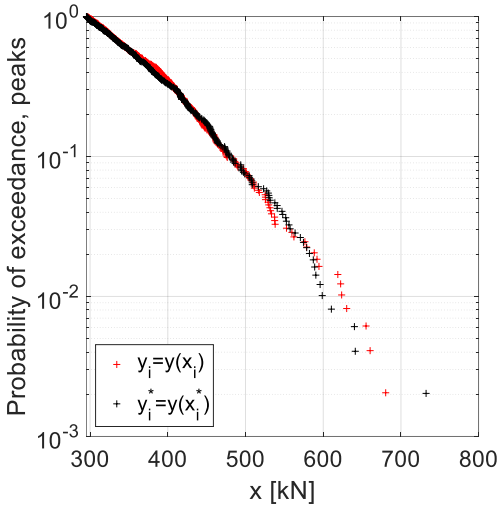

Fig. 8. Empirical CCDFs of peaks for horizontal load at the foundation, y_i and y_i^*.

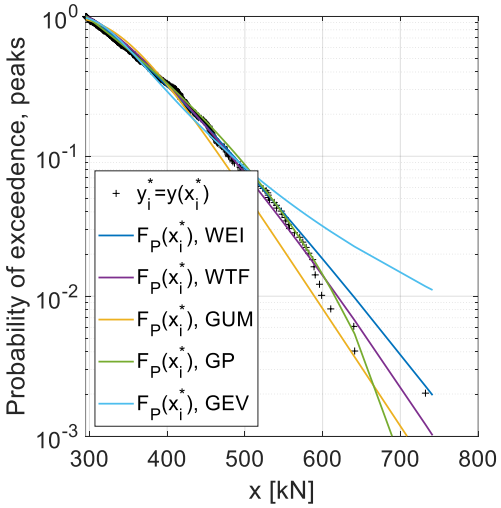

Fig. 9. Empirical CCDF of peaks for horizontal load at the foundation, y_i^*, and associated fitting distributions, $F_P(x_i)$.

V. DISCUSSION

The goodness-of-fit analysis presented in IV. C suggests that the hypothesis that the empirical peaks follow a GUM distribution is rejected by all tests, as the associated test statistics are consistently above the respective critical values. Such finding is aligned with what a visual inspection could have concluded, considering that the GUM function provides a poor visual fit (see e.g. Fig. 9). Similarly to what is observed for the GUM distribution, a visual inspection would have likely excluded the use of a GEV distribution. The hypothesis that the peaks follow a GEV distribution is rejected by all tests except the Kolmogorov – Smirnov test, noting however that the calculated test statistic is very close to the critical value.

Importantly, all tests except the Kolmogorov – Smirnov test reject the hypothesis that the peaks are well represented by a WTF distribution. This is somewhat in disagreement with what a visual inspection could have indicated, when observing e.g. Fig 6 and Fig. 9. The rejection of the WTF distribution is possibly related to the performance at the lower-tail and / or the central part of the distribution; this hypothesis can be investigated, for example for the Cramer – Von Mises test, by introducing a variable $\omega(x_i^*)$ which can be defined as:

$$\omega(x_i^*) = \left(F_P(x_i^*) - \frac{2i-1}{2n}\right)^2 \tag{13}$$

Fig. 10 illustrates estimates of $\omega(x_i^*)$ for all the tested distributions. It is clear from Fig. 10 that for the WTF, the highest values of $\omega(x_i^*)$ are associated with the lower tail of the distribution, corroborating the hypothesis that the WTF distribution was rejected mainly due to the poor agreement in the lower tail (for the Cramer – Von Mises test; for the Anderson – Darling test, emphasis in the quality of the fit at both tails is given). Similar conclusions apply to the GUM and GEV distributions.

Following the analysis of the results shown in Table V, the WEI and GP distributions are the only ones for which all the test statistics estimates are below the respective critical values. Such finding is consistent with what a visual inspection could have concluded, i.e. both distributions provide a good visual fit to the empirical observations. When considering the WEI and the GP distributions, and building on the results of the goodness-of-fit tests, the GP distribution may be considered as the most suitable, noting it can be consistently associated with the lowest test statistic estimates (and therefore a higher probability that the tested hypothesis is 'true').

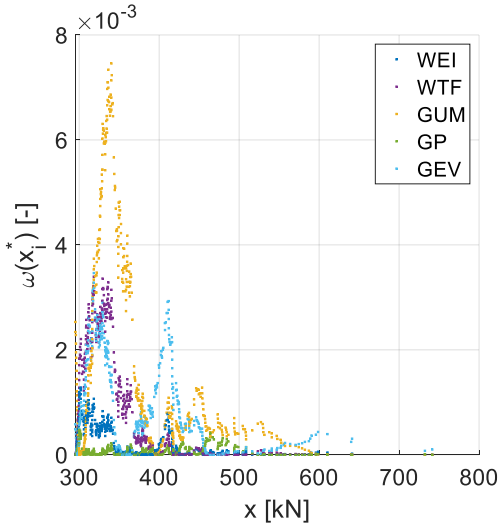

Fig. 10. Estimates of ω_i^* for the tested distributions.

Consequently, the most reliable / 'accurate' estimate of the 50-year return load is likely to be x_{50} =762kN. The WTF distribution, which provided a good visual fit but was rejected by goodness-of-fit testing, would have returned an extreme load estimate of 899kN - 18% higher than that calculated via a GP distribution. The WEI distribution, which passed the goodness-of-tests yet provided test statistics closer to the critical value(s) than those of the GP distribution, would have resulted in an extreme load estimate of 969kN - 27% higher than that calculated via a GP distribution.

Overall, this study suggests that, as a minimum, a visual inspection approach should be accompanied by goodness-of-fit testing to provide additional robustness to the extreme load analysis process. In addition, goodness-of-fit testing may be used to downselect from a group of preliminary suitable distributions, when two (or more) appear to be good candidates based on visual inspection only. Finally, and based on the definitions introduced in II.C, it is noted that the Kolmogorov – Smirnov test (via the statistic D) and the Kuiper test (via the statistic V) are less likely to be the most appropriate tests, as such tests rely on a point-measure which may be biased by e.g. influential observers. For example, the value of D for the WTF distribution is similar to that for the GEV distribution (see Table V), although the GEV distribution appears to perform significantly worse than the WTF distribution, based on both visual inspection and the other test statistics. The Cramer – Von Mises and Anderson – Darling tests are likely to provide a higher degree of robustness, being based on a quadratic distance across the entire distribution. The Anderson – Darling test has the potential to emphasise fit quality at the upper-tail; however, this is somewhat balanced by equal emphasis being given to the lower tail.

VI. CONCLUSIONS

A. *Summary of the Key Findings*

This paper presented a novel methodology, based on goodness-of-fit statistics, to reduce uncertainty in the evaluation of extreme loads acting on an offshore structure. Four goodness-of-fit tests were introduced: Kolmogorov – Smirnov, Kuiper, Cramer – Von Mises and Anderson – Darling. As an example, the tests were applied to evaluate the capability of Weibull, Weibull Tail-Fit, Generalised Extreme Value, Gumbel and Generalised Pareto distributions to estimate the extreme horizontal

load acting on the foundation of a reference WEC, in a parked design situation.

The main finding of this paper suggests that selecting a load peak distribution solely via visual inspection may lead to selecting a distribution which is not statistically representative of the empirical data, which in turn is likely to lead to unrealistic estimates of extreme loads. For the sample case analysed in this paper, selecting the extreme distribution for the horizontal load at the WEC foundation solely via visual inspection may have led to overestimating the 50-year return load by up to 27%. Therefore, goodness-of-fit tests are suggested as (at least) an add-on to visual inspection, to ensure that the load analysis process is more robust. The preliminary results indicate that the Cramer – Von Mises and Anderson – Darling tests may be more suitable than the Kolmogorov – Smirnov and Kuiper tests, as these are based on a quadratic distance across the entire range of data, rather than on single-point observations. Finally, and although numerically derived load time-series were used in this study, it is noted that the methodology is also applicable to physically derived load time-series.

B. Potential Future Work

Future work may focus on increasing the robustness of the presented methodology. This may include e.g. adapting statistical tests to the upper-tail of a distribution only; evaluating how results are affected by the number of available empirical load peaks, which can be related to the number of simulations performed and / or their duration, and by assessing the influence of different threshold value x_T, potentially setting it based on other measures of dispersion (e.g. Interquartile Range, IQR). Application of the methodology to other load sources and other design situations, simulated via either numerical or physical models, for multiple types of offshore structures, is also suggested as a next step – to stimulate the adoption of the methodology in a transition to a (more) probabilistic orientated design process.

Future work may also address the influence of the type of extreme value analysis method that is applied. The authors acknowledge that other methods may be followed when estimating extreme loads. For example, and from [7], the block-maxima method may provide extreme load estimates without fitting the load peaks with a specific distribution; when following a contour-based approach, pre-processing uncertainty is also introduced given the dependence on the method followed to generate the contours – see e.g. [15]; the contour followed prior to the extrapolation for the long-term period estimation also introduces a source of uncertainty, as e.g. nonlinearities may affect the structural response differently at different temporal and physical scales; and methods based on full environmental characterisation, which involve a large number of input sea states, aim to mitigate the dependence of the load estimate on a the data points along a contour. However, and when compared to the methodology described in this paper, some of the alternative methods often require an increased computational effort, which may not be compatible with earlier design stages, namely in the transition from concept to detailed design.

ACKNOWLEDGEMENT

The methodology proposed in this paper was developed within the VALID project, which has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101006927. The reference data used in this paper was generated within the IMPACT project, which has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101007071.

REFERENCES

[1] Marine energy - Wave, tidal and other water current converters - Part 2: Marine energy systems - Design requirements, IEC TS 62600-2:2019.
[2] Support structures for wind turbines, DNV-ST-0126, 2021.
[3] Floating wind turbine structures, DNV-ST-0119, 2021.
[4] Wind turbines – Part 1: Design requirements, IEC 61400-1:2005.
[5] Environmental conditions and environmental loads. DNV-RP-C205, 2021.
[6] Loads and site conditions for wind turbines, DNV-ST-0437, 2021.
[7] C. Michelen, R. Coe, Comparison of Methods for Estimating Short-Term Extreme Response of Wave Energy Converters, Sandia National Laboratories, SAND2015-6890C. [Online]. Available: https://www.osti.gov/servlets/purl/1306093
[8] M. Atcheson, J. Cruz, T. Martins, P. Johannesson and T. Svensson, T, "Quantification of load uncertainties in the design process of a WEC", in *Proc. 13th European Wave and Tidal Energy Conference* (*EWTEC 2019*), Naples, Italy.
[9] M. Stephens. "EDF Statistics for Goodness of Fit and Some Comparisons," *Journal of the American Statistical Association*, vol. 69, No. 347, p. 730-737, 1974.
[10] R. D'Agostino, M. Stephens, Goodness-of-fit techniques. New York: Marcel Dekker, 1986.
[11] A. Babarit, J. Hals, M. J. Muliawan, A. Kurniawan, T. Moan, and J. Krokstad, "Numerical benchmarking study of a selection of wave energy converters", *Renew. Energy*, vol. 41, pp. 44–63, 2012. DOI: 10.1016/j.renene.2011.10.002 [Online].
[12] A. Babarit, J. Hals, A. Kurniawan, T. Moan, and J. Krokstad, The NumWEC project: Numerical estimation of energy delivery from a selection of wave energy converters. Final Report. Ecole Centrale Nantes, Norwegian University of Science and Technology and Statkraft. [Online]. Available: https://www.researchgate.net/publication/275645326_The_NumWEC_project_Numerical_estimation_of_energy_delivery_from_a_selection_of_wave_energy_converters_-_final_report.
[13] A. Eckert-Gallup, C. Sallaberry, A. Dallman, V. Neary, "Application of principal component analysis (PCA) and improved joint probability distributions to the inverse first-order reliability method (I-FORM) for predicting extreme sea states", *Ocean Engineering*, vol. 112, pp. 307-319, 2016. DOI: 10.1016/j.oceaneng.2015.12.018 [Online].
[14] P. Moriarty, W. Holley, S. Butterfield, "Effect of Turbulence Variation on Extreme Loads Prediction for Wind Turbines", *ASME. J. Sol. Energy Eng*. Vol. 124(4), pp. 387–395, 2002. DOI https://doi.org/10.1115/1.1510137 [Online].
[15] A. Haselsteiner, R. Coe, L. Manuel, W. Chai, B. Leira, et al., "A benchmarking exercise for environmental contours", *Ocean Engineering*, vol. 236, 109504, 2021. DOI: 10.1016/j.oceaneng.2021.109504 [Online].