Required test durations for converged short-term wave and impact extreme value statistics–Part 2

Deck box dataset

Scharnke, Jule; van Essen, Sanne M.; Seyffert, Harleigh C.

DOI
10.1016/j.marstruc.2023.103411

Publication date
2023

Document Version
Final published version

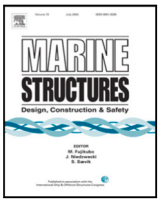
Published in
Marine Structures

Marine Structures 90 (2023) 103411

Contents lists available at ScienceDirect

Marine Structures

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



Required test durations for converged short-term wave and impact extreme value statistics–Part 2: Deck box dataset

Jule Scharnke [a,*], Sanne M. van Essen [a,b], Harleigh C. Seyffert [b]

[a] Maritime Research Institute Netherlands (MARIN), Wageningen, The Netherlands
[b] Department of Maritime & Transport Technology, Delft University of Technology, Delft, The Netherlands

ARTICLE INFO

Keywords:
Convergence
Waves
Offshore structure response
Wave impact loads
Extreme value statistics
Test duration
Short-term variability
Seed variations
Experiments
Wave-in-deck
Deck box

ABSTRACT

In the assessment of wave-in-deck loads for new and existing maritime structures typically model tests are carried out. To determine the most critical conditions and measure sufficient impact loads, a range of sea states and various seeds (realisations) for each sea state are tested. Based on these measurements, probability distributions can be derived and design loads determined. In air gap model testing usually only few, if any, impact loads occur per 3-hour seed. This can make it challenging to derive reliable probability distributions of the measured loads, especially when only a few seeds are generated. In addition wave impact forces, such as greenwater loading, slamming, or air gap impacts are typically strongly non-linear, resulting in a large variability of the measured loads. This results in the following questions: How many impacts are needed to derive a reliable distribution? How is the repeatability of individual events affecting the overall distribution? To answer these questions wave-in-deck model tests were carried out in 100 x 3-hour realisations of a 10,000 year North Sea sea state. The resulting probability distributions of the undisturbed wave measurements as well as the measured wave-in-deck loads are presented in this paper with focus on deriving the number of seeds and exposure durations required for a reliable estimate of design loads.

The presented study is Part 2 of a combined study on guidance for the convergence and variability of wave crests and impact loading extreme values. The data set of Part 1 ([1]) is based on greenwater loads on a sailing ferry and the data set of Part 2 on wave-in-deck loads on a stationary deck box.

1. Introduction

1.1. Background

The presented study is Part 2 of a combined study on guidance for the convergence and variability of wave crests and impact loading extreme values. The data set of Part 1 [1] is based on greenwater loads on a sailing ferry and the data set of Part 2 on wave-in-deck loads on a stationary deck box. A more general introduction regarding the addressed problem of statistical variability in the assessment of design loads is given in Part 1 of this study. The focus of the introduction of Part 2 will be mainly on the aspects specifically relevant for wave-in-deck type of loading.

For the design or reliability assessment of a maritime structure, it is interesting to know the distribution of wave induced responses for some defined duration. There will be some statistical variability to these distributions due to the stochastic nature

* Corresponding author.
E-mail address: j.scharnke@marin.nl (J. Scharnke).

https://doi.org/10.1016/j.marstruc.2023.103411
Received 30 November 2021; Received in revised form 6 February 2023; Accepted 24 February 2023
Available online 9 March 2023

Table 1
Test programme for the deck box at zero speed, where μ = heading (180 deg is head waves), H_s = significant wave height and T_p = peak wave period. All waves: long-crested JONSWAP wave spectra with peak enhancement factor 2.9.

Test condition	Duration (h)	μ (deg)	H_s (m)	T_p (s)	Set-up	# wave crests	# impact peaks
10,000 yr North Sea	300:00:00	180	17.0	15.9	With deck box	79 859	254
1 deterministic repeat of all seeds	300:00:00	180	17.0	15.9	With deck box	79 859	246
Undisturbed waves 10,000 yr North Sea	300:00:00	180	17.0	15.9	Wave only	79 859	–

of the ocean excitation. For strongly non-linear responses such as wave impact forces (e.g., greenwater, slamming, air gap impacts), this variability may be relatively large, and will further depend on the test duration itself. To derive a reliable distribution a certain number of measurement points are required. In air gap analysis typically only very few, if any, impacts occur in a 3-h realisation of a sea state. Therefore, to be able to derive design loads from these few occurrences a large number of 3-h realisation (or seeds) of the same sea state are required. For wave-in-deck type of model testing typically a maximum of 20 seeds (more commonly less) are used to determine deck impact loads. In the present study the objective was to determine the minimum required number of seeds for a reliable prediction of the expected extreme impact loads. For this purpose wave-in-deck model tests were carried out in 100 × 3-h realisation of a 10,000 year North Sea sea state. The resulting probability distributions of the undisturbed wave measurements as well as the measured wave-in-deck loads are presented in this paper.

In many impact load related model test campaigns a large variability of measured loads in repeat measurements is observed. This is due to the fact that especially local loads but also global load measurements (see [2] and [3]) are very sensitive to small variations in the incoming wave conditions, resulting in large variations of the measured impact loads. In order to analyse how the repeatability of individual events affects the overall distribution, the complete set of the impact load measurements with the 100 x 3 h seeds was repeated once (see also Table 1) and the resulting distributions and required seed variations of the two repeat data sets are compared.

1.2. Exposure duration and number of seed variations

Please refer to Part 1 [1] of this publication for a detailed elaboration regarding the topic of exposure duration and number of seed variations required for converged estimates of extreme values.

1.3. Rules, guidelines and results from literature

In the past decades a lot of research has been carried out in determining wave-in-deck impact loads. An elaborate air gap model test campaign for the Snorre TLP including a large number of seed variations for a range of 10,000 year and 100 year sea state is presented in [4]. Following this test campaign a lot of work has been carried out in the development of models for more realistic extreme waves and design methodology, as for instance within the framework of Joint Industry Projects (JIPs) such as the CresT JIP [5] and its follow-up project, the ShorTCresT JIP [6]. Extensive work on simplified loading models for the estimation of vertical deck loads in comparison with model tests and simulations was carried out as part of the WaveLand JIP (Phase I & II) and is described in [7,8] and [9].

More recently air gap analysis and related wave-in-deck impact loads are addressed in the DNV-GL Offshore Technical Guidance OTG-13 [10] and OTG-14 [11], which are the most widely-used references for air gap and impact load analysis for column stabilised mobile units. In [11] it is advised to use a total of at least 16 seeds with a duration of 3 h each for the governing sea states in air gap model testing, and to have at least a total of 8–10 impact events for a Weibull estimate. These are estimated values based on existing model test data. In the work of [12] it is shown that 50–60 seed variations (3 h duration) may be required for slamming type of impacts on a large-volume cylinder. In [2] more than 30 seed variations are applied for two different sea states in wave-in-deck model testing. The presented results showed differences in the underlying physical processes especially for the highest impact loads that were measured, resulting in a change in the relation of crest height and measured impact load compared to the rest of the measured impacts. This may lead to different loading regimes within the same data set and might require a different fitting procedure ('badly-behaving' problem, see [13]). Altogether the available publications on the topic indicate that even if a large number of seed variations are measured for very non-linear problems such as wave-in-deck loads, deriving design loads can still be challenging.

More references regarding rules and guidelines for extreme wave impact event analysis are given in Part 1 of this publication [1].

2. Objectives and approach

In the assessment of wave-in-deck loads typically the following questions are raised: How many impacts are needed to derive a reliable distribution? How is the repeatability of individual events affecting the overall distribution? In an attempt to answer the above questions, wave-in-deck model tests have been carried out with 100 × 3 h realisations of a 10,000 year North Sea state. The results of these impact measurements and the related undisturbed wave measurements are presented in this paper.

The main objective of the presented work is to:

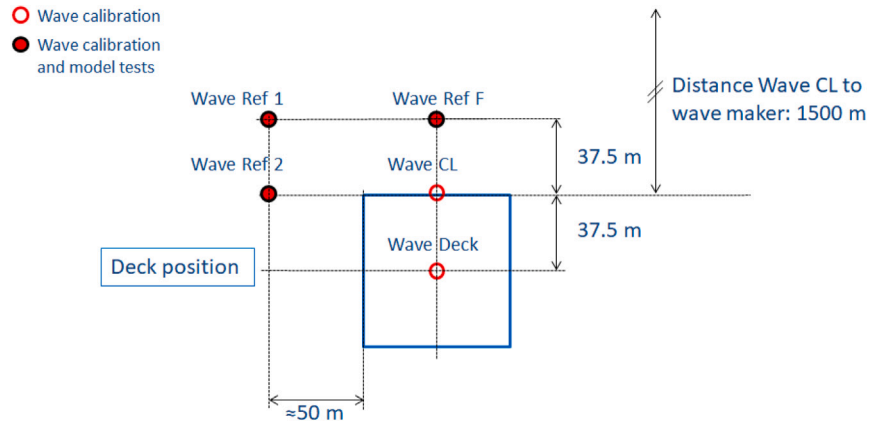


Fig. 1. Position of deck box and (reference) wave probes with respect to wave maker in MARIN's Concept Basin, top view.

- Quantify the convergence of short-term wave crest and impact force peak distributions based on experiments and for different exposure durations.
- Evaluate the effect of variability in repeat measurements on the resulting design loads.

Short-term extreme response values used in design are often required in the form of the most probable maximum (MPM) or a quantile value q. For offshore structures, quantile values of 85%–95% are considered a reasonable choice for design (e.g., [14,15]). For ships, the MPM is often considered as design value (e.g., [16]). It should also be noted that a higher quantile of a 1 h distribution is equivalent to a lower quantile of a 3 h distribution. There are therefore many ways to present the same results. In order to provide results that can be applied to the existing experimental data, we stuck as much as possible to the industry standard/common way that tests are performed in applied research institutes.

For comparison purposes with Part 1 of this publication [1], the analysis carried out in this paper is based on the most probably maximum (MPM) of wave crest heights and wave impact force peaks for different exposure durations (see also Section 4.2). The required number of seeds for convergence is found by analysing a long-duration wave impact experimental campaign. The total test duration was cut up into different combinations of exposure duration and number of seeds. The variability of the MPM values was then evaluated as a function of the number of seeds. Thus, in this case using MPM values is a means to evaluate the convergence of the distributions for different exposure durations. It is not the objective of this study to evaluate different approaches to derive design load values. For reference, also the 90% quantile values are presented for the full data set in Sections 5.2 and 5.4.

The model test setup, measured signals and test matrix are presented in Section 3. The data analysis methods are presented in Section 4. The results of this study are presented and discussed in Section 5 with the resulting wave crest distributions presented in Section 5.1, the impact load distributions in Section 5.2, the fitted distributions of both crests and impacts in Section 5.3 and a comparison of different approaches for deriving design loads in Section 5.4. In Section 5.5 the convergence of the extreme values of the wave crests and impact force peaks with and without fitting for an increasing number of seeds is presented and discussed, and in Section 5.6 the bias resulting from the fitting of the data sets. In Section 5.7 the results of the deck box data set are compared to the results of the ferry model tests presented in Part 1 of this publication [1]. In Sections 6 and 7 the conclusions and possible future work are presented.

3. Experiments

All results in the present publication are provided as full scale values. The model tests were carried out at scale of 1 to 50 in MARIN's Concept Basin. The Concept Basin has a length of 220 m, a width of 4 m and a depth of 3.6 m. A wave generator is fitted at the one end of the basin and a beach on the other end. The wave generator consists of 8 hinged flaps. The basin has a stiff overhead carriage which runs over the full width of the basin. For the wave-in-deck model tests the deck box model was fixed under the basin carriage in a fixed position with the front of the deck box at a distance of 1500 m from the wave maker. The test setup including positions of the deck box and (reference) wave probes is shown in Fig. 1. The wave probe signals used in the analysis presented in this paper are based on the measurements of the wave probe, which is positioned at the front of the deck box at the centre line (CL), wave probe Wave CL.

The deck box used for the wave-in-deck model tests was 75 m long and 75 m wide. The height of the deck box was selected based on the calibrated wave conditions, to ensure 2 to 3 wave impacts per three-h realisation. The bottom of the deck box was set at 17.5 m above the waterline based on the undisturbed wave measurements. An impression of the deck box is given in Fig. 2.

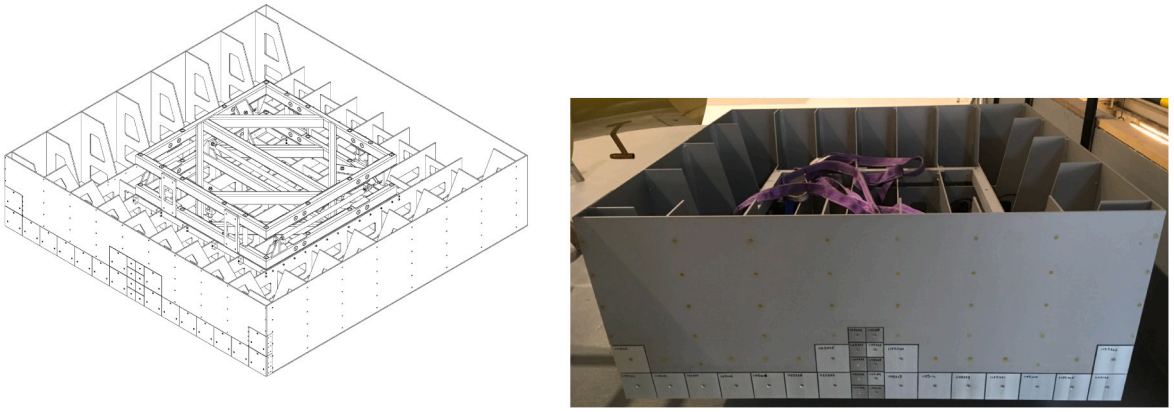


Fig. 2. Impression of deck box model used for wave-in-deck model tests.

3.1. Instrumentation

The following signals were measured during the impact load measurements:

- Global deck loads in 6 degrees of freedom
- Force panels in front side of deck box
- Pressure cells in underside of deck box
- Accelerations

The results regarding the impact load measurements presented in this paper are focused on the horizontal impact load measurements (Fx) based on the sum of the force panel measurements positioned in the front of the deck box. The global deck load measurements, accelerations and pressure cell measurements are not used for the data analysis and results presented in this paper. They are therefore not described in more detail.

3.1.1. Force panels

Two types of force panels were used: 50 × 50 mm model scale and 100 × 100 mm model scale (2.5 × 2.5 m and 5 × 5 m full scale). Both types of force panels were instrumented with one strain gauge type force transducer. The natural frequency of the small panel was analysed through FEM analysis and was estimated to be approximately 2000 Hz (model scale value). The natural frequency of the large panels is approximately 1780 Hz (model scale value) based on the FEM analysis. The natural frequencies of the force panels were verified by hammer tests carried out at the start of the test campaign. More details on the force panels used in this test campaign can be found in [17].

The locations of the force panels in the front of the deck box are shown in Fig. 3 (values are at model scale). The lower edge of the front of the deck box was covered over the entire width of the deck box with the larger force panels. At four locations, at the edges and in the middle, a second row of larger force panels was placed. This results in a total of 18 large force panels. 10 small force panels were placed in the middle of the deck box for a section with higher resolution of the measured impact loads. Between each force panel a clearance of 1.5 mm (model scale value) is introduced to ensure that there is no interaction between the panels. The sampling rate of the measured force panel signals was 19,200 Hz (model scale value) for all panels.

3.2. Test matrix

The experimental conditions for the deck box are listed in Table 1. A 10,000 year North Sea sea state was selected for the model tests to ensure realistic wave characteristics and crest heights. The first 3-h realisation was calibrated to match the according theoretical JONSWAP formulation. After the calibration, an additional 99 realisations of the same sea state were measured. All wave conditions were generated in one wave direction only, perpendicular to the front of the deck box.

With these 100 seed variations a series of measurements were carried out. First the undisturbed waves were measured at the locations shown in Fig. 1. Then the deck box was installed and the same 100 seed variations were generated to measure the impact loads on the deck box. To evaluate the repeatability of the resulting impact load distribution the same 100 seed variations were run again with the deck box in the same position. The results of the undisturbed wave measurements and the two sets of repeat measurements are presented in the following sections.

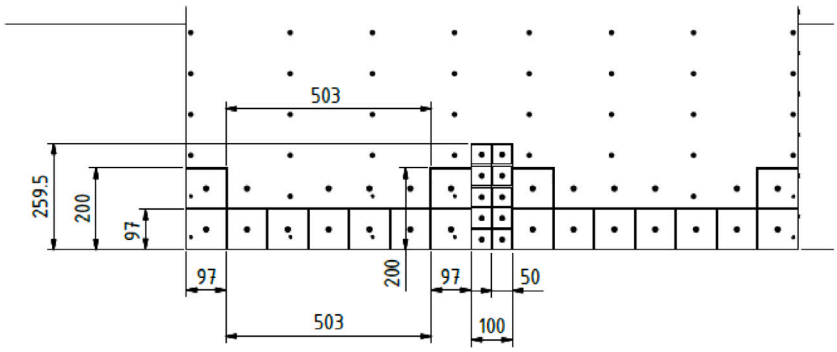


Fig. 3. Distribution of force panels in front of deck box (values at model scale).

4. Definitions and analysis methods

4.1. Concatenation of runs to generate long duration time traces

The tests with the deck box were done at zero speed, in three-hour intervals with a start-up of a half hour. The long duration tests given in Table 1 were obtained by removing start-up effects (so exactly three hours remain) and then concatenating all realisations.

4.2. Cutting long duration time traces in realisations with different exposure durations

The total time traces were divided into different combinations of exposure duration and number of seeds. Exposure durations of three and one hour were considered for the undisturbed wave measurements. The available tests allow for the following number of realisations of the undisturbed wave measurements (M):

- 100 × 3:00:00 h;
- 300 × 1:00:00 h.

Because of the low number of measured impacts for each seed (2–3 impacts per 3 h seed), the exposure durations for the measured impact loads had to be longer. They are divided in the following number of realisations (M):

- 5 × 60:00:00 h;
- 10 × 30:00:00 h;
- 20 × 15:00:00 h.

4.3. Peaks, ensemble maxima and distributions

As mentioned in Section 3, wave elevation η_{WCL} and force F_X are the most important signals in the present study. The crests in the wave elevation time traces were identified using a zero up-crossing analysis, where each peak and trough are after the corresponding zero up-crossing and before the next zero up-crossing. The force signal is not continuous; peaks were only measured during wave-in-deck events. A peak over threshold analysis was therefore used to identify its peaks. An ensemble maximum is defined as the maximum peak value in a certain exposure duration. Samples, peaks and ensemble maxima in general have different probability distributions, as illustrated for theoretical Gaussian wave crests in Fig. 4. The crests/peaks in wave elevation and accommodation force are indicated by $\eta_{WCL,C}$ and $F_{X,C}$, and their ensemble maxima by $\eta_{WCL,E}$ and $F_{X,E}$, respectively. In order to avoid contamination of the test results by the concatenation of the runs discussed in Section 4.1, all peaks and extremes within one wave period of a concatenation were removed from the data set. All wave results were furthermore non-dimensionalised with the significant wave height H_s. The wave impact forces are presented as absolute values, as there is no logical reference for them. In practice, an uncertainty requirement for structural design forces would probably be related to the maximum allowable design force on the structure.

We distinguish three types of exceedance probability distributions. Firstly, a DSR distribution is based on all crests or peaks in a single seed (or realisation): $\eta_{WCL,C}$ and $F_{X,C}$ in our case. The probability level in the DSR in the present study is based on the total number of wave encounters n_e for all signals. This is also done for the impact loads, even though they do not peak at every wave encounter. This is required in order to relate the number of impact peaks to a time duration — there are multiple force peaks per exposure duration so this information would be lost if the DSR probability level was plotted based on the number of impact peaks. The probability levels in the DSR are set to $1/n_e$ for the highest value and $(n_e - 1)/n_e$ for the lowest value. Secondly, the DNR distribution is based on all peaks or crests over a number of seeds (or realisations). The DNR is easily derived from multiple DSRs. All peaks in the available seeds are assembled, after which a new distribution is generated. This is effectively a long duration

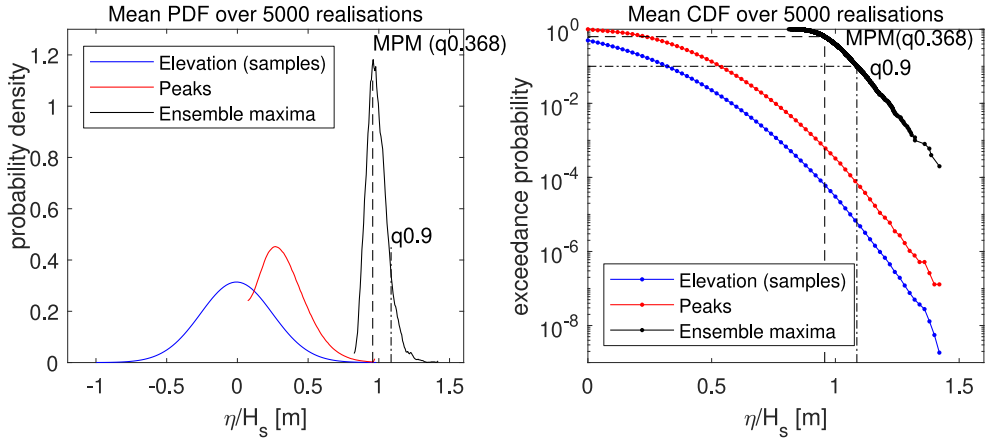


Fig. 4. Typical probability density function and inverse cumulative distribution of samples, peaks and ensemble maxima of wave and response signals — here for 5000 seeds of 3 h linear Gaussian waves with a JONSWAP spectrum and H_s 5 m, T_p 9 s, γ 3.3. The elevation and peak distributions are the mean distributions over all seeds. The MPM (36.8% quantile) and 90% quantile are also indicated.

distribution (e.g., 10 realisations of 3 h lead to a 30 h DNR). The probability level in the DNR is also based on the total number of wave encounters for all signals. Finally, a DEM distribution is based on ensemble maxima: $\eta_{WCL,E}$ and $F_{X,E}$ in our case. It uses the largest value of each seed for a given exposure duration. The probability level for the DEM in the present paper is based on the number of ensemble maxima (so the number of seeds N), both for the wave crests and the impact force peaks. The probability levels in the DEM are set to $1/N$ for the highest value and $(N-1)/N$ for the lowest value. For the waves, some theoretical formulations exist for these distributions. These are described in Section 4.5.

4.4. Extreme values

As already mentioned in Section 2, short-term extreme response values used in design are often required in the form of the MPM or a quantile value q. The point at which 90% of the extreme values in the DEM are smaller is indicated by $q = 90\%$, as also shown in Fig. 4. The MPM is the extreme value that is most likely to occur, which is equal to $q = 36.8\%$ for a linear Gaussian signal (see Fig. 4). The MPM can be estimated from data by taking the mode of the binned or fitted DEM, or by taking the extreme value with 63.2% exceedance probability (so cumulative probability quantile $q = 36.8\%$, assuming the extreme value distribution approximates a Gaussian distribution, see e.g., [18]). The 63.2% exceedance probability is a more robust definition for data sets that are not very large, as the mode definition is very sensitive to outliers, and the fitting definition may introduce extra fitting errors. Even though we do not consider purely Gaussian signals, the 63.2% exceedance probability definition is therefore applied in the present study. This is formulated in Eq. (1), where $S_E(N)$ are the DEM ensemble maxima values of N seeds, D is their exceedance probability distribution and $\widehat{S_E}(N)$ is the MPM value interpolated from N seeds. We use linear interpolation of the empirical distribution for this. Signal S can either be the non-dimensional wave elevation or the impact force (with the ensemble maxima values $\eta_{WCL,E}$ and $F_{X,E}$ defined above). As explained in the previous section, a probability of $1/N$ is used for the highest value in D and $(N-1)/N$ for the lowest value.

$$\begin{aligned} D(s) &= P\left(S_E(N) \geq s\right) \\ D\left(\widehat{S_E}(N)\right) &= P\left(S_E(N) \geq \widehat{S_E}(N)\right) = 0.632 \\ \widehat{S_E}(N) &= \phi_{0.632}\left[S_E(N)\right] \end{aligned} \tag{1}$$

For offshore structures, quantile values of 85%–95% are considered a reasonable choice for design (e.g., [14,15]). Eq. (1) can still be used in that case, with exceedance probability 0.15 or 0.05 instead of 0.632 and the result called quantile instead of MPM. For comparison purposes with Part 1 of this publication, van Essen et al. [1], the analysis carried out in this paper is based on the most probable maximum (MPM) of wave crest heights and wave impact force peaks for different exposure durations (see also Section 4.2). In Sections 5.2 and 5.4 the 90% quantile values are also presented for the full data set for reference purposes.

Alternatively, the MPM can be estimated using a DNR distribution. This is formulated in Eq. (2), where $n_{e,t}$ is the average number of wave encounters in the target exposure duration and P_t is the lowest available exceedance probability level for that exposure duration. $S_C(N)$ are the DNR crest/peak values of N seeds, H is their exceedance probability distribution, and $\widehat{S_C}(N)$ is the MPM value interpolated from N seeds. We use linear interpolation of the empirical distribution for this. Again, signal S can either be the non-dimensional wave elevation or the impact force (with the crest/peak values $\eta_{WCL,C}$ and $F_{X,C}$ defined above), or their fitted equivalents. As explained in the previous section, a probability of $1/n_e$ is used for the highest value in H and $(n_e-1)/n_e$ for the

lowest value (where n_e is the total number of wave encounters in all seeds combined).

$$\begin{aligned} P_t &= 1/n_{e,t} \\ H(s) &= P\left(S_C(N) \geq s\right) \\ H\left(\widehat{S_C}(N)\right) &= P\left(S_C(N) \geq \widehat{S_C}(N)\right) = P_t \\ \widehat{S_C}(N) &= \phi_{P_t}\left[S_C(N)\right] \end{aligned} \tag{2}$$

Again, the DNR can also be used to derive higher quantiles for offshore structures. The definition in [11] for peak-over-threshold quantiles is used for this, modified to account for the total number of wave encounters in all seeds n_e as basis for the exceedance probability. P_t in Eq. (2) is now substituted by P_q that is defined in Eq. (3) (the rest of Eq. (2) stays the same). M in this formulation is the total number of available seeds (see Section 4.2). It was shown in part 1 [1] that P_q approaches P_t for q = 36.8%.

$$P_q = 1 - q_{mod} = 1 - q^{M/n_e} \tag{3}$$

4.5. Theoretical wave distributions

Huang and Zhang [19] argue that the DSR of wave crest heights is often compared to theoretical distributions, whereas these are actually derived for the DNR. These authors therefore introduced new empirical 'Huang-Zhang' upper and lower limits for DSR distributions (for details see Appendix A). The wave crest height results in the present study are compared to these. For the DNR of the wave crest heights, some theoretical forms are available: Rayleigh, Forristall [20], Huang-Zhang and CresT [21]. Details of these distributions are provided in Appendix A. Note that there is a small difference between the DNR and the mean over all available DSR seeds at the same probability level (see also e.g. [19]).

4.6. Fitting

Measured distributions from experiments are often fitted with theoretical extreme value distributions. This is done in order to reduce the statistical variability of the single measured maxima or to extrapolate to a longer duration. Especially when experimental results with a short exposure duration are used and/or far extrapolations, this can lead to errors. Wave impacts can be 'badly-behaving problems' [13], where the phenomena that play a role for low-magnitude responses are different than those for high-magnitude responses (e.g., plunging or dambreak versus hammerfist green water events). This makes extrapolation difficult, if not impossible. However, fitting and extrapolation of the DSR based on only a few seeds are commonly used in applied research to obtain extreme wave impact design values. The present study therefore also evaluates the benefits of fitting to the DSR and DNR.

For measured DSR/DNR crest distributions in ship and offshore design, many authors successfully apply or recommend 2- or 3-parameter Weibull fitting (e.g., [16,22–24]). We therefore used the 3-parameter Weibull distribution (Eq. (4)). The distribution was fitted to the data using least-squares, as DNV [14] states that this often gives a better tail fit for Weibull fitting than the method of moments or maximum likelihood estimation. The influence of the higher peaks was further increased by only considering a certain percentage of the highest peaks in the fits. The applied percentage is indicated in each plot in Section 5.3.

$$P(x > X) = \exp\left(-\left(\frac{x-\theta}{\alpha}\right)^{\beta}\right) \tag{4}$$

In the remainder of this paper, we compare the following extreme values and their variability for an increasing number of seeds:

1. MPM from DEM that is based on the measured DSRs ('MPM from measured data').
2. MPM from DEM that is based on the Weibull-fitted DSRs ('MPM from fitted data').
3. MPM based on the measured DNR ('MPM from DNR').

A fourth one could be from a GEV fit to the empirical DEM. However, the objective of the present paper is to evaluate the convergence of the results by increasing the number of included seeds (see next section), which means that we start from very low number of seeds. Obtaining a MPM value from (in the limiting case) only 2 or 3 DEM values is already questionable, but fitting a GEV to only 2 or 3 data points and then obtaining the MPM was judged to be unfeasible. Options 1 to 3 also align with what is done in practice in wave basins to evaluate the MPM of marine structure response. We therefore stuck to the three options above.

4.7. RMSE convergence metric for MPM values

The main objective of the present work is to evaluate how many seeds are required for a given extreme value accuracy. To this end, the root mean square error (RMSE) for the MPM value over a given number of seeds is introduced. The RMSE defined here expresses the difference between the MPM over a randomly selected number of seeds (1*, 2*, …, N) and the MPM over all available seeds (M). The latter is seen as the 'true' value (some comments about this assumption can be found in Section 4.9). The indices with a star indicate a randomly picked set of N seeds out of the available M (random permutation without repeating elements). The number N is variable, and M was fixed during the performance of the experiments. By increasing the number of seeds included in the calculation N, its effect on the RMSE can be evaluated.

Each time we randomly pick N seeds leads to a different result. We will call this the number of 'seed picking realisations' W, not to be confused with the random wave seed (also called wave realisation) itself. The lowest number of available seeds M is 5 (60 h impact force measurements) and the highest number is 100 (3 h wave measurements). The number of combinations for an 'unordered sampling without replacement' seed picking routine is given in Eq. (5). Calculating the MPM over all combinations is not feasible: $N = 50$ and $M = 100$ for instance gives 1.0×10^{29} options. Instead, we randomly pick 500 times for all combinations of N and M ($W = 500$). This provides a decent indication of the RMSE for larger N: the curves of the RMSE plots in Section 5.5 became quite smooth, whereas they were very peaked for low numbers of W.

$$n_{options} = \frac{M!}{N!(M-N)!} \tag{5}$$

Finally, the RMSE was calculated over W. This is formulated in Eq. (6), where the MPM values $\widehat{S_E}(N)$ are defined as in Eq. (1) and index w refers to the seed picking realisations. We do this for all possible values of N, in order to obtain the convergence RMSE for the MPM values as a function of N. The seed picking routine is 'unordered sampling without replacement': we draw N samples from the set of M, such that ordering does not matter and repetition is not allowed. We used $W = 500$ for our analysis. For example, 100 seeds are available for the 3 h exposure duration of the wave measurements (so $M = 100$). We randomly pick 10 seeds out of these 100 ($N = 10$), and repeat this 500 times ($W = 500$). We then calculate the RMSE over these 500 seed picking realisations, for the MPM over the 10 seeds compared to the MPM value over the fixed 100 seeds. This can be done for any value of N between 1 and M.

$$R(N) = \sqrt{\frac{\sum_{w=1}^{W}\left[\widehat{S_{E,w}}(N) - \widehat{S_E}(M)\right]^2}{W}} = \sqrt{\frac{\sum_{w=1}^{500}\left[\widehat{S_{E,w}}(N) - \widehat{S_E}(M)\right]^2}{500}} \tag{6}$$

The procedure above was used to evaluate the convergence of the extreme values in the measured original data (option 1 in Section 4.6). For the DSR-fitted data (option 2 in Section 4.6), there are two different things to evaluate: the convergence over an increasing number of seeds, and the accuracy of the fitted data compared to the measured data (the bias). The bias is discussed in Section 4.8. The convergence of the fitted data was evaluated in the same way as for the measured data. This was done using Eq. (6), but with S_E in this equation and in the MPM definition of Eq. (1) substituted by the ensemble maxima in the fit (option 2 in Section 4.6) at the same probability levels S_E^f. This results in Eq. (7).

$$R^f(N) = \sqrt{\frac{\sum_{w=1}^{W}\left[\widehat{S^f_{E,w}}(N) - \widehat{S^f_E}(M)\right]^2}{W}} = \sqrt{\frac{\sum_{w=1}^{500}\left[\widehat{S^f_{E,w}}(N) - \widehat{S^f_E}(M)\right]^2}{500}} \tag{7}$$

The RMSE values were calculated for the measured wave and impact data and for the fitted wave and impact data, for exposure durations 3 h and 1 h for wave conditions and 60 h, 30 h and 15 h for the measured impact forces. The results can be compared to an arbitrary convergence criterion, in order to determine how many seeds are required. This value will depend on the application; it can be a relative uncertainty criterion (e.g. 2% of H_s for the wave crests), an absolute uncertainty criterion, a criterion based on the derivative of the RMSE with respect to the number of seeds (the slope of the lines presented in Section 5.5) or a combination.

4.8. RMSE bias metric for MPM values

To evaluate the bias of the fits compared to the measured values, a similar formulation was used. This bias RMSE was again defined based on the MPM defined in Eq. (1), and W is again 500. S_E^f here indicates the ensemble maxima in the fit (option 2 in Section 4.6), and S_E again the ensemble maxima in the measured signal. This RMSE value for the fitted data does not converge to zero for larger N, but to a constant value: the bias introduced by fitting the DSRs.

$$B^f(N) = \sqrt{\frac{\sum_{w=1}^{W}\left[\widehat{S^f_{E,w}}(N) - \widehat{S_E}(M)\right]^2}{W}} = \sqrt{\frac{\sum_{w=1}^{500}\left[\widehat{S^f_{E,w}}(N) - \widehat{S_E}(M)\right]^2}{500}} \tag{8}$$

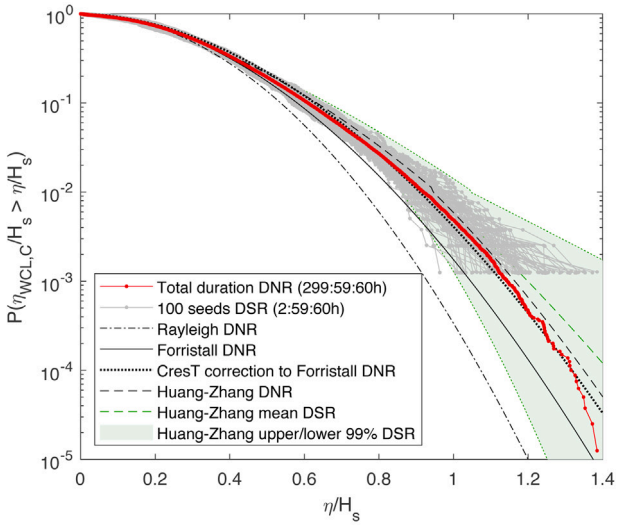
4.9. Assumptions

As mentioned, some assumptions were made in the RMSE convergence and bias methods. Firstly, it is assumed that the MPM value over the maximum number of available seeds M is converged and the 'truth'. This is not necessarily true of course, but it is the best estimate that we have based on the available data.

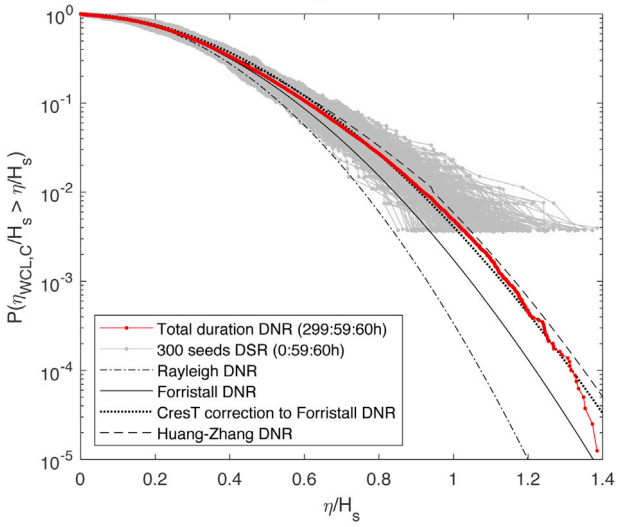
5. Results and discussion

5.1. Wave crest distributions (DSR and DNR)

The DSR and DNR wave crest distributions measured at wave probe Wave CL (see Fig. 1) were derived using the procedure explained in Sections 4.1–4.3 and 4.5. The result is shown in Fig. 5 for 3 h and 1 h exposure duration (in grey). No repeat



(a) 100 × ∼3h exposure duration

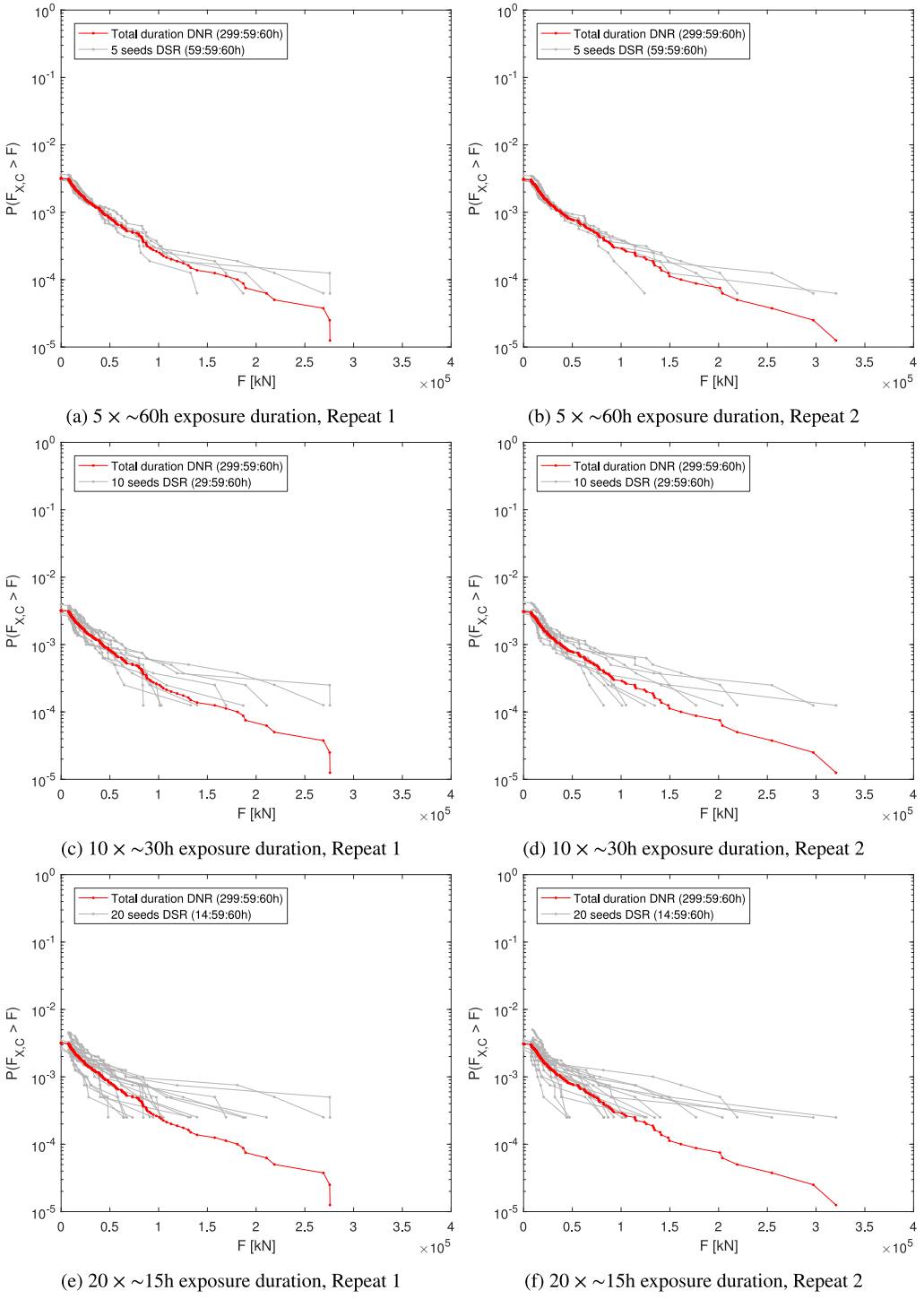


(b) 300 × ∼1h exposure duration

Fig. 5. DSR and DNR distributions of wave crests at wave probe Wave CL $\eta_{WCL,C}$ for different exposure durations and number of seeds (including theoretical crest height distributions).

measurements were carried out of the undisturbed waves. Repeat measurement results are only available for the impact load measurements presented in the following sections. The DNR is shown in red in Fig. 5. The measured crest distributions are compared to the Huang-Zhang DSR limits for three hours exposure and to the Rayleigh, Forristall and Huang-Zhang DNR and CresT distributions (see Appendix A). The comparison of the measured data with these theoretical distributions is discussed in Appendix A.1. The variation in DSR over all seeds decreases with increasing exposure duration; more crests are measured per seed, so the variability between the seeds decreases. This can be observed in Fig. 5, by comparing the range of the highest crest heights of each DSR at the lowest probability level. For 1 h exposure duration this range is much larger than for 3 h exposure duration. This variability is also reflected in the convergence results comparing the different exposure durations presented in Section 5.5.

As expected, the plots also show that the ensemble maximum wave crest heights (the largest value in each seed) increase with increasing exposure duration (see also Section 5.4). This is a direct consequence of the stochasticity of waves. It confirms again that the exposure duration should be selected based on the expected steady state duration at sea: this will deliver realistic extreme values.



(a) 5 × ∼60h exposure duration, Repeat 1

(b) 5 × ∼60h exposure duration, Repeat 2

(c) 10 × ∼30h exposure duration, Repeat 1

(d) 10 × ∼30h exposure duration, Repeat 2

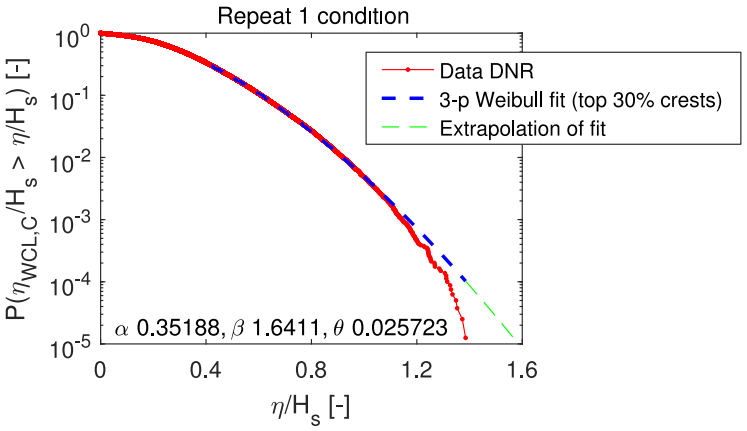
(e) 20 × ∼15h exposure duration, Repeat 1

(f) 20 × ∼15h exposure duration, Repeat 2

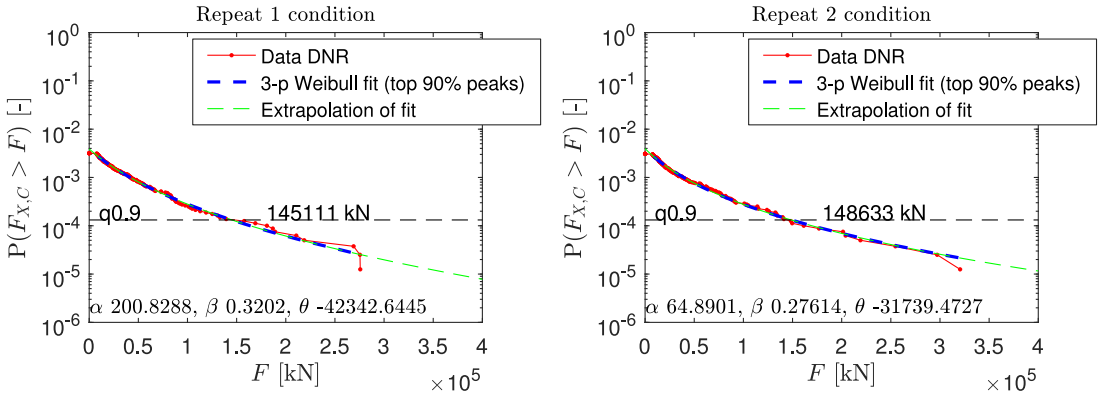
Fig. 6. DSR and DNR distributions of total wave impact force peaks on deck box F_X, for different exposure durations and number of seeds.

5.2. Impact force peak distributions (DSR and DNR)

Similar DSR and DNR distributions are shown in Fig. 6 for all wave impact force peaks on the deck box (see Section 3 for a description of the measured and analysed forces). In the case of the measured impact forces the results are shown for 60 h, 30 h and 15 h duration as explained in Section 4.2. The exposure is longer for the measured impact loads, because only very few (2–3) impacts



(a) Waves, 100 × ∼3h exposure duration



(b) Impact loads, 100 × ∼3h exposure duration, Repeat 1 (c) Impact loads, 100 × ∼3h exposure duration, Repeat 2

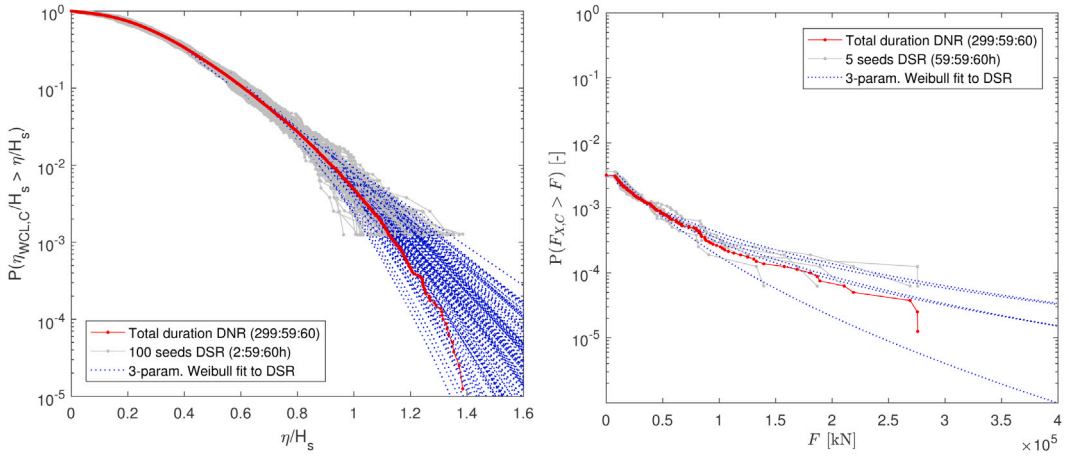
Fig. 7. DNR least-squares fitting result for wave crests measured at probe Wave CL (top) and wave-in-deck impact forces (bottom) with 3-parameter Weibull, considering 30% of the highest crests and 90% of the highest impact peaks.

occur for each 3 h wave realisation. No theoretical distributions are available for the distribution of the impact loads. However, similar trends are visible as for the wave crest distributions:

- An increasing exposure duration decreases the variability between the seeds, and increases the mean of the ensemble maximum peak values over all seeds (see also Section 5.4).
- The variability in the highest peak forces is a lot higher than for the wave crest heights, especially for shorter exposure duration. This gives a clear indication that a long exposure duration is required for a reliable estimate of the impact loads.
- The DNR distributions of the two sets of repeat measurements are quite similar except for the highest impact loads in the tail of the distribution. The highest measured impact of Repeat 2 is clearly higher than the highest measured impact load of Repeat 1. This also results in a larger variability of the DSR distributions, mainly for the 30 h duration case. This is also reflected in the convergence results discussed in Section 5.5 and presented in Fig. 10.

5.3. Fitting of crest and peak distributions

In Section 4.6 the fitting procedure is presented. For reference, we first apply this procedure to the DNR distributions. This provides an idea of the suitability of the Weibull fit over a large range of crest heights and impact loads. The result is shown in Fig. 7 for the wave crests at probe Wave CL and wave-in-deck impact force peaks for both repeat data sets. For the wave measurements the 30% highest peaks were used for the fit. Since less data points are available for the impact loads, the 90% highest peaks are used for the fit. The figures show that Weibull distribution fits the wave crest data very well, up to a certain limit. This is likely due to the influence of wave breaking for the highest and steepest crests. Based on this, Weibull is expected to give a slight overestimation of the wave crest heights when extrapolated to lower probabilities. For the impact force peaks, the fit seems very good for both sets of repeat measurements with some deviations of individual impacts in the tail of the distribution. However, no indications of a very badly-behaving problem (see [4,13]) are found for the full data set of 300 h wave data and impacts. Even though the highest



(a) Waves, 100 × ~3h exposure duration (b) Impacts, 5 × ~60h exposure duration

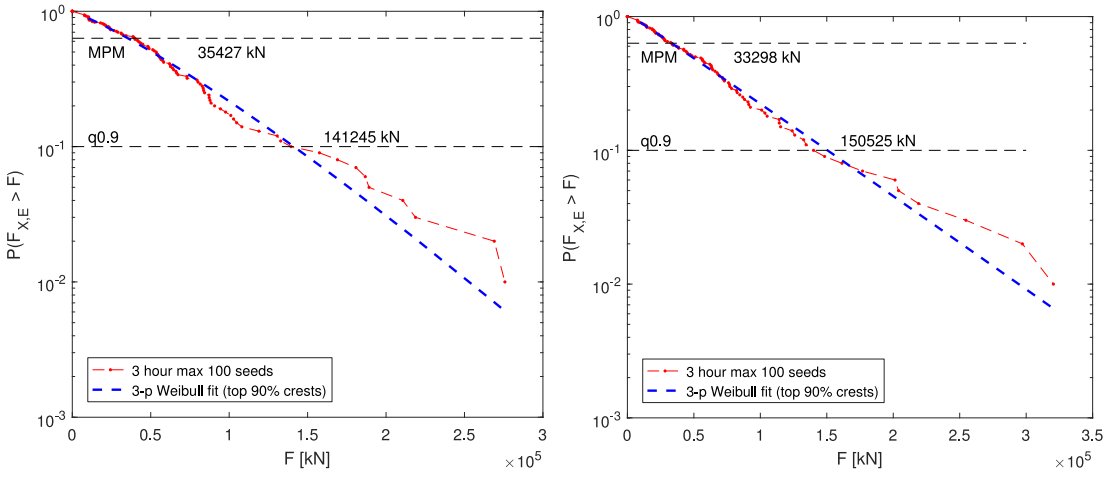
Fig. 8. DSR and DNR of η_{WCLC} (left) and $F_{X,C}$ (right), crests for 3 h exposure duration, impact loads for 60 h exposure duration. Including 3-parameter Weibull fits to top 30% crests (waves) or top 90% peaks (forces) in the DSR distributions.

measured impact load of Repeat 2 is higher than that of Repeat 1, the overall distributions look very similar. The highest measured impact loads are approximately 275,000 kN for Repeat 1 and 320,000 kN for Repeat 2 respectively. The impact loads at the 90% quantile level of the fitted distributions of the two repeat measurements are approximately 145,000 kN for Repeat 1 and 149,000 kN for Repeat 2 respectively. Thus, based on the fitted data the 90% quantile level of the impact force of the second data set is approximately 3% higher than that of the first data set. Considering the range of variation in the individual events in the tail of the distributions (the highest event of Repeat 2 is about 16% higher than the highest event of Repeat 1) a 3% uncertainty seems rather small. This suggests that if sufficient data is available, the variability of individual impact events in repeat measurements does not affect the overall distribution and resulting 90% quantile level impact load significantly. It must be noted that this comparison is based on the horizontal deck load measurements (sum of force panels in the front of the deck box). The variability between repeat measurements of local load measurements (e.g., individual force panels) can be much higher than the presented horizontal deck loads. And even more impact peaks may be required for converged design values between repeat measurement data sets. The available data set also includes local load measurements, but the results of these measurements are not included in the present study.

Fitting the DSR data could be used to reduce the uncertainty of the measured maxima and to extrapolate results to longer exposure durations. Each short-duration DSR seed distribution was therefore fitted in the same way as the DNR. For the waves, the 30% highest crests were used for fitting. Again fewer events are available per seed for the wave-in-deck impact load measurements, so the 90% highest force peaks were used for fitting. The results for the waves and wave-in-deck impact loads for 3 h (waves) and 60 h (impacts) exposure duration are shown in Fig. 8. The results of the remaining exposure durations are presented in Appendix B. These plots seem to confirm that the variability between the seeds may reduce when proper fitting methods are applied to the DSR. This would be good news: compared to the unfitted results, it may be possible to get closer to the longer duration DNR value based on a faster convergence. There are however also a few remarks. Firstly, the plots show that while the fits may in some cases decrease the variability, they may also increase the mean deviation from the DNR at most probability levels. Secondly, the quality of the fit plays a large role. This quality is higher for fits based on measured data with a longer exposure duration. This is quantified in Section 5.6. It should also be kept in mind that we use one fitting procedure, where many are possible. Another choice of fit distribution type, fitting procedure or included percentage of peaks may lead to different results. However, the present procedure complies with common practice (as explained in Section 4.6), so it provides a fair indication of what can be expected.

5.4. Deriving design loads based on DEM vs. DNR

As described in Section 2 design values are typically derived in the form of most probable maximum (MPM) or a quantile value q. For offshore structure quantile values of 85%–95% are typically applied. These design values would typically be based on the DEM (Distributions of Extremes, see Section 4.3). If a 3-h MPM value is required this would however require a large number of realisations for a proper fit. In OTG-14 [11] for example a minimum of 8–10 events is recommended for a Weibull estimate. If not enough realisations are available a peak over threshold approach can be applied, selecting all impact loads that exceed a threshold value (DNR). This is the approach selected in this study, to ensure that enough impact peaks are available to study the effect of different exposures. To be able to relate the probability of exceedance of the assembled impact peaks to the underlying wave condition, the probability level of the total wave encounters is applied (see also Section 4.3). The peak-over-threshold quantile levels are modified to account for the number of available seeds at a certain duration in relation to the total number of wave encounters (see Eq. (3) in Section 4.4).



(a) 100 × ~3h exposure duration Repeat 1 (b) 100 × ~3h exposure duration Repeat 2

Fig. 9. Least-square fitting results for 100 × 3 h max impact load values, including resulting MPM and 90 percentile values.

In this study a large data set is available with 100 3-h realisations with on average 2–3 impact events per 3-h realisation. This allows for a comparison of the different approaches and resulting estimates of the design loads. The present section compares the MPM and 90% quantile level values between the approaches based on the DNR (using all peaks) and on the DEM (using only the maximum peak of each seed). In Fig. 9 the distribution of the maximum impact load per 3-h realisation (DEM) is presented for both repetitions of the impact load measurements. In addition the 3 parameter Weibull fits as well as the resulting MPM and 90% quantile values are presented. In Fig. 7, presented and discussed in Section 5.3, the DNR distributions, resulting fits and 90% quantile values of all measured impact loads of the presented 10,000 year sea state are presented.

Based on the results presented in Figs. 7 and 9 it can be observed that the derived 90% quantile level values between the DEM and DNR are very similar (within 3% and 2% for the two repeat data sets). This indicates that sufficient data points are available for an analysis based on DNR results as well DEM. Furthermore, it also indicates that the individual impact loads are independent from each other and it is justified to assemble them in one distribution (DNR). If the derived quantile levels would deviate between the two distributions this could indicate one of the following cases:

1. There are insufficient impact peaks for a DEM and we cannot derive a proper distribution. In this case [11] suggests to use all measured impact loads above a defined threshold.
2. The data contains a mix of impact peaks from different loading regimes (badly-behaving problem). This can result in different distributions for DEM and DNR, depending on how many cases of each loading regime are in each distribution. Depending on which loading regime is dominant for each distribution the resulting 90% quantile level value may be quite different for DEM vs DNR. An example of this case is given in [25].
3. The impact peaks of the DNR are not independent from each other.

In line with the observation made in Section 5.3 the resulting design values of the two repeat sets are very close to each other, with a slightly larger variation of approximately 6% compared to the 3% difference based on the DNR results. This is likely related to the fact that fewer data points are available for the DEM than for the DNR.

5.5. MPM convergence: original and fitted data

In this section the convergence of the distributions with varying duration is presented. The main objective of the present study is to determine how many seeds are typically required for converged extreme values of wave crests and impact force peaks. Based on the RMSE convergence metric $R(N)$ defined in Section 4.7, calculated for an increasing number of seeds N this can be evaluated for the presented undisturbed wave measurements and wave-in-deck loads. The results are shown in Fig. 10 for the undisturbed wave data and measured impact loads for each exposure duration. The results on the left are based on the measured data; the results on the right are based on the fitted data. The plots show the mean result over 500 seed picking realisations, as explained in Section 4.7.

The presented figures provide an indication of the accuracy that can be expected for the wave crest or impact peak MPM values in different exposure durations and wave conditions, for an increasing number of seeds. They can also be used to derive the required number of seeds in order to obtain extreme wave crest and impact peak results with an arbitrary RMSE uncertainty criterion (see Section 4.7) for the present test case. The following can be observed from Fig. 10:

- The influence of fitting on the required number of seeds for convergence is small for most cases. The 60 h case of Repeat 2 of the impact measurements seems to show a faster convergence.

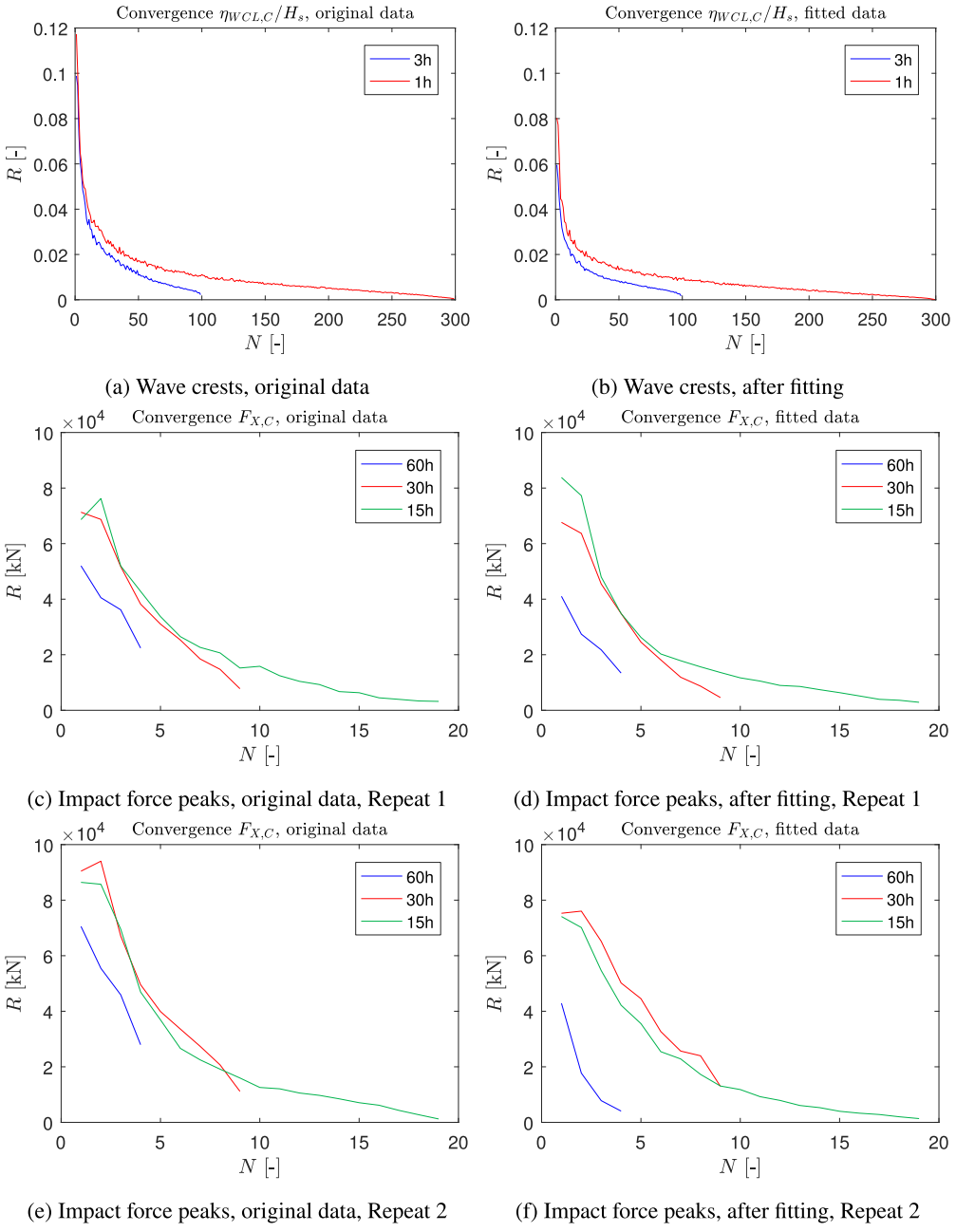
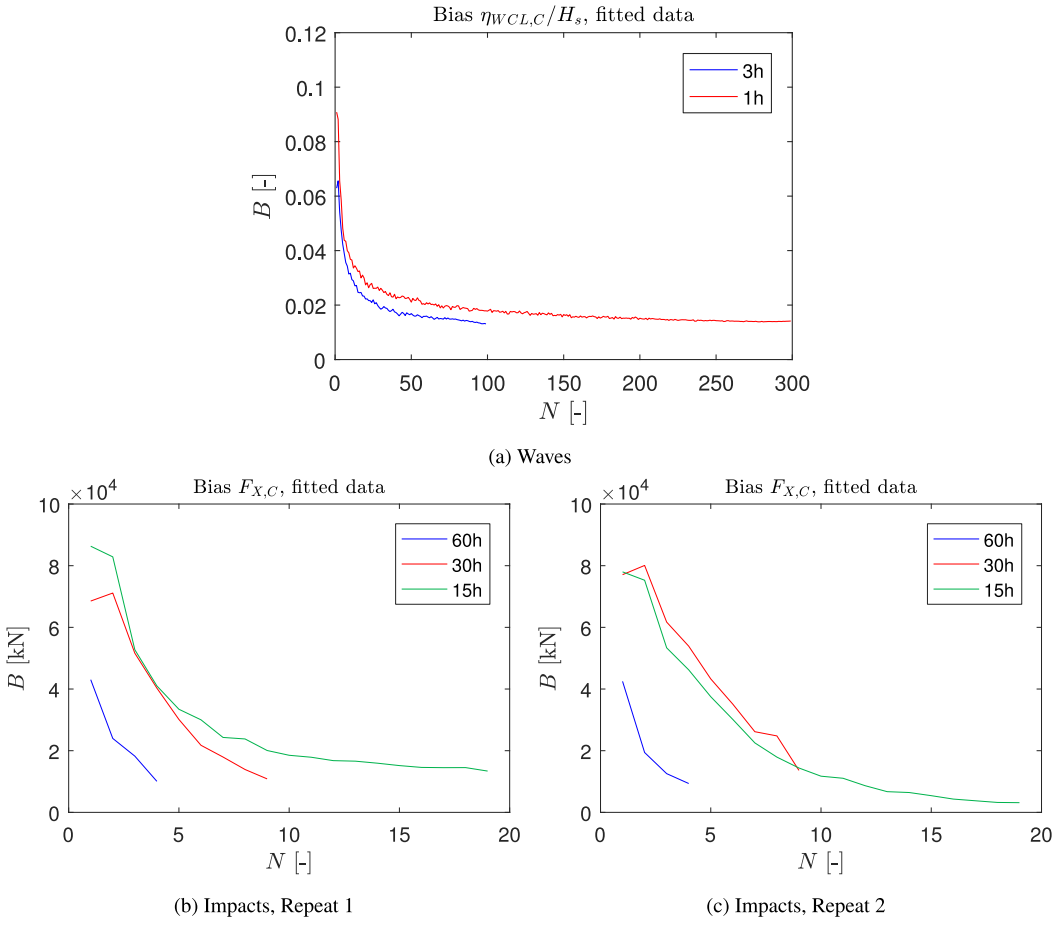


Fig. 10. Convergence metric $R(N)$ for the MPM value for waves and impacts, original and fitted data. Mean of 500 random seed picking realisations.

- For longer exposure durations, convergence is reached with fewer seeds. The total required test time however, may not be shorter.
- The fitted data of the wave measurements initially appear to converge faster up to approximately 50 seeds. However, this also results in a larger bias of the fit as shown in Section 5.6.
- For the two repeat measurement sets very similar results are achieved for most cases duration. The 30 h duration results of the fitted data however converge more slowly for Repeat 2 compared to Repeat 1. This may be related to the larger variability of the 30 h DSRs as also observed in Section 5.2. This also results in a larger bias of the fit as presented in Section 5.6.

It might be expected that the required number of seeds is lower for more linear problems (less steep waves) and higher for more non-linear or 'badly-behaving' problems (where other physical phenomena play a role in causing low (or no impacts if e.g. the air



(a) Waves

(b) Impacts, Repeat 1

(c) Impacts, Repeat 2

Fig. 11. Bias metric $B(N)$ for the MPM value for waves and impacts, fitted data. Mean over 500 random seed picking realisations.

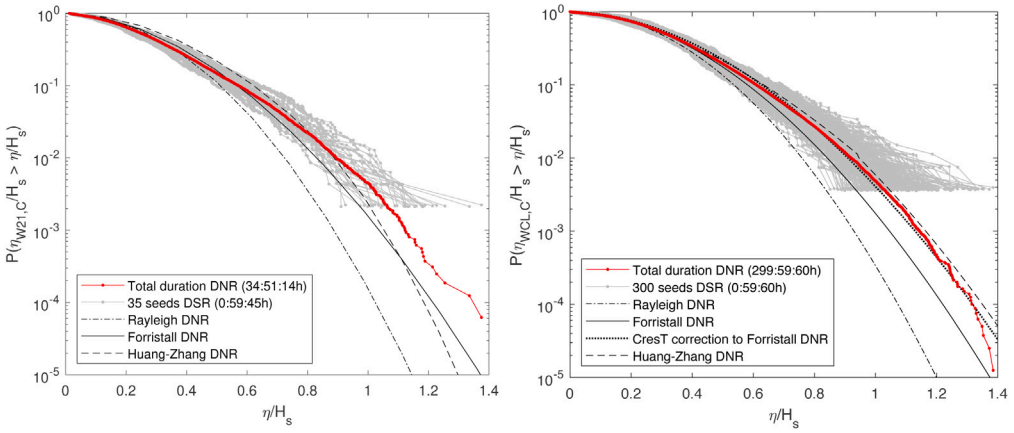
gap is not exceeded) and high impact load peaks). For lower sea states it is hard to say what is expected; the waves and impact events are expected to be more linear and well-behaved, but there will also be fewer impact peak events per seed (if any).

5.6. MPM bias due to fitting

Fig. 10 showed the convergence metric for the RMSE of the exposure duration MPM $R(N)$ Eq. (6) for the original experiments and the fits, which is a good measure for the convergence of the data. However, Fig. 8 and Appendix B show that the fits may also introduce a bias in the MPM with respect to the original data. In Fig. 11 the bias of the MPM based on these fits for the waves as well as the two impact peak data sets are shown as defined in Eq. (8). In the figures it can be observed that the data no longer converges to zero but to the bias of the fit compared to the original data. In case of the wave data both exposure durations converge to a value close to 0.015 (approximately 1.5% of H_s). For the impact loads the bias varies depending on the duration and between the two repeat data sets. For the first repeat set the bias seems quite small for 15 h duration and is above 1000 kN. For Repeat 2 the bias of the 15 h duration data is lower, below 500 kN. The bias of the 30 h duration case of Repeat 2 is clearly higher than of Repeat 1. Thus, based on these results, it appears that the quality of the fit can vary also between repeat measurements. This means that variations of individual impact loads between repeat measurements have an effect on the quality of an applied fit. In this case only 3 parameter Weibull fits are applied. This effect may be different for different types of fits.

5.7. Comparison of the deck box and ferry results

Similar analyses and results are presented for greenwater impact loads on a sailing ferry in [1] (Part 1 of the present study). The results presented in Part 1 also include wave elevations and wave impact forces measured using similar instrumentation in long duration tests. However, there are also some important differences that need to be considered in the interpretation and comparison of the results:



(a) Ferry, $S = 0.092$ (van Essen et al. (2023)) (b) Deck box, $S = 0.067$ (fig. 5)

Fig. 12. DSR and DNR distributions of wave crest height around 1 h exposure duration for the ferry and deck box (van Essen et al. [1]).

Table 2
Wave conditions for the ferry and deck box. All: irregular, long-crested, JONSWAP, $\gamma = 3.3$. Wave steepness $S = H_s/T_p^2$, λ_p = peak wave length, x = (average) distance to the wave generator, NS = North Sea.

Campaign	Condition	Duration [h]	H_s [m]	T_p [s]	S [m/s^2]	λ_p [m]	x [m]	x/λ_p	# wave crests	# impacts
Ferry	Heavy	34:51:16	8.1	9.4	0.092	138	~3000	21.7	16 113	515
Ferry	Extreme	23:42:49	8.3	10.0	0.083	156	~3000	19.2	12 096	1001
Deck box	10,000 yr NS	300:00:00	17.0	15.9	0.067	395	1500	3.8	79 859	254

1. In the wave-in-deck impact loads on average only 2–3 impacts on the deck box were measured per 3 h seed, whereas for the greenwater loading on the sailing ferry were on average 44 (heavy condition) and 127 (extreme condition) green water impacts per 3 h seed were measured. Obviously, this makes a difference for the convergence of the impact forces, which is not only related to exposure duration, but also to the number of events in that duration. For this reason, the deck box impact results were presented for longer exposure durations: 60 h, 30 h and 15 h. The wave crests were evaluated for 3 h and 1 h in both case studies.
2. The steepness of the wave conditions was different (see also Table 2 and Fig. 12). Usually it is assumed that some wave crests start breaking around steepness values of 0.06–0.08 m/s^2. The deck box conditions were in the transition area from non-breaking to breaking waves, whereas the ferry conditions are above the breaking limit. There may therefore be many breaking crests in one deck box seed, and few in another. The waves for the ferry are above the breaking limit and breaking was frequently observed in all seeds. This is supported by the variation between DSRs in Fig. 12: there is less variation between the DSRs for the ferry than for the deck box. For the ferry, most distributions clearly show the downward trend expected for breaking waves in the tail of the distribution (observed by e.g., [21,26,27]). For the deck box data this is missing for some seeds and present for others. This could introduce more variability and slower convergence of waves with a steepness around the start of breaking.
3. The wave crests directly hit the deck box, whereas they first run over deck before impacting the ferry accommodation. Direct impacts increase the influence of the wave steepness, air entrapments and other complicated impact phenomena (see e.g., [3,28]). Steep almost breaking wave crests that are badly timed with respect to the structure may lead to another load regime for the deck box than the other waves, in line with a badly-behaving problem. This is less likely for the green water loads, which makes them easier to fit and extrapolate.
4. For the measurement of the crest heights different types of instrumentation were used in the two test campaigns. For the ferry test campaigns acoustic type wave probes were used. These types of wave probes are typically used for wave measurements at forward speed. For the deck box wave measurements resistant type wave probes were used. The two types of wave probes are based on different measurement techniques. Differences in measured crest heights can be observed especially in very steep and breaking waves, as presented in [29]. The type of measurement probe might therefore also have an effect on the variability of the crests.
5. The deck box tests were done at zero speed and the ferry tests at forward speed. In theory this should not matter, but the tests at forward speed were done further from the wave generator on average than the tests at zero speed (Table 2). As wave energy can only be inputted at the generators, this may have some influence (especially for steep near-breaking waves, see

Table 3
Number of seeds required for convergence of the wave crest MPM (original data) following the analysis in Section 4.7, for convergence criterion $R(N) = 3\%\ H_s$ and the experimental campaigns in part 1 and 2 of this study.

Exposure duration	Required N (P1, Heavy)	Required N (P1, Extreme)	Required N (P2, 10,000 yr NS)
30 min	23	15	–
1 h	17	8	22
3 h	6	5	14

e.g., [30]). The distance to the wavemaker × divided by the wave length λ_p is much lower for the deck box than for the average ferry distance (Table 2). Canard et al. [31] showed that higher-order non-linear wave effects converge slowly with distance to the wave generator (even for identical wave spectra at different locations), and that the wave non-linearity is converged around $x/\lambda_p = 20–30$. The non-linearity of the waves is therefore probably more converged at the average ferry distance than at the deck box distance. This may cause additional variability in the deck box wave crests.

6. The tests were done in different basins, with different wave generators and instrumentation. Both are position-controlled flap-type wave generators, controlled with similar software. The water depths of both basins can be considered deep at the tested scales. This is therefore not expected to introduce differences in the results.

For the purpose of comparison of the non-dimensional wave crest MPM convergence in the two data sets, the convergence criterion for the RMSE in Section 4.7 is set at 0.03 (so at 3% of the H_s values) for both test cases. Based on this criterion a required number of seeds can be derived from the convergence results. The results are presented in Table 3 for the two data sets. The table shows that waves in both case studies converge at a similar rate, although the deck box waves seem slightly slower. This can probably be attributed to points 2, 3, 4 and 5 above.

Because the deck box air-gap impacts are much rarer than the ferry green water impacts (point 1 above) and because there is no obvious reference force to relate the absolute forces in both cases, these results cannot be directly compared.

6. Conclusions

The main objective of the presented work is determining the minimum required number of seed variations for a reliable design load estimate for wave-in-deck type of impact loads. In an attempt to achieve this objective, wave-in-deck model tests have been carried out with 100 × 3 h realisations of a 10,000 year North Sea state. Convergence plots were made, showing the RMSE in MPM values as a function of the number of seeds (or realisations) for different exposure durations. These convergence plots can be used to obtain an idea of the expected error in the extreme values when a certain number of seeds is tested in an experiment, or vice versa, to determine how many seeds are required in order to obtain a target uncertainty. Based on this analysis it was concluded that the wave crests and the impact force statistics do not convergence equally quickly in the considered condition. This is of course related to the fact that not all wave crests result in a deck impact as a result of the deck clearance. For longer exposure durations, convergence is reached with fewer seeds. The total required test time however, may not be shorter.

Proper fitting of the distributions of individual seeds may reduce the variability of the extremes, as indicated by the fitting plots presented and discussed in Section 5.3, but this has a limited effect on the required number of seeds based on the convergence study and introduces a bias with respect to the measured data. The distributions and resulting fits of the two complete repeat data sets of the impact loads suggest that, if sufficient data is available, the variability of individual impact events in repeat measurements is not affecting the overall distribution and resulting 90% quantile level of the impact load significantly. The convergence based on the RMSE of the MPM at different exposure durations was similar for the two repeat data sets.

The same analysis was carried out for green water impacts on a sailing ferry presented in Part 1 of this study (see [1]). The required number of seeds for impact force peak MPM convergence cannot be directly compared, because the deck box impacts were much rarer than the green water impacts on the ferry (on average 1 versus 15–42 impacts per hour). Taking an RMSE criterion of 3% H_s for the wave crest MPM values results in 5–14 required 3 h seeds for both cases, or 8–22 required 1 h seeds. This is based on two test campaigns with in total three wave conditions in relatively steep waves. The difference between the wave crest MPM convergence of the two test campaigns is larger than between the two wave conditions in one campaign. This can have several possible causes, including differences in wave steepness, wave measurement equipment, average distance to the wave generator, or possibly other differences between the basins. The convergence of the green water force MPM values on the ferry presented here is representative for structures with a relatively high impact frequency, whereas the deck box results are more representative for structures with a low impact frequency.

7. Future work

The results presented in this paper suggest that a large amount of 3 h seed variations are required in order to obtain converged impact load distributions. In practice this is very time and cost consuming. Thus, more efficient ways to achieve proper estimation of the design impact loads need to be considered.

One approach would be to screen the undisturbed wave data for possible events causing an impact load, then only run the part of the time trace containing the extreme event for the impact load measurements. For a fixed structure such as the deck box presented in this paper, this is rather straight forward, as a crest needs to be high enough to reach the deck box. Thus, in this case a screening criterion could simply be the crest height vs deck height. However, for moving objects especially at forward speed, more aspects start to play a role, such as wave steepness, low frequency response of the floater, etc (see [32]). For the present data set, it would be very interesting to see if the same results for the distribution of impact loads can be achieved by extracting the extreme events from the undisturbed wave measurements and only rerunning short sequences containing these events for impact measurements. The resulting impact load distributions should then be reviewed also in view of how much measurement time can be saved by this approach.

In addition to the screening approach described above, numerical tools can also assist in the derivation of design loads. Numerical tools can either be used to assist in screening for extreme events or for determining impact loads for pre-selected wave events. In an ongoing research project the existing data set is simulated via CFD with the objective to see whether the results can be reproduced via CFD [33]. If the results can be reproduced this then may show the potential to use CFD for similar problems in the future.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The authors do not have permission to share data.

Acknowledgements

Part of this research received funding of the project ‘‘Multi-fidelity Probabilistic Design Framework for Complex Marine Structures’’ (project number TWM.BL.019.007) of the research programme ‘‘Topsector Water & Maritime: the Blue route’’ which is (partly) financed by the Dutch Research Council (NWO).

Appendix A. Theoretical wave crest distributions

The Rayleigh DNR wave crest distribution for Gaussian narrow-banded wave spectra is provided in Eq. (9), where H_s is the significant wave height and η is the wave crest height.

$$P_{Rayleigh}(\eta_c > \eta) = \exp\left(-\frac{8\eta^2}{H_s^2}\right) \tag{9}$$

The Forristall DNR wave crest distribution [20] for second-order waves is provided in Eq. (10a). It is based on the Weibull distribution, with α and β parameters that are based on fitting to second-order wave simulations. These parameters are provided for long-crested waves in Eqs. (10b) and (10c), where S_1 is a wave steepness parameter (Eq. (11)) and U_r is the Ursell number (Eq. (12), characterising the effect of water depth on wave non-linearity). T_1 in these formulations is the mean wave period, d the water depth and k_1 the corresponding wave number.

$$P_{Forristall}(\eta_c > \eta) = \exp\left(-\left(\frac{\eta}{\alpha H_s}\right)^{\beta}\right) \tag{10a}$$

$$\alpha = 0.3536 + 0.2892S_1 + 0.1060U_r \tag{10b}$$

$$\beta = 2 - 2.1597S_1 + 0.0968U_r^2 \tag{10c}$$

$$S_1 = \frac{2\pi H_s}{gT_1^2} \tag{11}$$

$$U_r = \frac{H_s}{k_1^2 d^3} \tag{12}$$

There are two versions of the Huang-Zhang crest wave crest distribution: for DSR limits (mean and upper and lower 99% bands) and for DNR [19]. For reference, DSR is called 'PDSR' in this publication, and DNR is called 'PDER' (the present publication sticks to DSR and DNR). The basic formulation for all these distributions is the Weibull formulation (provided in Eq. (13a)). The same definitions for wave steepness parameter and the Ursell number are still used. The values of coefficients a_0 to a_6 and b_0 to b_5 were based on non-linear wave simulations. They were defined separately for probabilities below and above 10^{-2}, see Table A.4. NB. A small mistake in the original publications was corrected: the a_4 value in Eq. (13b) was given to the power of four. This power was removed in order to arrive at realistic distributions.

$$P_{HuangZhang}(\eta_c > \eta) = \exp\left(-\left(\frac{\eta}{\alpha' H_s}\right)^{\beta'}\right) \tag{13a}$$

$$\alpha' = a_0 + a_1S_1 + a_2S_1^2 + a_3S_1^3 + a_4S_1^4 + a_5U_r + a_6U_r^2 \tag{13b}$$

$$\beta' = b_0 + b_1S_1 + b_2S_1^2 + b_3S_1^3 + b_4S_1^4 + b_5U_r^2 \tag{13c}$$

Table A.4
Coefficients Huang-Zhang distributions [19].

Type	DSR, mean	DSR, upp99%	DSR, low99%	DSR, mean	DSR, upp99%	DSR, low99%	DNR	DNR
P	$<10^{-2}$	$<10^{-2}$	$<10^{-2}$	$>10^{-2}$	$>10^{-2}$	$>10^{-2}$	$<10^{-2}$	$>10^{-2}$
a_0	0.2894	0.1334	0.3752	0.3712	0.3516	0.3708	0.3242	0.3733
a_1	12.3011	13.0432	9.1269	1.0087	2.3892	0.9835	11.7467	0.9398
a_2	−662.632	−751.0935	−446.8393	−43.0667	−105.4167	−39.9404	−652.147	−40.0095
a_3	12 153.3466	14 727.6571	7837.572	567.5292	1557.3914	598.3819	12 308.3001	512.0601
a_4	−68 031.8045	−87 711.081	−42 682.1177	−1173.1204	−5807.4512	−2144.492	−70 529.2504	−849.0734
a_5	0.3779	0.084	0.8351	0.1276	0.0791	0.2427	0.3785	0.1294
a_6	−3.7904	−5.2045	−5.7942	0.3115	0.355	−0.5379	−3.8837	0.2882
b_0	1.5277	0.7965	2.4637	2.006	1.62	2.212	1.7321	2.0411
b_1	67.2118	44.9229	82.5688	4.841	19.72	10.951	72.7179	3.6068
b_2	−3683.1338	−2525.6956	−4300.5362	−321.181	−909.665	−637.673	−4093.6916	−272.2806
b_3	63 759.0846	47 184.1324	68 857.4691	846.332	10 465.556	6707.056	72 132.2957	−58.9649
b_4	−336 712.3631	−270 084.1075	−337 914.7489	27 223.189	−22 952.443	−10 244.141	−386 504.7502	32 786.2976
b_5	−8.1382	−13.316	2.5661	0.832	−3.726	4.505	−9.9594	0.9003

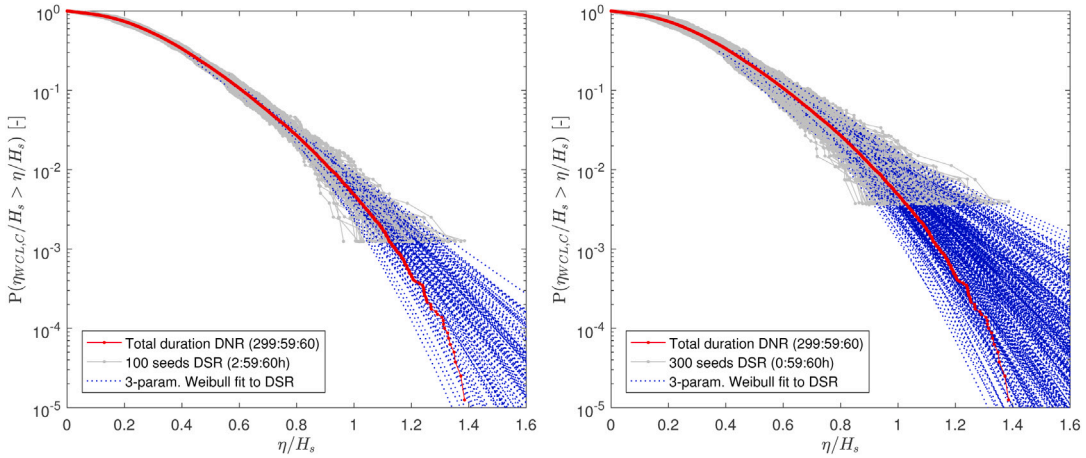
A.1. Comparison theoretical wave distributions with measurements

Based on the results presented in Fig. 5 compared to the theoretical crest height distributions, the following observations are made:

- The DNR distribution of the measured crest heights seem to be very close to the distribution based on the CresT correction as well as the Huang-Zhang DNR distribution. The measured crest heights are right between the two theoretical lines for most parts of the distribution.
- Only the crests with the lowest probabilities, starting from approximately 10^{-4} fall below those two theoretical distributions. This is typically observed with occurrence of wave breaking (see [6]).
- The range of the measured crest distributions of the individual seeds (DSR) falls with two exemptions right between the upper and lower limit of the Huang-Zhang DSR distribution (green area).
- When the measured wave signals are divided up in groups with shorter duration (1 h) the variation in crest heights at the lowest probability levels increases.

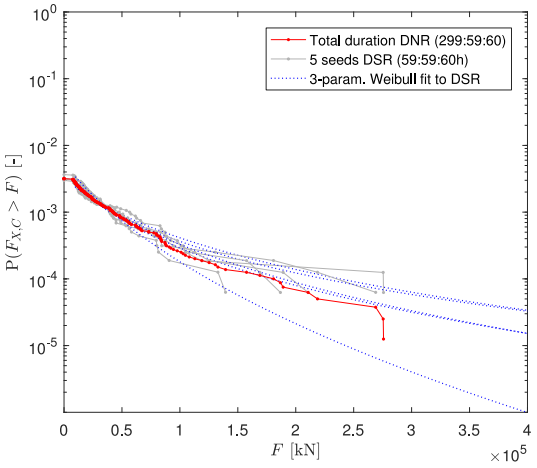
Appendix B. Weibull fit plots

Fig. B.13 provides all Weibull fit plots for the wave crests and Fig. B.14 provides the same plots for wave impact force peaks.

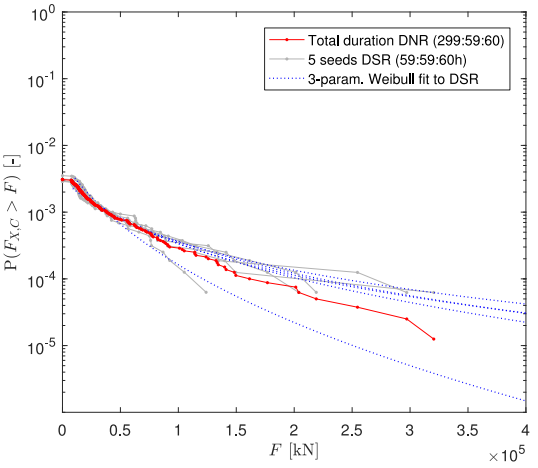


(a) Wave crests, 100 × 3h exposure duration (b) Wave crests, 300 × 3h exposure duration

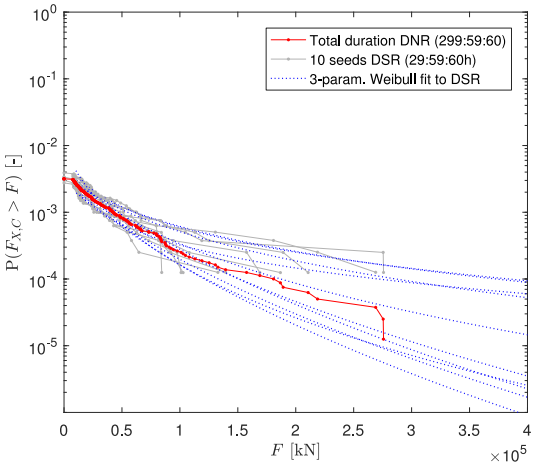
Fig. B.13. DSR and DNR distributions of $\eta_{WCL,C}$ for different exposure durations and number of seeds. Including 3-parameter Weibull fits to top 30% crests in the DSR distributions.



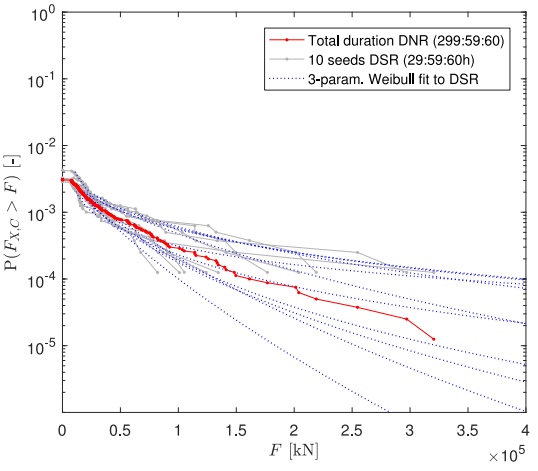
(a) Impact peaks, 5 × 60h exposure duration, Repeat 1



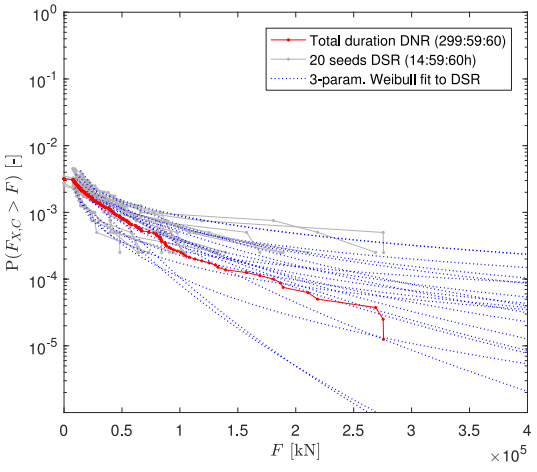
(b) Impact peaks, 5 × 60h exposure duration, Repeat 2



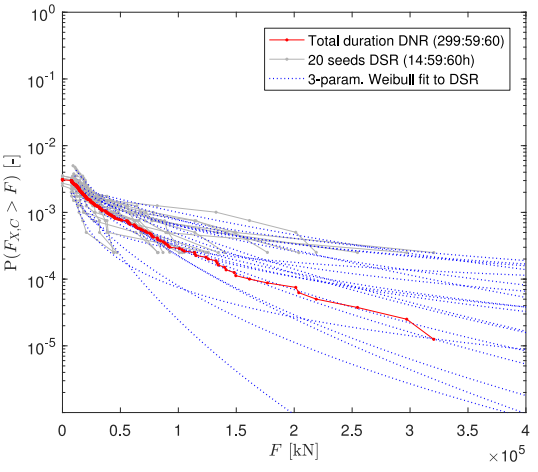
(c) Impact peaks, 10 × 30h exposure duration, Repeat 1



(d) Impact peaks, 10 × 30h exposure duration, Repeat 2



(e) Impact peaks, 20 × 15 h exposure duration, Repeat 1

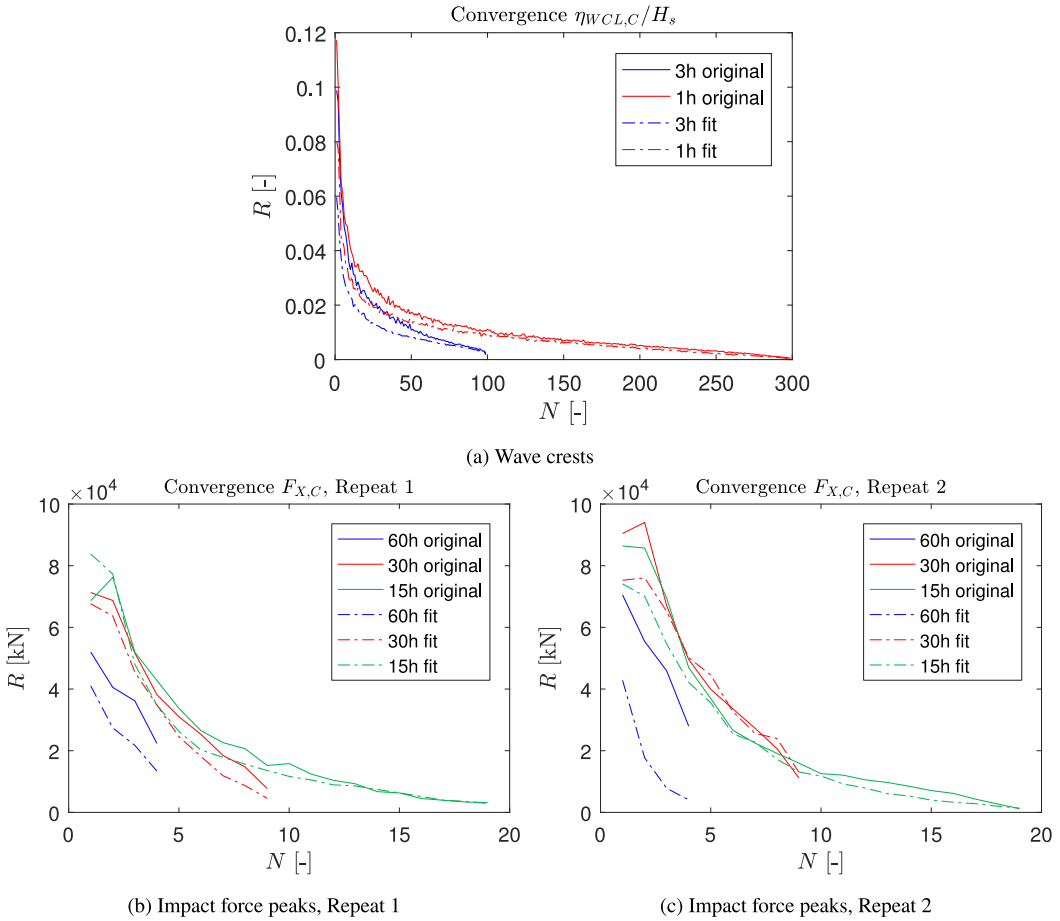


(f) Impact peaks, 20 × 15 h exposure duration, Repeat 2

Fig. B.14. DSR and DNR distributions of $F_{X,C}$ for different exposure durations and number of seeds. Including 3-parameter Weibull fits to top 90% peaks in the DSR distributions.

Appendix C. Re-ordered convergence and bias plots

In Section 5.5, the convergence plots are arranged to show the results of the original and fitted data separately. In Fig. C.15, the convergence results of the original and fitted data are shown in one figure for direct comparison. Similarly, Fig. C.16 show the results for the bias and convergence of the fits in the same plot.

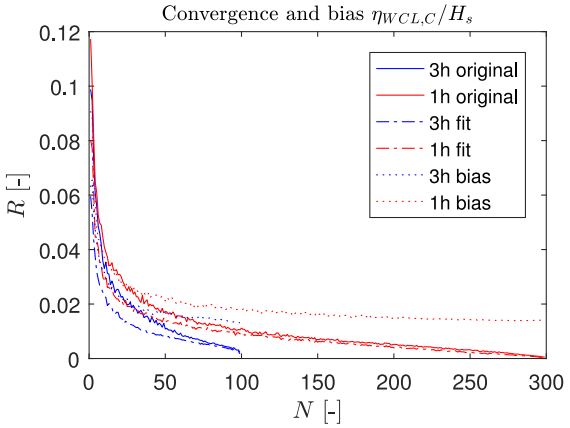


(a) Wave crests

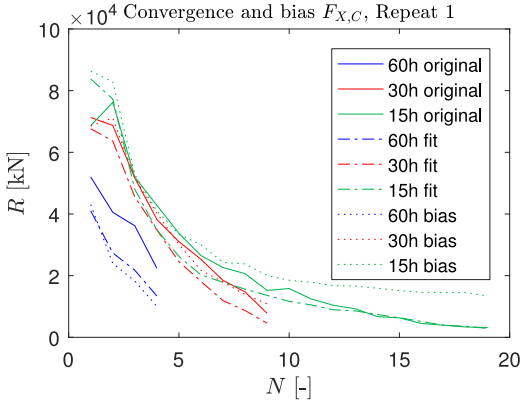
(b) Impact force peaks, Repeat 1

(c) Impact force peaks, Repeat 2

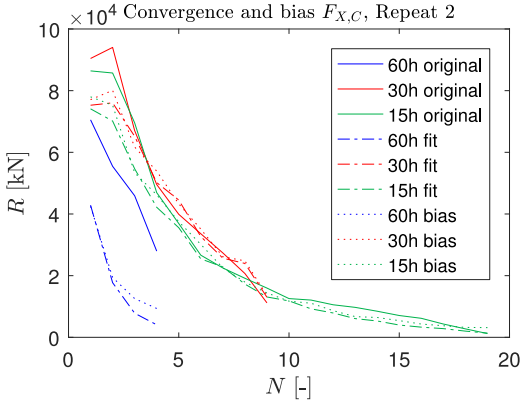
Fig. C.15. Convergence metric $R(N)$ for the MPM value for waves and impacts, original and fitted data. Mean of 500 random seed picking realisations.



(a) Wave crests



(b) Impact force peaks, Repeat 1



(c) Impact force peaks, Repeat 2

Fig. C.16. Convergence metric $R(N)$ and Bias metric $B(N)$ for the MPM value for waves and impacts, original and fitted data. Mean of 500 random seed picking realisations.

References

[1] van Essen SM, Scharnke J, Seyffert HC. Required test durations for converged short-term wave and impact extreme value statistics - Part 1: ferry dataset. Mar Struct 2023. http://dx.doi.org/10.1016/j.marstruc.2023.103410.

[2] Scharnke J, de Wilde J, Vestbøstad TM, Haver S. Wave-in-deck impact load measurements on a fixed platform deck. In: Proceedings of ASME 33rd international conference on ocean, offshore and arctic engineering. 2014, http://dx.doi.org/10.1115/OMAE2014-23180.

[3] Scharnke J. Elementary loading processes and scale effects involved in wave-in-deck type of loading - a summary of the BreaKin JIP. In: 38th Int. conf. ocean offshore arct. eng. (OMAE). Glasgow,UK; 2019, http://dx.doi.org/10.1115/OMAE2019-95004.

[4] Johannessen T, Haver S, Bunnik T, B. B. Extreme wave effects on deep water TLPs lessons learned from the snorre a model tests. In: Deep offshore technology conference, DOT2006. 2006.

[5] Buchner B, Forristall G, Ewans K, Christou M, Hennig J. New insights in extreme crest height distributions. In: Proceedings of OMAE 2011 30th international conference on offshore mechanics and arctic engineering. 2011, http://dx.doi.org/10.1115/OMAE2011-49846.

[6] Hennig J, Scharnke J, Swan C, Hagen Ø, Ewans K, Tromans P, Forristall G. Effect of short-crestedness on extreme wave impact - summary of findings from the Joint Industry Project 'ShorTCresT'. In: Proceedings of ASME 2015 34th international conference on ocean, offshore and arctic engineering. 2015, http://dx.doi.org/10.1115/OMAE2015-41167.

[7] Stansberg C, Baarholm R, Fokk T, Gudmestad O, Haver S. Wave amplification and possible deck impact on gravity based structure in 10^{-4} probability extreme crest heights. In: Proceedings of OMAE04 23rd internation conference of offshore mechanics and arctic engineering. 2004, http://dx.doi.org/10.1115/OMAE2004-51506.

[8] Baarholm R. A simple numerical method for evaluation of water impact loads on decks of large-volume offshore platforms. In: Proceedings of OMAE2005 24th internation conference of offshore mechanics and arctic engineering. 2005, http://dx.doi.org/10.1115/OMAE2005-67097.

[9] Baarholm R. Experimental and theoretical study of three-dimensional effects on vertical wave-in-deck forces. In: Proceedings of ASME 2009 28th internation conference of offshore mechanics and arctic engineering. 2009, http://dx.doi.org/10.1115/OMAE2009-79560.

[10] DNV-GL. Offshore technical guidance, prediction of air gap for column stabilised units, DNVGL-OTG-13. Oslo, Norway: Det Norske Veritas; 2019.

[11] DNV-GL. Offshore technical guidance, horizontal wave impact loads for column stabilised units, DNVGL-OTG-14. Oslo, Norway: Det Norske Veritas; 2019.

[12] Lian G. Slamming loads on large volume structures from breaking waves (Ph.D. thesis), Stavanger, Norway: University of Stavanger; 2018.

[13] Haver S. Freak waves: a suggested definition and possible consequences for marine structures. In: Work. rogue waves. Brest, France: Ifremer; 2004, URL: http://www.ifremer.fr/web-com/stw2004/rw/fullpapers/haver.pdf.

[14] DNV. Det norske veritas, recommended practice DNV-RP-C205: Environmental conditions and environmental loads. Oslo, Norway: Det Norske Veritas; 2019.

[15] NORSOK. Standard N-003 actions and action effects. Standards Norway; 2007.

[16] DNV. Det norske veritas, class guideline DNVGL-CG-0130: Wave loads. Oslo, Norway: Det Norske Veritas; 2018, URL: https://rules.dnv.com/docs/pdf/DNV/CG/2018-01/DNVGL-CG-0130.pdf.

[17] van de Bunt E, Dekker J, Scharnke J, Jaouen F. Applying force panels for wave impact measurements. Ocean Eng 2021;232.

[18] Ochi M. Applied probability & stochastic processes in engineering & physical sciences. Wiley series in probability and mathematical sciences, 1990.

[19] Huang Z, Zhang Y. Semi-empirical single realization and ensemble crest distributions of long-crest nonlinear waves. In: Proc. of 37th OMAE conf.. Madrid, Spain: ASME; 2018, http://dx.doi.org/10.1115/OMAE2018-78192.

[20] Forristall GZ. Wave crest distributions: Observations and second-order theory. Phys Oceanogr 2000;30(8):1931–43. http://dx.doi.org/10.1175/1520-0485(2000)030<1931:WCDOAS>2.0.CO;2.

[21] Buchner B, Forristall GZ, Ewans K, Christou M, Hennig J. New insights in extreme crest height distributions (a summary of the 'CresT' JIP). In: 30th Int. conf. ocean offshore arct. eng. (OMAE). Rotterdam, The Netherlands: ASME; 2011, http://dx.doi.org/10.1115/OMAE2011-49846.

[22] ITTC. Recommended practice 7.5-02-07-02.6: global loads seakeeping procedure. International Towing Tank Conference; 2017, URL: https://www.ittc.info/media/8109/75-02-07-026.pdf.

[23] Haver S. Metocean modelling and prediction of extremes. Stavanger, Norway: May, Haver & havet, University in Stavanger, NTNU; 2017, p. 1–256.

[24] Mauro F, Braidotti L, la Monaca U, Nabergoj R. Extreme loads determination on complex slender structures. Int Shipbuild Prog 2019;66:57–76. http://dx.doi.org/10.3233/ISP-180256.

[25] Johannessen TB, Hagen O. Estimating design levels for strongly nonlinear response. In: Proceedings of ASME 2012 31st international conference on ocean, offshore and arctic engineering. 2012, http://dx.doi.org/10.1115/OMAE2012-83947.

[26] Latheef M, Swan C. A laboratory study of wave crest statistics and the role of directional spreading. Proc R Soc A Math Phys Eng Sci 2013;469(2152). http://dx.doi.org/10.1098/rspa.2012.0696.

[27] Hennig J, Scharnke J, Swan C, Hagen Ø, Ewans K, Tromans PS, Forristall GZ. Effect of short-crestedness on extreme wave impact - a summary of findings from the joint industry project "ShorTCresT". In: 34th Int. conf. ocean offshore arct. eng. (OMAE). St. John's, Canada: ASME; 2015, http://dx.doi.org/10.1115/OMAE2015-41167.

[28] Bogaert H. An experimental investigation of sloshing impact physics in membrane LNG tanks on floating structures (Ph.D. thesis), Delft, The Netherlands: Delft University of Technology; 2018, http://dx.doi.org/10.4233/uuid:96870b88-e07b-4ec2-8bd4-ef2cd3713568.

[29] Hennig J, Scharnke J, Schmittner C, van den Berg J. ShorTCresT: Directional wave measurements at MARIN. In: Proceedings of ASME 2015 34th international conference on ocean, offshore and arctic engineering. 2015, http://dx.doi.org/10.1115/OMAE2015-41169.

[30] van Essen SM. Influence of wave variability on ship response during deterministically repeated seakeeping tests at forward speed. In: Okada T, Suzuki K, Kawamura Y, editors. 14th int. symp. practical design of ships and other floating structures (PRADS2019). Lecture notes civil eng, 63, Yokohama, Japan: Springer Nature Singapore; 2021, p. 899–925. http://dx.doi.org/10.1007/978-981-15-4624-2_54.

[31] Canard M, Ducrozet G, Bouscasse B. Varying ocean wave statistics emerging from a single energy spectrum in an experimental wave tank. Ocean Eng 2022;246:110375. http://dx.doi.org/10.1016/j.oceaneng.2021.110375.

[32] Bunnik THJ, Stansberg CT, Pákozdi C, Fouques S, Somers L. Useful indicators for screening of sea states for wave impacts on fixed and floating platforms. In: 37th Int. conf. ocean offshore arct. eng. (OMAE). Madrid, Spain: ASME; 2018, http://dx.doi.org/10.1115/OMAE2018-78544.

[33] Bunnik T, Scharnke J. Statistical variation of the 3-h maximum crest height in a survival sea state. In: Proceedings of ASME 2022 41st international conference on ocean, offshore and arctic engineering. 2022, http://dx.doi.org/10.1115/OMAE2022-79045.