Applied Ocean Research 108 (2021) 102402

Contents lists available at ScienceDirect

Applied Ocean Research

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

# On the physical constraints for the exceeding probability of deep water rogue waves

S. Mendes [a,b], A. Scotti [a], P. Stansell [c]

[a] Department of Marine Sciences, University of North Carolina at Chapel Hill, Chapel Hill, NC, 27599 USA
[b] Group of Applied Physics and Institute for Environmental Sciences, University of Geneva, Blvd. Carl-Vogt 66, Geneva, Switzerland
[c] Independent scholar, Edinburgh, UK

ARTICLE INFO

*Keywords:*
Rogue wave
Exceeding probability distribution
Water wave
Storms
Nonlinearity
Physical bounds

ABSTRACT

Nearly four decades have elapsed since the first efforts to obtain a realistic narrow-banded model for extreme wave crests and heights were made, resulting in a couple dozen different exceeding probability distributions. These models reflect results of numerical simulations and storm records measured off of oil platforms, buoys and more recently satellite data. Nevertheless, no consensus has been achieved in either deterministic or operational approaches. Moreover, a minor issue with the established distributions is that they are not bounded by more than one physical limit while others are not bounded at all. Though the literature is rich in physical bounds for single waves, here we describe physical limits for the ensemble of waves that have not yet been addressed. As previous studies have shown, the exceeding probability distribution does not depend unequivocally on one sea state parameter, thus, this work supplies a combination of sea state parameters that provide guidance on the sea state influence on rogue wave occurrence. Based on specific bounds, we conjecture the dependence of the expected maximum of normalized wave heights (also known as abnormality index) and crests on the aforementioned sea-state parameters instead of the total number of waves in the wave record. Finally, we introduce a new dimensionless parameter that is capable of explaining the uneven distribution of rogue waves in the different storms pointed out by Stansell [74].

## 1. Introduction

Since the first attempt to give an expectation for the tallest waves in a narrow-banded sea by Longuet-Higgins [48], several other distributions have been proposed in an attempt to improve the prediction of extreme waves in rough seas (see Dysthe et al. [17], Pelinovsky and Kharif [65] for a review). The first articles on the topic were mainly descriptive [15,16,42,53], but starting with the precise measurement of the Draupner wave [34,35,87] the literature became more technical. The drive to develop better tools to predict the occurrence of extreme waves goes beyond the academic realm, since rogue waves are believed to be responsible for the majority of ship losses associated with unknown causes [19,85]. Moreover, offshore operations have relied on a statistical description of the environmental conditions to characterize the operating envelope. In the years since, engineers and oceanographers have tried, with varying degrees of success, to predict extreme wave heights for a wide range of sea states [24,33,77]. Although the sea state can be well described by the hindcast of weather records, a proper study of long-term wave height and crest heights probabilities require integration over a distribution of sea states, which is not quite feasible. Nevertheless, the main features of the operating envelope (as specified by the design wave) are obtained from the Longuet-Higgins [48], which in turn is contrasted with wave spectra from ocean states that do not assume narrow-banded behavior, making the design wave incompatible with realistic sea states. One essential problem is the lack of a definition of a suitable sample size for the measurements of sea surface elevation. Typically, distributions found in the literature analyze a very large set of waves with large variations in sea-state parameters while neglecting homogeneous smaller samples. One possible consequence of this is the apparent disagreement between several studies regarding the prediction of rogue wave occurrence: some studies report less rogue wave heights (not to be confused with the number of wave records) [e.g. [11,13,14,24,27,30,33,47,67]] while others report more rogue waves [57,72,73] and super-rogue waves [46,63] or the same statistics [18,59,94] predicted by Longuet-Higgins [48], sometimes a combination of the three in the very same study [9,52,74]. Therefore, this study focuses on the implementation of physical constraints for a target exceeding probability of both wave heights and crest heights.

*E-mail addresses:* saulo.dasilvamendes@unige.ch (S. Mendes), ascotti@unc.edu (A. Scotti), paulstansell@gmail.com (P. Stansell).

https://doi.org/10.1016/j.apor.2020.102402
Received 16 April 2020; Received in revised form 22 August 2020; Accepted 5 October 2020
Available online 4 February 2021

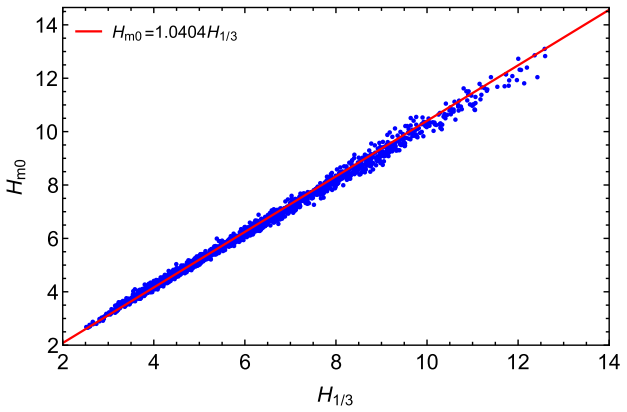


**Fig. 1.** Scatter plot of $H_{1/3}$ versus $H_{m0}$ (in meters). Each dot shows the height variables measured from each 20-min record of each storm of Stansell [74]. For the range $H_{1/3} > 6$m we find that the ratio among these two definitions obeys $0.893 < H_{1/3}/H_{m0} < 1.033$.

Intuitively, rogue waves are waves taller and steeper than expected by standard model predictions [41]. Unfortunately, this is a rather vague formulation of the phenomenon and there is no precise theoretical definition based on wave spectra or dynamical principles. Dean [12] was one of the first to propose a statistical definition for rogue waves as any wave that is at least twice as large as the significant wave height of the sea state in which it occurs. Despite the arbitrariness of the factor of two, such definition has been widely accepted [35] or slightly modified [66], and therefore, we shall adhere to it here. In order to separate the targeted group of tall waves from the ordinary, we define the normalized wave height $\alpha$ as,

$$\alpha \equiv \frac{H}{H_{1/3}} \equiv \frac{\mathscr{Z}_c + \mathscr{Z}_t}{H_{1/3}}, \tag{1}$$

where $H$ is any individual wave height, $\mathscr{Z}_c$ and $\mathscr{Z}_t$ are the crest height and trough depth, respectively, and $H_{1/3}$ is the measured significant wave height, defined as the mean height among the 1/3 highest waves in the time series. $H_{1/3}$ is distinct from the spectral counterpart, $H_{m0}$, that is obtained from the spectrum (see Appendix B and Fig. 1). The return period, $\mathscr{N}_\alpha$, is defined as is the inverse of the exceeding probability, and the total number of rogue waves that can be observed in a series of $N$ waves is the product of $N$ and the exceeding probability. In terms of normalized wave heights, the exceeding probability proposed by Longuet-Higgins [48] for narrow-banded seas, e.g. the Rayleigh distribution (henceforth abbreviated as RD), is expressed as (see Cartwright and Longuet-Higgins [8]):

$$\mathscr{R}_\alpha \equiv \mathscr{R}\left(H > \alpha H_{1/3}\right) = e^{-2\alpha^2}. \tag{2}$$

## 2. Observational data

In this study we use data collected from three infrared laser altimeters mounted on three corners of the North Alwyn oil platform. The platform is situated in the northern North Sea in a water depth of approximately 130 m. The logging system was configured with a 5 Hz logging frequency. Every 20 minutes, the logging apparatus calculated $H_{m0}$. If $H_{m0}$ exceeded 3 m all th 20-min sea surface records were saved to optical disk for further analysis. Table 1 summarizes the data. In this dataset a storm is defined as the period between the start of the first record and the end of the last of a contiguous sequence of 20-min records each satisfying $H_{m0} > 3$ m. The raw data were stored as 2381 20-min records of surface elevation measurements. For ease of discussion, we divide waves in to the five groups listed in Table 2. This is inspired by the work of Liu and MacHutchon [46] who divided extreme waves into "typical" rogue waves when $2 < \alpha < 4$ and "uncommon" rogue waves whenever $\alpha > 4$ whilst Didenkulova and Anderson [13] suggested that the latter group should be classified as "abnormal" rogue waves when $\alpha > 3$.

While most studies focus on the Longuet-Higgins's prediction for $\alpha > 2$, Table 1 surprisingly shows Longuet-Higgins's imprecision for sub-rogue waves, producing hundreds of false positives (see more inconsistencies in Fig. 2). Moreover, the data shows that there are plenty of dangerously tall waves ($H > 15$ m) even among sub-rogue waves (Fig. 3). Interestingly, Table 1 shows the first four storms having rogue wave return period, $\mathscr{N}_{\alpha=2}$, lower than that predicted by Longuet-Higgins [48] whereas the remaining storms show equal or longer return periods.

**Table 2**
Alternative definitions for the upper, middle and lower tail of the distribution of ocean waves and the total counts and percentage of waves. While rogue waves can also be understood as the maximum wave within a storm regardless of the threshold $\alpha$, in this study we refer to any wave that exceeds the above thresholds.

| | Ordinary Wave | Large Wave | Sub-Rogue Wave | Rogue Wave | Super-Rogue Wave |
|---|---|---|---|---|---|
| **Definition** | $\alpha \in (0, 1]$ | $\alpha \in (1, 1.75]$ | $\alpha \in (1.75, 2]$ | $\alpha \in (2, 3]$ | $\alpha > 3$ |
| **Total Count** | 304,371 **(86%)** | 48,771 **(13.8%)** | 480 **(0.14%)** | 105 **(0.03%)** | 1 **(0.0003%)** |

**Table 1**
Summary of the data from [74]. The numbers in the main largest section of the table are counts of numbers of waves for each storm. The storms are presented in ascending order of return period $\mathscr{N}_{\alpha=2}$. $\|\alpha\|$ denotes a storm's maximum observed normalized wave height. Numbers in bold font show Longuet-Higgins's prediction for the total number of rogue waves, maximum $\alpha$ and return period.

| Storm ID | $\alpha > 0$ | $\alpha > 1.75$ | $\alpha > 2$ | $\alpha > 2.25$ | $\alpha > 2.5$ | $\alpha > 3.0$ | $\|\alpha\|$ | $\mathscr{N}_{\alpha=2}$ |
|---|---|---|---|---|---|---|---|---|
| 29 | 13,610 | 30 | 12 | 1 | 0 | 0 | 2.30 | 1134 |
| 149 | 52,766 | 95 | 26 | 6 | 1 | 0 | 2.50 | 2029 |
| 90 | 44,867 | 111 | 20 | 3 | 1 | 0 | 2.65 | 2243 |
| 172 | 23,591 | 54 | 9 | 2 | 1 | 1 | 3.19 | 2621 |
| 132 | 45,056 | 68 | 14 | 3 | 0 | 0 | 2.30 | 3218 |
| 28 | 22,155 | 30 | 6 | 1 | 0 | 0 | 2.38 | 3693 |
| 146 | 15,109 | 20 | 4 | 1 | 0 | 0 | 2.46 | 3777 |
| 23 | 25,068 | 42 | 5 | 0 | 0 | 0 | 2.08 | 5014 |
| 26 | 27,774 | 33 | 4 | 0 | 0 | 0 | 2.16 | 6944 |
| 127 | 14,845 | 15 | 2 | 0 | 0 | 0 | 2.09 | 7423 |
| 25 | 16,896 | 21 | 2 | 1 | 1 | 0 | 2.59 | 8448 |
| 27 | 20,379 | 23 | 1 | 1 | 0 | 0 | 2.40 | 20,379 |
| 124 | 21,737 | 30 | 0 | 0 | 0 | 0 | 1.97 | — |
| 195 | 9875 | 9 | 0 | 0 | 0 | 0 | 1.95 | — |
| Total | 353,728 | 581 | 105 | 19 | 4 | 1 | 3.19 | 3369 |
| Total × $\mathscr{R}_\alpha$ | — | **774** | **119** | **14** | **1** | **0.005** | **2.58** | **2,981** |

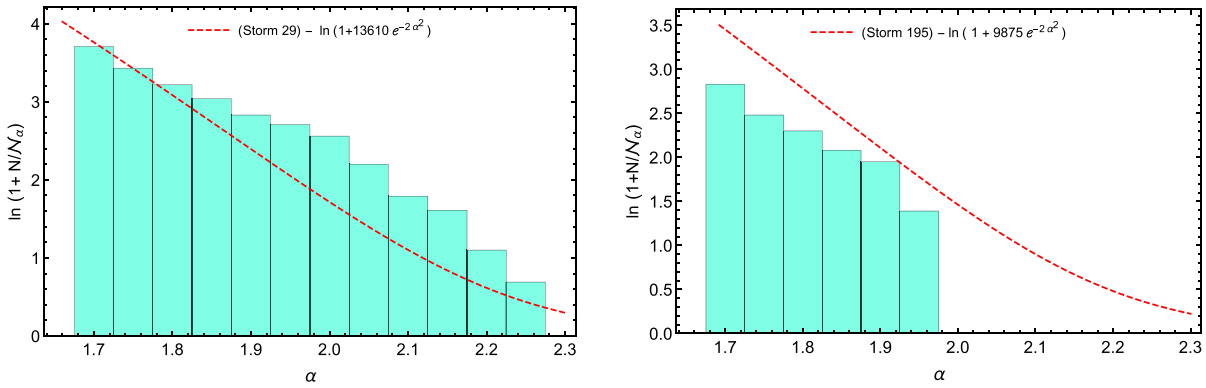


**Fig. 2.** Histogram of the number of waves in natural logarithm scale as a function of normalized wave height: we compare the observed number of waves for each threshold $\alpha$ (light green bars) and Longuet-Higgins's prediction (dashed red lines) for the storms with greatest (finite) and least return periods on Table 1. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

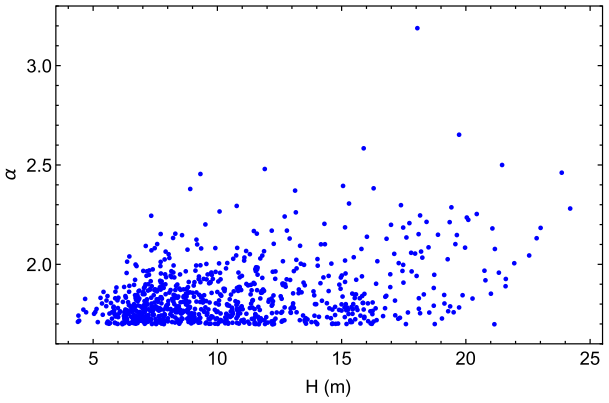


**Fig. 3.** Scatter plot of normalized height against wave height in meters.

Combining all storms into a single dataset demonstrates that the prediction accuracy of $\mathscr{R}_\alpha$ quickly decreases (either underpredicting or overpredicting) away from its center at $H > 2H_{1/3}$. The RD overpredicts the number of $H > 1.5H_{1/3}$ waves by 25% while it underpredicts the number of $H > 2.5H_{1/3}$ waves by 75%. Clearly, the inaccuracy grows faster for larger rogue wave heights. However, that is only if we consider the data set as a whole. Considering individual storms, the inaccuracy is even greater. For instance, for small clusters such as for one single storm, Longuet-Higgins [48] can either underpredict, overpredict or be equal to the observed number of waves with varying $\alpha$. This inserts a bias into the conclusions of any study [26].

The main argument against the standard approach is that a large sample of waves can be drawn from storms with very different metocean conditions and therefore can cover an enormous variability in its sea-state parameters. Using this approach can lead to distributions that do not account for the combination of sea-state parameters. Therefore, conclusions based on this procedure on the rogue wave occurrence most likely leads to contradictory claims that Longuet-Higgins [48] overpredicts, underpredicts or accurately predicts it. Therefore, small data sets tend to show uneven rogue wave occurrence while it approaches an average as the number of waves, $N$, becomes very large.

Unfortunately, most known distributions will also fail to explain single storm data due to their lack of flexibility with changing sea states, even with a small variability of sea-state parameters (see Table 3). Whilst a clear definition on the suitable time series length for accurate modelling is yet to be provided, it is not prudent to analyze only the full set of records while neglecting the distinctiveness of individual storms.

## 3. Standard distributions

Several authors have proposed models to address the shortcomings of $\mathscr{R}_\alpha$ in predicting the occurrence of rogue waves. One of the most common theoretical distributions used for the prediction of extreme waves was derived by Tayfun [77], who applied Stokes's second-order theory for the water surface elevation to calculate the exceeding probability as (see Appendix A and Appendix B):

$$\mathscr{T}_\alpha \equiv \mathscr{T}\left(H > \alpha H_{1/3}\right) = \exp\left[-\frac{8}{\breve{s}^2}\left(\sqrt{1+\alpha\breve{s}}-1\right)^2\right], \tag{3}$$

where $\breve{s} = 2\pi H_{1/3}/\lambda_1$ is the significant steepness and $\lambda_1$ is the wavelength obtained from the mean wave period as defined by eqs. (C.2) and (B.4). Naturally, as the Stokes model recovers the Airy solution for small steepness, the above distribution does recover RD in the limit of $\breve{s}\to 0$. The steepness of individual waves is limited by the Miche limiting condition [32,56]

$$\frac{H_i}{\lambda_i} \leqslant \frac{1}{7}\tanh\left(\frac{2\pi D}{\lambda_i}\right), \tag{4}$$

where $H_i$ is the $i$-th individual wave height and $\lambda_i$ its wavelength. Taking Miche's limit on steepness into account, which can be converted to its significant counterpart $H_{1/3} \leq \lambda_1/7$ [11,79], the probability should approach zero as $\breve{s}$ approaches 0.9 from below. Alas, $\mathscr{T}_\alpha$ breaks this limit by assigning a quite high probability of finding very steep waves (Fig. 4, red curve). This can be clearly observed by the fact that $\mathscr{T}_\alpha$ predicts that 1% of all waves in Table 1 will simultaneously exceed $\alpha > 2$ and $\breve{s} > 2\pi/7$ (or equivalently $s_{s,j} > 22/105$, see eq. (12)) which amounts to nearly 3500 rogue waves, however, there are no 20-min records with such a high significant steepness (see Fig. 14). Though presenting a theoretical improvement, $\mathscr{T}_\alpha$ is unable to produce sub-Rayleigh regime (i.e. probabilities lower than calculated by $\mathscr{R}_\alpha$), hence not being capable of explaining the North Alwyn observations.

Another widely used model is due to Haring et al. [33]. Unlike the previous one, it includes the water depth. Converted into crest-to-trough heights (to be consistent with eq. (3)), Haring et al. exceeding probability is given by,

$$\mathscr{H}_\alpha = \exp\left[-2\alpha^2\left(1-1.24\,\epsilon\alpha+1.09\,\epsilon^2\alpha^2\right)\right], \epsilon \equiv \frac{H_{1/3}}{D}, \tag{5}$$

where $D$ is the water depth and $\epsilon$ the *significant height-to-depth* parameter. The model is based on a previous version [36] that attempted to fit the data from Hurricane Camille in 1969. Given that the typical depth-limited breaking is of $H \leq 0.8D$ [32] and considering that $\|H\| > H_{1/3}$, we conclude that the limit distribution should approach zero in the limit $\epsilon\to 0.8^-$. However, $\mathscr{H}_\alpha$ shows some degree of flexibility

**Table 3**
North Alwyn's major sea state parameters (see Appendix B and Appendix C for nomenclature). The *nonlinearity* $\eta_{1/3}$ is defined in eq. (11), the spectral bandwidth $\nu$ in eq. (B.2), $\lambda_{1/3}$ and $H_{1/3}$ are the wavelength and wave height of the 1/3 tallest waves in meters and $T_z$ is the zero-crossing period in seconds.

| Storm ID | $H_{1/3}$ | $\langle H_{1/3}\rangle$ | $\langle T_z\rangle$ | $\lambda_{1/3}$ | $\langle\lambda_{1/3}\rangle$ | $\eta_{1/3}$ | $\langle\nu\rangle$ | $S_1^{-1}$ | $10^3\times$ Ur | $U_{00}$ | $U_1$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 29 | 5.07 | 5.85 | 8.06 | 147.9 | 157.8 | 1.234 | 0.740 | 27.02 | 1.302 | 1.855 | 1.933 |
| 149 | 6.68 | 6.76 | 8.90 | 191.4 | 185.6 | 1.241 | 0.779 | 28.88 | 2.081 | 1.863 | 1.937 |
| 90 | 4.91 | 5.85 | 8.02 | 141.3 | 150.9 | 1.245 | 0.626 | 24.18 | 1.190 | 1.840 | 1.925 |
| 172 | 6.02 | 5.78 | 8.13 | 174.9 | 152.2 | 1.228 | 0.797 | 24.26 | 1.196 | 1.841 | 1.926 |
| 132 | 4.51 | 4.70 | 7.43 | 147.8 | 143.4 | 1.151 | 0.874 | 32.50 | 0.864 | 1.879 | 1.944 |
| 28 | 4.57 | 5.67 | 8.01 | 146.0 | 154.3 | 1.181 | 0.654 | 25.26 | 1.206 | 1.846 | 1.928 |
| 146 | 4.01 | 4.01 | 7.07 | 167.1 | 154.8 | 1.109 | 0.936 | 36.55 | 0.859 | 1.891 | 1.951 |
| 23 | 6.50 | 6.69 | 8.55 | 190.4 | 192.9 | 1.192 | 0.696 | 25.55 | 2.224 | 1.847 | 1.929 |
| 26 | 3.70 | 3.77 | 6.75 | 139.3 | 126.6 | 1.119 | 0.847 | 32.46 | 0.540 | 1.879 | 1.945 |
| 127 | 4.94 | 5.28 | 8.76 | 177.5 | 174.6 | 1.127 | 0.487 | 28.09 | 1.438 | 1.860 | 1.936 |
| 25 | 5.57 | 5.68 | 7.82 | 166.0 | 165.9 | 1.167 | 0.770 | 26.97 | 1.397 | 1.855 | 1.933 |
| 27 | 5.50 | 5.35 | 8.08 | 172.6 | 166.0 | 1.153 | 0.758 | 30.19 | 1.317 | 1.870 | 1.940 |
| 124 | 7.97 | 7.53 | 9.56 | 214.3 | 198.7 | 1.178 | 0.466 | 23.08 | 2.656 | 1.832 | 1.921 |
| 195 | 6.44 | 6.93 | 8.80 | 203.0 | 202.3 | 1.163 | 0.774 | 27.94 | 2.534 | 1.859 | 1.935 |
| mean | 5.46 | 5.71 | 8.14 | 170.0 | 166.2 | 1.178 | 0.729 | 28.06 | 1.486 | 1.858 | 1.934 |

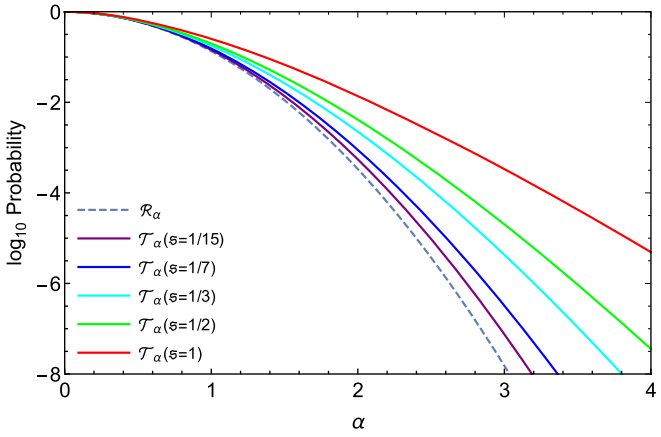


**Fig. 4.** $\mathcal{T}_\alpha$ compared with $\mathcal{R}_\alpha$ for different values of $\alpha$ and $\breve{s}$.

and varies around $\mathcal{R}_\alpha$ only for very high values of $\epsilon$ (Fig. 5) that are not allowed, such that $\mathcal{H}_\alpha$ can not explain the uneven spread of the North Alwyn dataset either.

The last distribution that we consider here is a modification of a distribution originally obtained by Forristall [24], based on hurricane data in the Gulf of Mexico:

$$\mathcal{T}_\alpha \equiv \mathcal{T}\left(H > \alpha H_{1/3}\right) = \exp\left[-2.263\,\alpha^{2.126}\right]. \tag{6}$$

In common with $\mathcal{R}_\alpha$, $\mathcal{T}_\alpha$ does not depend on any physical variable. Analyzing second-order numerical simulations that reproduced measurements and observations, Forristall [25] introduced a modified version that includes both depth and nonlinear effects for wave crests (converted to wave heights):

$$\mathcal{T}^*_\alpha = \exp\left[-\left(\alpha/2U_0\right)^{U_1}\right] \equiv \exp\left[-U_{00}\,\alpha^{U_1}\right], \tag{7}$$

with coefficients set to

$$\begin{aligned} U_0 &= 0.3536 + 0.2568S_1 + 0.08\mathrm{Ur}, \\ U_1 &= 2 - 1.7912S_1 - 0.5302\mathrm{Ur} + 0.284\mathrm{Ur}^2, \end{aligned} \tag{8}$$

with $S_1$ being the steepness and *Ur* the Ursell number (see Appendix B for the definitions of $T_1$, $k_1$ and $\lambda_1$):

$$S_1 \equiv \frac{2\pi H_{1/3}}{gT_1^2} = \frac{H_{1/3}}{\lambda_1} \quad ; \quad \mathrm{Ur} = \frac{H_{1/3}\lambda_1^2}{4\pi^2 D^3}. \tag{9}$$

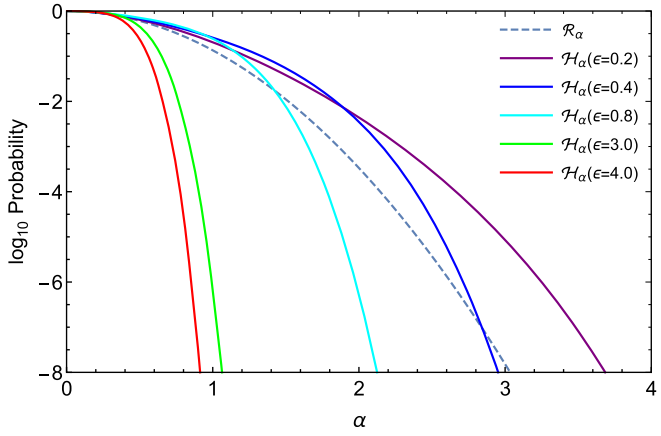


**Fig. 5.** $\mathcal{H}_\alpha$ compared with $\mathcal{R}_\alpha$ for different values of $\alpha$ and $\epsilon$.

As shown in Fig. 6, both of Forristall's distributions demonstrate a lack of flexibility in comparison to $\mathcal{R}_\alpha$ in deep water. Though $\mathcal{T}^*_\alpha$ generates higher variability than $\mathcal{T}_\alpha$, it assigns high probabilities to waves past the breaking Miche's limit. Moreover, these distributions are incapable of producing both sub-Rayleigh and super-Rayleigh regimes in deep water due to the very small variability in the coefficients $U_{00}$ and $U_1$, thus not being able to explain the North Alwyn data. Interestingly, Forristall [25] shows a complete different picture in shallow water (see Fig. 7): while the distribution can cross the Rayleigh regime, rogue waves become as likely as ordinary waves, e.g. the distribution asymptotically becomes a step function with a very large Ursell number, with a rogue wave chance leaping from one in every 2980 waves (RD) to one in every 60 ($S_1 = 0.1, \mathrm{Ur} = 10$). This anomalous behavior (Fig. 7) is due to intrinsic model limitations [80], so that variables must obey the relations (see eq.()):

$$S_1 < \frac{1}{7}, \quad \mathrm{Ur} < \frac{10\sqrt{m_0}}{D} \approx \frac{5}{2}\epsilon. \tag{10}$$

Applying the bound of eq. (10) into the definitions in eq. (9) we find that $\mathcal{T}^*_\alpha$ is limited to $\lambda_1 \leqslant \pi\sqrt{10}D \approx 10D$, thus not being a relevant model for shallow water waves either. This shortcoming highlights how the mathematical modelling of a narrow physical range can yield strange results if the limiting conditions of the sea state variables are not properly bounded. Nevertheless, [25] is regarded as one of the most useful distributions for design purposes. For instance, Gibson et al. [27] reports that the crest height version of eq. (7) is commonly employed in the industry, due to its dependency on variables that are readily available from the hindcast of sea state parameters, as opposed to parameters that are only available by tracking the storm evolution in time.

This brief review of distributions shows that the validity of probability models with respect to the sea state parameters are as important as

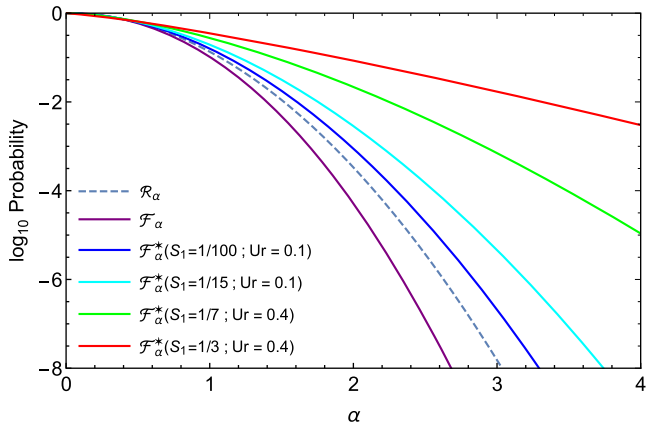


**Fig. 6.** $\mathcal{F}_\alpha$ and $\mathcal{F}^\star_\alpha$ compared with $\mathcal{R}_\alpha$ for different values of $\alpha$, $S_1$ and Ur.

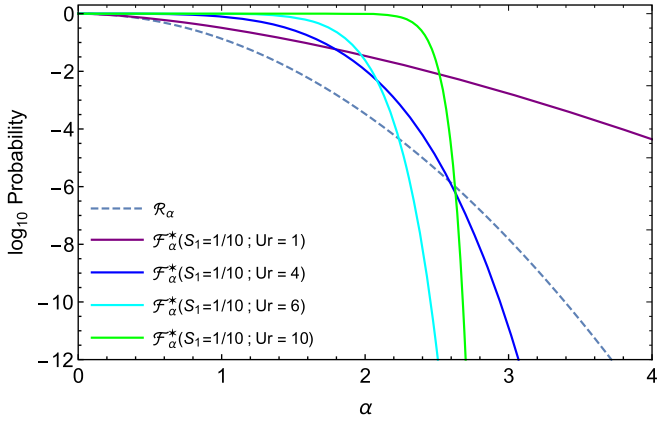


**Fig. 7.** Shallow water regime for the Forristall [25] model (large Ursell).

the data fit. Furthermore, this exercise was meant to show which regimes some of the most important distributions are capable of reproducing, such that following this work we can present a model that cover the widest possible set of sea states. In this regard, the mathematical structures of Forristall [24, 25], Haring et al. [33], Tayfun [77] make it unfeasible to model both super and sub-Rayleigh regimes while also producing unrealistic physical scenarios, whether formulated for wave heights or crest heights. Except for the first two empirical ones [24,33], is not surprising that the latter distributions are unable to provide the Sub-Rayleigh regime, as they account only for second-order bound modes who increases the likelihood. Therefore, Forristall [25], Tayfun [77] were devised to be limited to a certain range, as most distributions in the field, such that our previous discussion does not imply an unexpected deficiency. The same issue appears in several other distributions designed for deep water waves. For example, Fedele et al. [22] third-order model either exceeds or equals the likelihood assigned by Tayfun [77], thus not being able to produce sub-Rayleigh regimes (underprediction). On the other hand, The Boccotti [5, 6], Naess [62], Tayfun [78] distributions were devised to produce the Sub-Rayleigh regime and are identical in mathematical structure and occurrence probability [52,82], replacing the term $2\alpha^2$ in eq. (2) by $4\alpha^2/(1 \pm r(T))$, where $r(T)$ are slightly different ways to measure the autocorrelation function of the spectrum [39], thus being expectedly incapable of producing a distribution with Super-Rayleigh regime (overprediction) but recovering RD when $r(T) = \pm 1$. Moreover, Forristall [24] underpredicts rogue waves without the necessity of spectral shape and Longuet-Higgins [51] showed that a finite bandwidth correction is enough to display the same features, such that the autocorrelation dependent distributions have no theoretical or numerical advantage, in addition to not being more flexible than those presented in detail.

## 4. Physical bounds for sea state parameters

As discussed in the previous section, the major issue common to the distributions is the assignment of high finite probabilities to unrealistic physical scenarios. Although the limiting values of the major ratios in both infinite and finite depths are well established for individual waves, the analysis of sample averages and their respective ratios are not known. We aim to find empirical relations, as well as parameter limits for sea state variables that are capable of controlling the empirical exceeding probability distribution. We start by defining the *effective steepness* and the *nonlinearity* as, respectively (see Appendix B and Appendix C for the statistical notation):

$$S_{\mathrm{Eff},i} = \frac{H_i}{\langle\langle\lambda_2\rangle\rangle}, \quad \eta_{1/3} \equiv \left(\frac{\langle\mathscr{Z}_c\rangle}{\langle\mathscr{Z}_t\rangle}\right)_{H>H_{1/3}}. \tag{11}$$

We define the $j-$ th record significant steepness as:

$$s_{s,j} \equiv \frac{2\pi\,\langle\langle H_{m0}\rangle\rangle}{g\langle\langle T_z\rangle\rangle^2} \equiv \frac{\langle\langle H_{m0}\rangle\rangle}{\langle\langle\lambda_2\rangle\rangle}. \tag{12}$$

The entire storm steepness can be written as:

$$s_s = \frac{\langle H_{m0}\rangle}{\langle\lambda_2\rangle}, \quad \langle s_s\rangle = \left\langle\frac{\langle\langle H_{m0}\rangle\rangle}{\langle\langle\lambda_2\rangle\rangle}\right\rangle, \quad \varepsilon \equiv \frac{\langle H_{1/3}\rangle}{\langle\lambda_{1/3}\rangle}. \tag{13}$$

### 4.1. Empirical findings

Upon the thorough analysis of the fourteen storms of Table 1, a correlation between average of the skewness $\mu_3$ in eq. (C.4) and the nonlinearity $\eta_{1/3}$ emerges (Fig. 8):

$$\eta_{1/3} \approx 1 + \langle\mu_3\rangle, \quad \langle\eta_{1/3}\rangle \approx 1 + \langle\langle\mu_3\rangle\rangle. \tag{14}$$

Then, another relation provides insights on the scaling between important maxima of a storm (see Fig. 9):

$$\| \alpha \| \cdot \langle s_s\rangle \approx 0.96 \| S_{\mathrm{Eff}} \|. \tag{15}$$

In addition, the relationship between normalized effective steepness and wave heights reads (Fig. 10):

$$\frac{S_{\mathrm{Eff}}}{\langle S_{\mathrm{Eff}}\rangle_{1/3}} \approx \frac{5}{6}\alpha \equiv \frac{5H}{6H_{1/3}}, \tag{16}$$

which suggests that that rogue waves are more "effectively" steeper than the ordinary waves. If we measured the individual steepness instead of

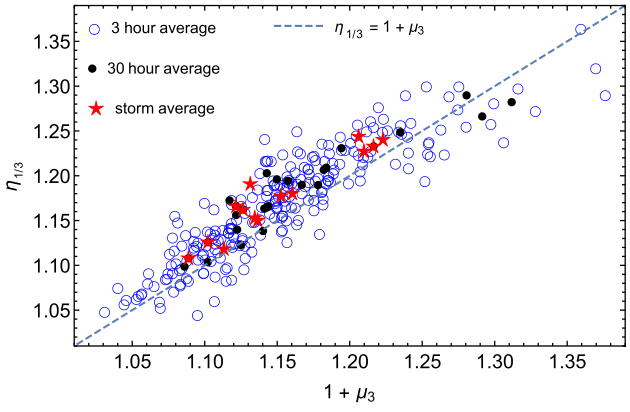


**Fig. 8.** Relation between nonlinearity and skewness. The correlation seems to be strong for both averagings $\langle\langle\mu_3\rangle\rangle$ and $\langle\mu_3\rangle$.

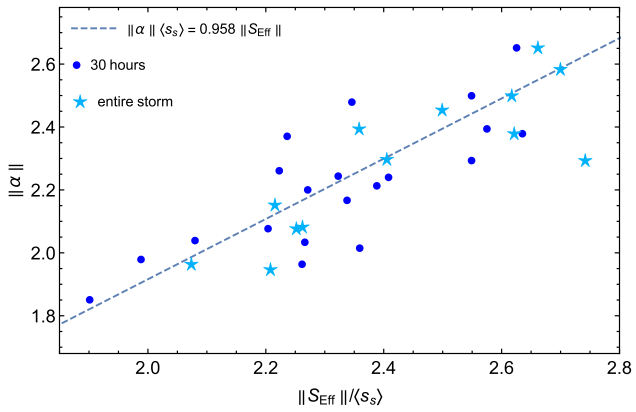


**Fig. 9.** The relationship between maximum normalized wave heights, maximum effective steepness and average significant steepness.

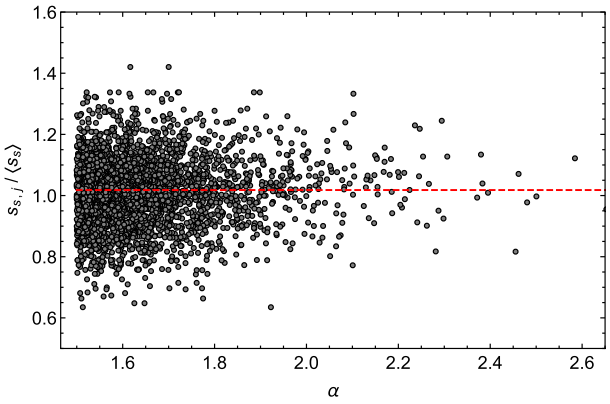


**Fig. 11.** Normalized significant steepness plotted against normalized heights.

its effective counterpart, the steepest waves are not necessarily rogue waves. On the other hand, as shown in Fig. 11, the normalized significant steepness does not grow with normalized heights and vice-versa, e.g. contrary to standard theories, rogue waves do not necessarily appear or are more likely to appear in very steep seas. Finally, a combination of eqs. (15–16) leads to the useful relation $\langle S_{\mathrm{Eff}}\rangle_{1/3} \approx 1.14\langle s_s\rangle$ (see Fig. 12).

### 4.2. Wave breaking limits

Focusing on variables reflecting the physics of 20-min records, an alternative formulation of Miche's limit reads:

$$\frac{7H_i}{\langle\langle\lambda_2\rangle\rangle} \leqslant \tanh\left(\frac{2\pi D}{\langle\langle\lambda_2\rangle\rangle}\right) \therefore \frac{H_i}{D} \leqslant \frac{\langle\langle\lambda_2\rangle\rangle}{7D}\ \tanh\left(\frac{2\pi D}{\langle\langle\lambda_2\rangle\rangle}\right). \tag{17}$$

Fig. 13 shows that such model fails to describe Stansell's data. Instead, we have found that a numerical modification proposed by Weggel [92], which is able to recover Miche [56] in the shallow water limit, provides a superior fit for both 20-min records and for entire storms:

$$\frac{H_i}{D} \equiv \epsilon\alpha \leqslant \| \epsilon\alpha \|_\infty \equiv \frac{16\pi}{55}\left[1 + \frac{4D}{\langle\lambda_2\rangle}\tanh\left(\frac{2\pi D}{\langle\lambda_2\rangle}\right)\right]^{-1}. \tag{18}$$

Notice that when the terms inside the bracket reach unity in shallow water ($\lambda_2 \to \infty$) we obtain $H \leq 0.91D$ as opposed to Miche's asymptotic value of $0.90D$. Though it bounds the vast majority of waves, the data analysis show that when converted to depth-limited breaking, the Miche-Stokes limit is not suitable to bound half a dozen rogue waves.

In addition, Miche [56] is also breached in its original form given by eq. (4). Fig. 14 shows how several waves violate the standard 1/7 limit in deep water whereas the significant steepness is well below the

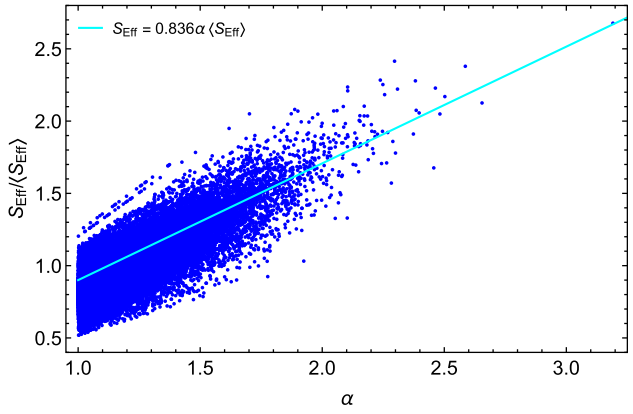


**Fig. 10.** The ratio $S_{\mathrm{Eff}}/\langle S_{\mathrm{Eff}}\rangle$ as a function of the dimensionless height.

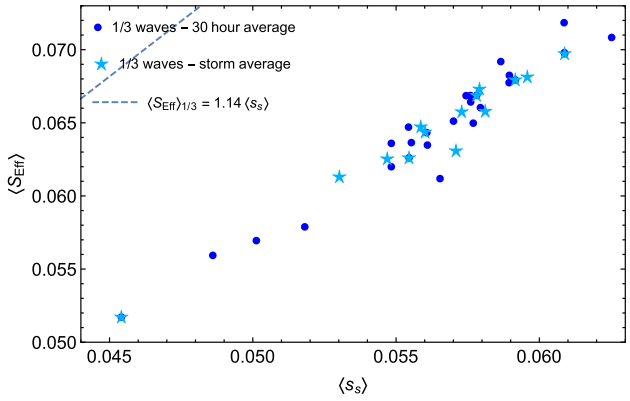


**Fig. 12.** The comparison of the average effective steepness amongst the 1/3 tallest waves and significant steepness of each averaging cluster.

threshold, noticing that only two rogue waves taller than 10 m exceeded Miche's rule. The significant steepness is more stable vertically and horizontally compared to the individual steepness, providing a better assessment of the sea state than if we computed the average individual steepness. As attested by Fig. 11, the significant steepness does not increase with high normalized wave heights, implying that high $s_s$ does not necessarily produce high $\alpha$, as seen in Fig. 14. Most notably, Fig. 14 confirms that rogue waves are typically steeper than the average wave, but also how hundreds of waves become as steep or steeper than themselves, in agreement with similar observations found by Christou and Ewans [11]. On average, rogue waves are not more than 20% steeper than sub-rogue waves (providing another argument for creating this subdivision) but are typically twice as steep as waves with $H \leq 1.75H_{1/3}$. In fact, only about 40% of rogue waves were at least twice as steep as the overall storm average. Since Miche's limit can not bound rogue waves either by individual or effective steepness, Fig. 15 provides motivation for the usage of the effective version of the individual steepness. Indeed, $S_{\mathrm{Eff}}$ diminishes the wild variations in water wave regime (deep or shallow). More importantly, it will also significantly lower the steepness of waves in very deep water: Stansell's data shows that typically 1.5% of all waves have very small wavelengths and individual steepness of the order of 1, but because $\lambda_2$ will not be so small the effective steepness is typically not higher than one seventh.

### 4.3. Bounding the effective steepness

Even though eq. (4) describes the limit for single waves, Miche's version for the effective steepness is breached by a handful of waves. Thus, instead of trying to modify Miche [56] or [3], we rather conjecture a limit for the effective steepness. Accordingly, motivated by the

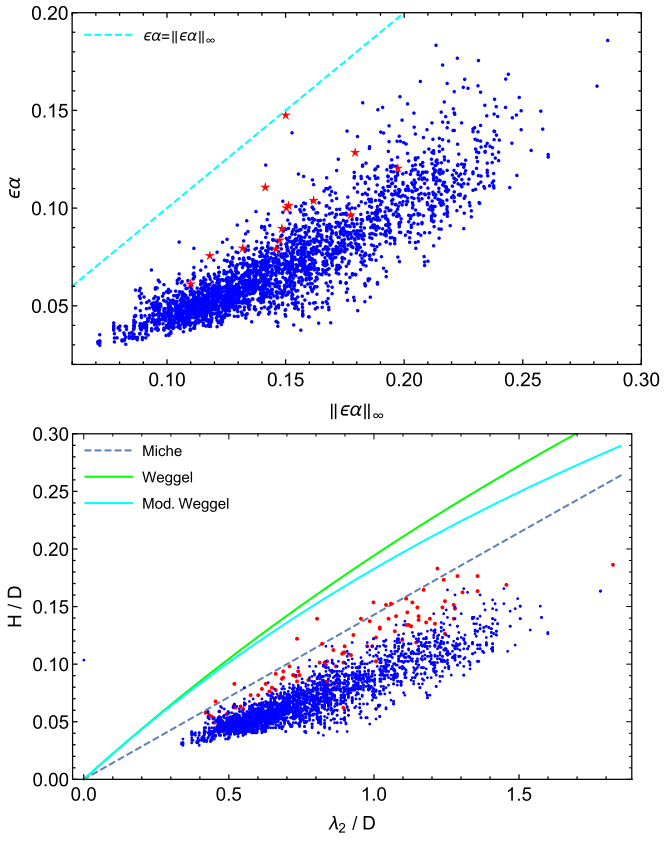


**Fig. 13.** The fitting of the depth-limited breaking limit: (top) scatter plot of observed height-to-depth ratio versus the theoretical maximum ratio and (bottom) models [56,92] versus data: blue dots are waves that exceed $H_{1/3}$ while red dots are rogue waves and red stars storm averages. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

expected bound on the significant steepness that is formulated in terms of $\lambda_1$ in eq. (10), we start this task by writing the range for individual waves as:

$$\lambda_i \approx (1 \pm \langle\langle\delta_1\rangle\rangle)\cdot\langle\langle\lambda_1\rangle\rangle \equiv \langle\langle\lambda_1\rangle\rangle \pm \sigma_{\langle\langle\lambda_1\rangle\rangle}. \tag{19}$$

Following eq. (10) and using Christou and Ewans's conclusion that the average significant steepness also obeys Miche [56] almost exactly as the individual steepness for the vast majority of cases, we extend [56] in eq. (4):

$$\| \frac{H_i}{\lambda_i} \| \approx \| \frac{H_{1/3}}{\langle\langle\lambda_1\rangle\rangle} \| \leqslant \frac{1}{7} \tanh\left(\frac{2\pi D}{\langle\langle\lambda_1\rangle\rangle}\right) \tag{20}$$

Hence, for a given minimum $i$-th wavelength we can estimate:

$$\| \frac{H_i}{\lambda_i} \| \sim \frac{H_i}{\min\lambda_i} \leqslant \frac{1}{7} \tanh\left(\frac{2\pi D}{\langle\langle\lambda_1\rangle\rangle}\right) \tag{21}$$

By means of eq. (19) is straightforward to arrive at the limit:

$$\frac{H_i}{\langle\langle\lambda_1\rangle\rangle} \leqslant \frac{1}{7}\left[1 - \langle\langle\delta_1\rangle\rangle\right] \tanh\left(\frac{2\pi D}{\langle\langle\lambda_1\rangle\rangle}\right) \tag{22}$$

For practical purposes, however, it is more convenient to use the second moment. Thus, by means of eq. (11) and eq. (B.5), we obtain the bound:

$$S_{\mathrm{Eff}} \leqslant \frac{22}{105}\left[1 - \langle\langle\delta_2\rangle\rangle\right] \tanh\left(\frac{30\,\pi D}{22\langle\langle\lambda_2\rangle\rangle}\right). \tag{23}$$

Remarkably, the coefficient 22/105 seems more suitable for Fig. 15, showing that the naive modification of Miche [56] in eq. (17) is replaced by eq. (23) as the suitable limit.

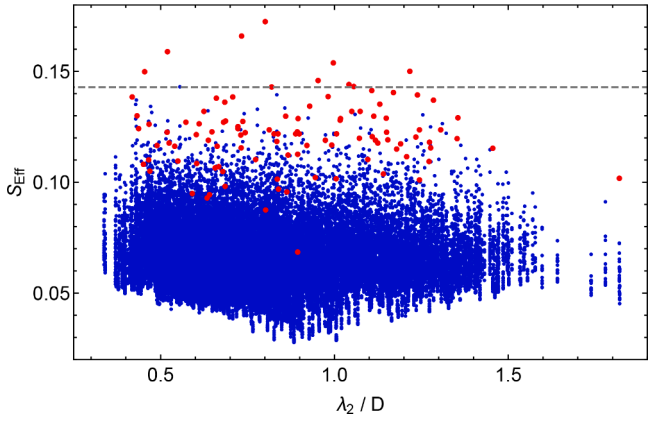


**Fig. 15.** Effective steepness of all waves taller than $H_{1/3}$ (blue dots) and rogue waves (red dots) versus the adjusted Miche's limit (dashed line) of eq. (17). As for the significant steepness in Fig. 14, the effective steepness provides a more homogeneous distribution of waves in regards of the water wave regime (whether deep or intermediate). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

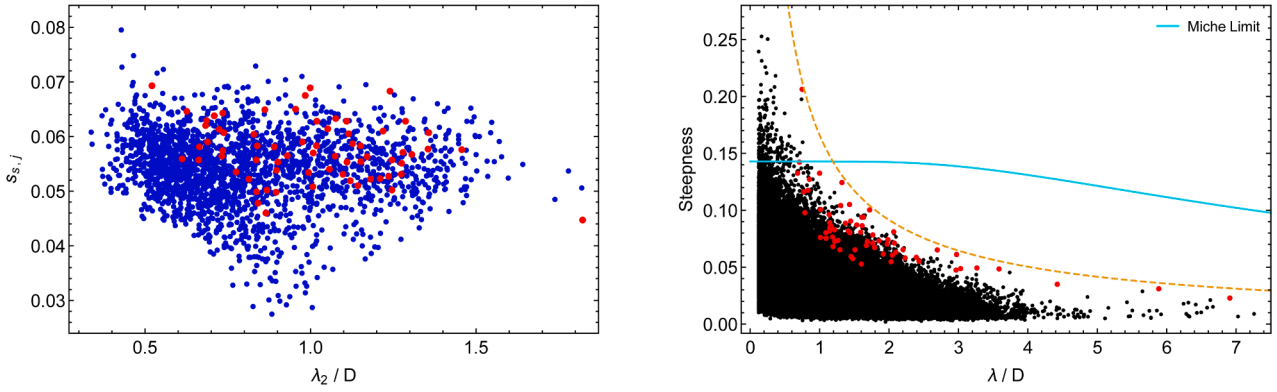


**Fig. 14.** Both (left) 20-min record significant steepness and (right) individual wave steepness versus water wave regime. For the first plot, 20-min records with or without rogue waves taller than 10 m are respectively denoted by red and blue dots, whereas the second plot has respectively red and black dots. The figure on the right has a range $0.125D \leq \lambda \leq 7.5D$, accounting for nearly 300,000 waves. The dashed curve shows the asymptotics of the data and is expressed as $6s \leqslant (\lambda/D)^{-0.86}$. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

## 5. Conjectured models for global maxima

Given the lack of theoretical studies on the maximal normalized heights rogue waves can reach, we attempt to shed some light into this problem through statistical reasoning based on the empirical relations of the previous section. Further analysis of these relations suggests that combining eqs. (15) and (23) one might estimate the dimensionless height peak. Therefore, if one applies the expected maximum operator $\mathbb{E}$ into eq. (15) and uses the bound in eq. (23) for $\mathbb{E}(\| S_{\text{eff}} \|)$, we obtain (see Appendix C and eq. (27) for nomenclature):

$$\mathbb{E}(\| \alpha \|) \equiv \alpha_\star \left(1 - \frac{\sigma_\lambda}{\overline{\lambda_2}}\right) \approx \alpha_\star \left[1 - \frac{6}{5}\left\langle \frac{\sigma_{\langle\lambda\rangle}}{\langle\langle\lambda_2\rangle\rangle}\right\rangle\right], \tag{24}$$

where the factor 1.2 comes from the fourteen storm mean ratio $\delta_2/\langle\langle\delta_2\rangle\rangle$ and $\alpha_\star$ denotes the pseudo maximum height $\alpha$ (steepness dependent), readily calculated for the North Alwyn data:

$$\alpha_\star \approx \frac{22}{105\langle s_s\rangle} \tanh\left(\frac{30\,\pi D}{22\langle\lambda_2\rangle}\right)\cdot(0.96 \pm 0.07). \tag{25}$$

Consequently, the central task is to obtain the overall storm variance of wavelengths when the value of every single wavelength is unknown. Even if every single wavelength is known, it is a good exercise because it will lead to modelling the variance in terms of sea state parameters. However, this procedure must be done carefully, as the modelling will serve both the effective steepness bound and the normalized wave height bound. Towards this modelling, the selection of sea parameters is paramount for the pattern of the rogue wave occurrence.

### 5.1. Rogue wave occurrence patterns

Studies have shown that individually, rogue wave occurrence finds no correlation with sea state parameters [11,74]. It is less common to investigate correlation with a combination of parameters. For instance, Table 4 confirms that there is no indication of individual sea state parameter control over the storms return periods, and yet a remarkable qualitative coherence between the latter, $\mathscr{N}_\alpha$, and the proposed dimensionless parameter is found (see Fig. 16):

$$\aleph_1 = \frac{\langle\lambda_{1/3}\rangle}{\eta_{1/3}D} = \frac{\langle\epsilon\rangle}{\varepsilon\,\eta_{1/3}} \equiv \frac{1}{\varepsilon}\,\aleph_2, \quad \aleph_2 = \frac{\langle H_{1/3}\rangle}{\eta_{1/3}D}. \tag{26}$$

While the whole North Alwyn dataset has the mean value $\aleph_1 \approx 1.1$ and is associated with an average probability (Longuet-Higgins's expected return period), the four storms with lowest return period have $\aleph_1 \approx 1.0$, whereas the mean ratio for the seven storms with highest return periods is $\aleph_1 \approx 1.2$. The discrepancy is particularly extreme for storms 124 and 195 with no rogue waves and $\aleph_1 > 1.3$. Fig. 17 shows how the combined parameter $\aleph_1$ follows the increase in return period of Table 1, where we assigned a return period for $\alpha = 2$ higher than the total number of waves in storms 124 and 195, respectively of 30,000 and 40,000 waves, with the latter being much higher due to the fact that the sub-rogue wave return period for storm 195 was 50% higher than in storm 124. The correlation coefficient for $\aleph_1$ of storm averages and their return period was found to be 0.704 whereas for $\aleph_1/\eta_{1/3}$ is 0.723 and for $\aleph_1\eta_{1/3}$ estimated as 0.629, suggesting that a combination of $\aleph_1$ and $\eta_{1/3}$ provides very little numerical improvement and no qualitative advantage over $\aleph_1$ alone. Besides, the term $\aleph_1\eta_{1/3}$ would be problematic in deep and shallow water, with a step-function occurring when $\lambda_{1/3}/D \to 0$ (e.g. similar to the extreme cases of Fig. 7) and the impossibility of rogue wave formation when $\lambda_{1/3}/D \to \infty$, thus, requiring the rejection of such parameter.

In spite of this congruence, one can not be sure of the exact dependence of the return period on $\aleph_1$ solely based on Table 4, as one might be inclined to notice a somewhat weaker correlation between the return period and the nonlinearity $\eta_{1/3}$. Fig. 8 assures us that the nonlinearity factor can be obtained from the skewness of the surface elevation, which in turn might be extracted from spectral variables [1,80]. In fact, is very common to associate rogue wave occurrence with the skewness or kurtosis of the sea surface [20,37,58,60] in addition to calculating probabilities based on such coefficients [49,61,81]. Though Stansell [74] already ruled out any individual correlation between rogue wave occurrence and spectral bandwidth, significant steepness, skewness (hence, the *nonlinearity* $\eta_{1/3}$) and kurtosis, this was done for individual waves and wave records, whereas any fair contrast to be drawn against $\aleph_1$ must be done for storm averages. Consequently, we plotted in Fig. 18 their possible relation to an increase in return period, thus extending Stansell's conclusions to storm averages. In fact, the correlation coefficient between the significant steepness $\langle s_s\rangle$ and the return period was of -0.039 and for $\varepsilon$ approximately -0.051, while for $\langle\nu\rangle$ it measured -0.253 and -0.247 for the nonlinearity $\eta_{1/3}$.

**Table 4**
The summary of patterns of Table 1. We examine how large (qualitatively) the three variables $\langle\epsilon\rangle$, $\varepsilon$ and $\eta_{1/3}$ are in comparison to their averages in Table 3 for every individual storm (first column). The colored boxes show the likelihood of finding a rogue wave compared to Longuet-Higgins's prediction.

| *Storm ID* | $\varepsilon$ | $\langle\epsilon\rangle$ | $\eta_{1/3}$ | $\varepsilon\,\eta_{1/3}/\langle\epsilon\rangle$ | $\varepsilon\,\eta_{1/3}^2/\langle\epsilon\rangle$ | $\mathscr{N}_\alpha^{-1}$ |
|---|---|---|---|---|---|---|
| **29** | High | Ave. | High | Very High | Very High | Very High |
| **149** | High | Vhigh | High | Ave. | High | Very High |
| **90** | High | Ave. | High | Very High | Very High | Very High |
| **172** | High | Ave. | High | Very High | Very High | Very High |
| **132** | Ave. | Vlow | Low | High | Ave. | Ave. |
| **28** | High | Ave. | Ave. | High | Ave. | Low |
| **146** | Vlow | Vlow | Vlow | Very Low | Very Low | Low |
| **23** | Ave. | Vhigh | Ave. | Low | Low | Very Low |
| **26** | Low | Vlow | Vlow | Low | Very Low | Very Low |
| **127** | Low | Low | Vlow | Very Low | Very Low | Very Low |
| **25** | Ave. | Ave. | Ave. | Ave. | Low | Very Low |
| **27** | Ave. | Low | Low | Ave. | Low | Very Low |
| **124** | High | Vhigh | Ave. | Low | Low | Very Low |
| **195** | Ave. | Vhigh | Ave. | Very Low | Very Low | Very Low |

In addition, parameters that are not accurately obtained from the spectral shape are a clear disadvantage for prediction purposes [81], as the skewness or kurtosis are often modeled from the spectrum not as accurately as $H_{1/3}$ or $T_z$ (see for example Annenkov and Shrira [1], Mori et al. [60] for modelling). As Gibson et al. [27] pointed out, exceeding probabilities should be dependent on the sea state parameters available from hindcast (e.g. accurately from the spectral shape), but even if the *nonlinearity* $\eta_{1/3}$ could not be obtained by hindcast, the new parameter $\aleph_1$ mitigates the spectral shape modelling dependence to a minimum. Regarding the forecast of rogue waves and building a exceeding probability (or complementary functions, such as the probability density) from hindcast coupled with HOSM models, see Bitner-Gregersen et al. [4].

The advantage of combining parameters into a single $\aleph_1$ is twofold: for one it does not depend solely on the spectral shape modelling and secondly, it gives new insights about varying physical settings. For instance, this parameter points to the realization that it is easier to form rogue waves in deep water than in shallow water wave regime. In fact, the works of Barbariol et al. [2], Chien et al. [10], Didenkulova and Anderson [13], Didenkulova and Rodin [14] provide evidence that depth shoaling has negative effects on rogue wave statistics. However, it does not become problematic in shallow water, as $\eta_{1/3}$ is expected to grow in this regime, weakening the growth of $\lambda_{1/3}/D$. Therefore, guided by the empirical results, we shall discuss the pattern presented in Tables 1 and 3 and build a framework for inserting physical parameters into eq. (24).

### 5.2. Subsample variance

In order to evaluate the bound for $S_{\text{eff}}$ we need to estimate $\langle\delta_2\rangle$ in eq. (23), thus, we write the storm variance of $\lambda$ calculated from 20-min

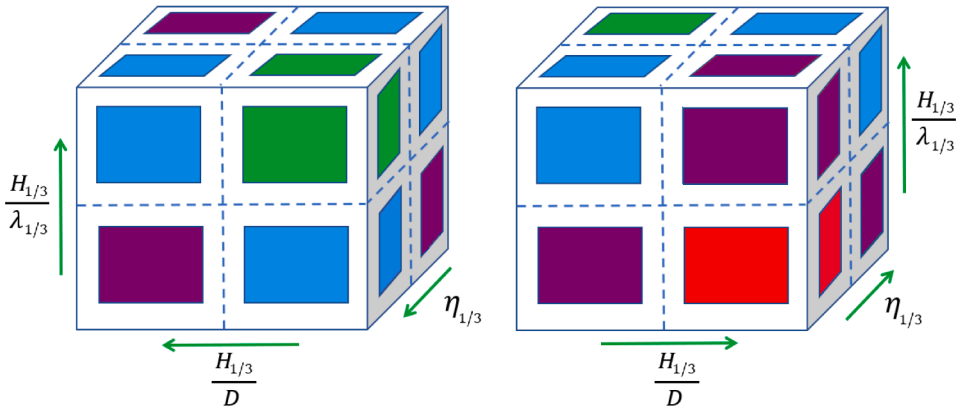


**Fig. 16.** Graphic display of the probability dynamics shown in Table 4 with the same color gradient conveying the relative probability [48]. Green arrows depict the direction where the variables increase.

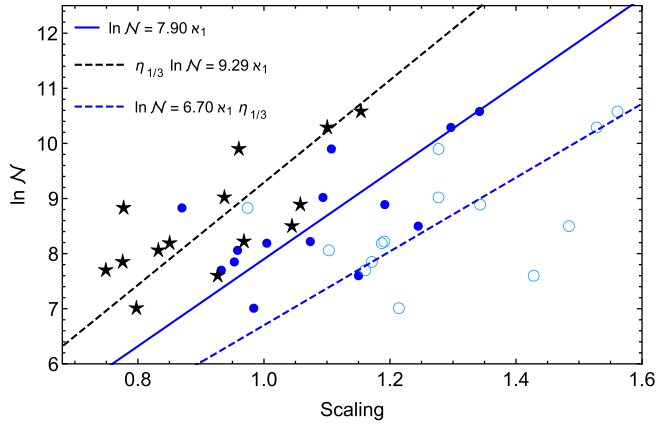


**Fig. 17.** Strong correspondence between rogue wave return periods of Table 1 and the scalings $\aleph_1$ (blue dots), $\aleph_1/\eta_{1/3}$ (black stars) and $\aleph_1\eta_{1/3}$ (blue circles). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

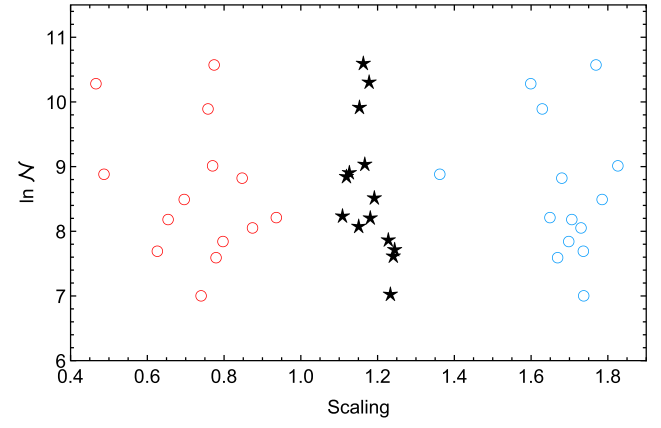


**Fig. 18.** Weak correspondence between rogue wave return periods of Table 1 and spectral parameters $30\langle s_s\rangle$ (blue circles), $\langle\nu\rangle$ (red circles) and $\eta_{1/3}$ (stars). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

record averaged wavelengths:

$$\sigma^2(\langle\lambda\rangle) = \frac{1}{p}\sum_{j=1}^{p}\Big(\lambda_j - \langle\lambda\rangle\Big)^2 \equiv \frac{1}{p}\sum_{j=1}^{p}\Big(\langle\langle\lambda\rangle\rangle_j - \langle\lambda\rangle\Big)^2, \tag{27}$$

where $p$ is the number of 20-min records. Our task is to obtain the exact value of each $\sigma(\langle\langle\lambda\rangle\rangle)$ from the data, but since we are not calculating the exact value of the zero-crossing wavelengths of all waves, one can assume an even partition of the record variance so that we estimate ($\tau$ is an input coefficient):

$$i\,(\lambda_{ij} - \lambda_j) = \pm\tau\,\sigma(\langle\lambda\rangle) \equiv \pm\tau\,\sigma_{\langle\lambda\rangle}. \tag{28}$$

Thus, we may write:

$$N_j\,\sigma^2_{\langle\langle\lambda\rangle\rangle} = \sum_{i=1}^{N_j}\Big(\lambda_{ij} - \lambda_j\Big)^2 = 2\tau^2\sum_{i=1}^{(N_j-1)/2}\frac{\sigma^2_{\langle\lambda\rangle}}{i^2},$$

where $N_j$ is the number of waves in the $j$-th 20-min record. Then, we can estimate for large $N$:

$$\frac{N_j\,\sigma^2_{\langle\langle\lambda\rangle\rangle}}{2\sigma^2_{\langle\lambda\rangle}} \overset{N\to+\infty}{\longrightarrow} \sum_{i=1}^{+\infty}\frac{\tau^2}{i^2} = \frac{\pi^2\tau^2}{6}, \tag{29}$$

hence, on average, the 20-min record variance reads:

$$\big\langle\sigma^2_{\langle\langle\lambda\rangle\rangle}\big\rangle \sim \sum_{j=1}^{p}\frac{N_j}{N}\,\sigma^2_{\langle\langle\lambda\rangle\rangle} \approx \frac{\pi^2\,\tau^2}{3\langle n\rangle}\sigma^2_{\langle\lambda\rangle}, \quad p\,\langle n\rangle = N\,. \tag{30}$$

Since the 20-min record average zero-crossing period is found through $\langle n\rangle\langle T_z\rangle \approx 1200$, we may write:

$$\left\langle\frac{\sigma_{\langle\langle\lambda\rangle\rangle}}{\langle\langle\lambda_2\rangle\rangle}\right\rangle \approx \frac{\pi\tau}{60}\langle T_z\rangle^{1/2}.\frac{\sigma_{\langle\lambda\rangle}}{\langle\lambda_2\rangle}. \tag{31}$$

### 5.3. Expected maximum dimensionless height

After inferring what physical variables could appear in $\delta_1$ and obtaining an expression for the subsample variance, we are ready to construct a full model for eq. (24):

$$\mathbb{E}(\,\|\,\alpha\,\|\,) \equiv \alpha_\star\,\mathbb{E}^\dagger_\alpha \approx \alpha_\star\left[1 - \frac{\pi\tau}{50}\langle T_z\rangle^{1/2}.\frac{\sigma_{\langle\lambda\rangle}}{\langle\lambda_2\rangle}\right]. \tag{32}$$

Following the patterns of Table 4, we intend to model the deviation in the following manner:

$$\sigma_{\langle\lambda\rangle} \approx \mathscr{A}\,\eta^a_{1/3}\,\aleph_1^b\langle\lambda_2\rangle. \tag{33}$$

The reason for the absence of second auxiliary parameter in the equation above is that $\aleph_2 << \aleph_1$, which would make $\mathbb{E}_\alpha^\dagger$ be always near to unity. Consequently, this would invalidate the model for expected maxima since the standard deviation of the storm averaged significant steepness within the data set is of the order of only 6.8% and can not explain the 13.9% standard deviation of $\| \alpha \|$ in Table 1 alone. Hence, we can merge the equations (32) and (33) to obtain the general structure for:

$$\mathbb{E}_{\alpha,0}^\dagger = 1 - \frac{\mathscr{A}\pi\tau}{50}\eta_{1/3}^{a-b}\left(\frac{\langle\lambda_{1/3}\rangle}{D}\right)^b \langle T_z\rangle^{1/2}. \tag{34}$$

Therefore, our task is to obtain a good model for $(a, b, \mathscr{A})$ so that it describes the observation of both $\langle\langle\delta_2\rangle\rangle$ and $\| \alpha \|$ properly. On the other hand, the term $T_z^{1/2}$ can grow quickly for wind waves, with the possibility of producing an unrealistic scenario of $\mathbb{E}_\alpha^\dagger \leqslant 0$. Therefore, we adjust equations (32–33) with:

$$\langle\langle\delta_2\rangle\rangle^* \longrightarrow \sqrt{\frac{\aleph_1^3}{2\eta_{1/3}^7}\cdot\langle\langle\delta_2\rangle\rangle}, \tag{35}$$

adjusting the physical implications of $\langle\delta_2\rangle$ and maintaining the fourteen storm average. The importance of this adjustment is twofold: Firstly, it makes the term $(a - b)$ in eq. (34) become very negative, which is desirable as $\eta_{1/3}$ grows in shallow water and it counters the growth of the ratio $\lambda_{1/3}/D$. Secondly, we have better control of the zero-crossing period. Accordingly, we would obtain an adjusted version of eq. (34):

$$\mathbb{E}_\alpha^\dagger = 1 - \sqrt{\frac{6.\mathscr{A}\pi\tau}{500}}\eta_{1/3}^{(a-b-10)/2}\left(\frac{\langle\lambda_{1/3}\rangle}{D}\right)^{(3+b)/2}\langle T_z\rangle^{1/4}. \tag{36}$$

A different but crucial aspect of this optimization is the shallow water wave constraint. The structure of equation (24) shows that $\mathbb{E}_\alpha^\dagger$ will ultimately be inserted in the upper limit of the steepness, since $\mathbb{E}_\alpha^\dagger$ contains $1 - \langle\delta_2\rangle$ to find the expected maximum and $1 + \langle\delta_2\rangle$ for the upper bound. In view of the limit of large $\lambda_2/D$, we would obtain a steepness approaching zero in accordance with Miche [56], thus, forcing one to obtain:

$$\lim_{\lambda_2/D\to+\infty}\left[1 + \langle\delta_2\rangle\right]\tanh\left(\frac{30\pi D}{22\langle\lambda_2\rangle}\right) \equiv \lim_{\lambda_2/D\to+\infty}\mathscr{W} = 0\ . \tag{37}$$

Therefore, having eq. (C.2) in mind, the limit is evaluated:

$$\lim_{\lambda_2/D\to+\infty}\frac{\tanh\left(\frac{30\pi D}{22\langle\lambda_2\rangle}\right)}{\tanh^{1/8}\left(\frac{2\pi D}{\langle\lambda_2\rangle}\right)}\left(\frac{\langle\lambda_2\rangle}{D}\right)^\rho \leqslant 1\ ,\ 2\rho = 3 + b\,. \tag{38}$$

In the shallow water limit we have $\tanh(s_s/\epsilon) \approx s_s/\epsilon$, hence:

$$\mathscr{W}(\rho) \sim \frac{30\pi}{22\cdot(2\pi)^{1/8}}\cdot\left(\frac{\langle\lambda_2\rangle}{D}\right)^{\rho-7/8} \sim \frac{7}{2}\left(\frac{\langle\lambda_2\rangle}{D}\right)^{\rho-7/8}. \tag{39}$$

Accordingly, the limit of eq. (38) will be fulfilled whenever $\rho < 7/8$ (see Fig. 19), which requires the condition $b < -5/4$. Then, the best model with the structure of eq. (33) that will provide a good fit for $\langle\delta_2\rangle$ and obey the constraint of eq. (37–38) sets $(a = 0, b = -2, \mathscr{A} = 2/7)$. A similar model that obeys eq. (34) and (38), now with $b < 7/4$, can be found, such that:

$$\langle\delta_2\rangle_{\mathscr{J}} = \frac{2}{7\aleph_1^2}\ ;\ \langle\delta_2\rangle_{\mathscr{J}}^* = \sqrt{\frac{\aleph_1}{7\eta_{1/3}^7}}\ ;\ \langle\delta_2\rangle_{\mathscr{JJ}} = \frac{\eta_{1/3}^3}{6\aleph_1}\,. \tag{40}$$

The method in eq. (35) re-scales the scatter plot and diminishes the number of outliers while improving the correlation significantly, now reading 0.492. However, as discussed earlier, the physical interpretation is as important as the data fit in Fig. 20 and paramount for eq. (32). For instance, if we switch $\langle\lambda_{1/3}\rangle$ by its median $\underline{\lambda_{1/3}}$, we find a correlation of

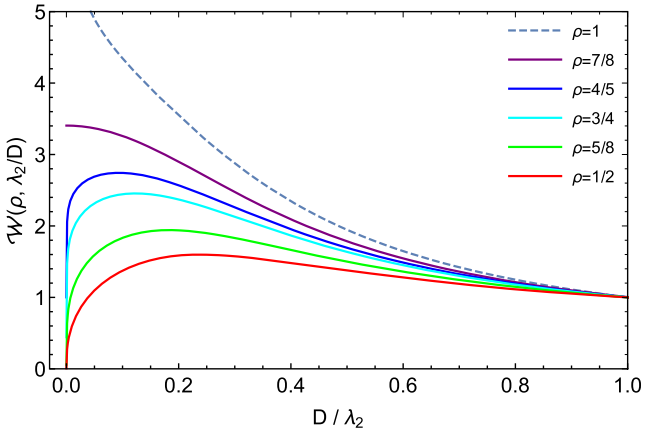


**Fig. 19.** Outcomes for the limit eq. (38) with varying $\rho$. The function reaches zero when $\rho < 0.875$ and otherwise can either become finite or diverge.

0.503, but of no qualitative advantage (see Fig. 21). The model $\langle\delta_2\rangle_{\mathscr{JJ}}$ produces similar correlation without the adjustment, but will likely overpredict the maximum normalized heights. Besides the fitting and the shallow water constraint, the model $\langle\delta_2\rangle_{\mathscr{J}}^*$ also reverses the decreasing trend of $\langle\delta_2\rangle$ with return period, which will result in higher $\|\alpha\|$ for storms with smaller return period due to eq. (32). Though it is expected that storms with higher rogue wave occurrence to produce higher $\|\alpha\|$, this is not always true, as seven storms with higher $\mathscr{N}_{\alpha=2}$ than storm 29 displayed $\|\alpha\|$, consequently, the growth of $\langle\delta_2\rangle$ with $\mathscr{N}_\alpha$ should not be monotonic or exponential, making of $\langle\delta_2\rangle_{\mathscr{J}}^*$ a more suitable model.

### 5.3.1. *Storm geometry and τ*

In the last section we learned how to constrain the structure for the expected maximum normalized height, obtaining a reasonable model for the wavelength deviation $\sigma_{\langle\lambda\rangle} = \sigma(\langle\lambda_2\rangle, \aleph_1)$. The final part of the puzzle consists in estimating the parameter $\tau$ inserted in eq. (30). If $\tau$ is just a constant, we shall not see large variations for its value regardless of which 20-min records within any storm from Table 1 we choose. We report that the dimensionless deviation of all wavelengths $\delta_2$ is typically twice of $\langle\delta_2\rangle$, which for storm 29 (with 89 records) was of 1.96, whereas $\delta_2/\langle\langle\delta_2\rangle\rangle \approx 1.18$. Then, using eq. (30) we obtain:

$$\tau_{29} \geqslant \frac{1.96}{1.18\pi}\sqrt{458.76} \approx 11.22. \tag{41}$$

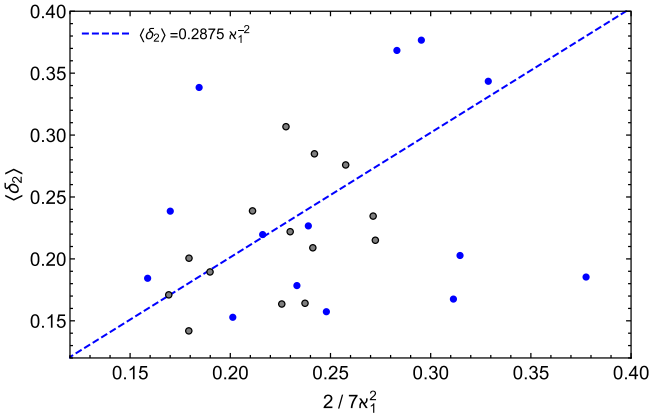


**Fig. 20.** Observed deviation (vertical axis) modeled with (grey dots) and without (blue dots) the adjustment of eq. (35). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

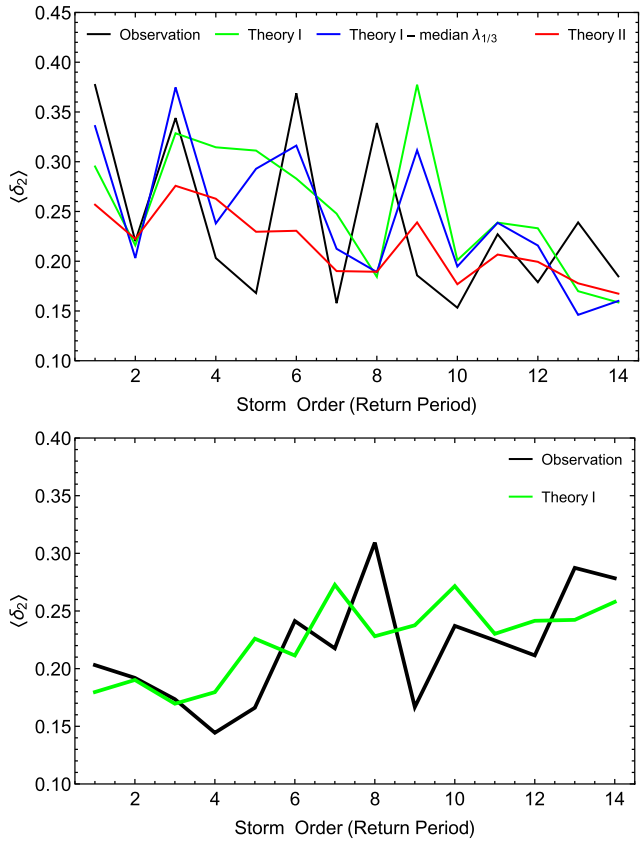


**Fig. 21.** The observed dimensionless deviation $\langle\delta_2\rangle$ plotted with increase of return period of all fourteen storms (as seen in Table 1): (top) solutions without adjustment shown in eq. (40) and (bottom) with the adjustment of eq. (35).

Nevertheless, such approach presents the same difficulty of modelling probabilities with skewness and kurtosis [81], as it depends on ratios not known *a priori* and the size of the sample, e.g. how large is $\langle n\rangle$. However, the latter is equivalent to an inverse dependence on the zero-crossing period, such that we want to model $\tau$ with sea state parameters and a calibrated ratio $\langle\langle\delta_2\rangle\rangle/\langle\delta_2\rangle$. A better approach would be calculating the storm parameter for all 20-min records and average them out over the whole storm. We combine eq. (30) with eq. (33) as in Fig. 20, obtaining:

$$\frac{\sigma_{\langle\langle\lambda\rangle\rangle}}{\langle\langle\lambda_2\rangle\rangle} \cong \frac{\pi\tau}{60}T_z^{1/2}.\frac{\sigma_{\langle\lambda\rangle}}{\langle\lambda_2\rangle} \approx \frac{\pi\tau\,\eta_{1/3}^2 D^2}{210}.\frac{\langle\langle T\rangle\rangle_z^{1/2}}{\langle\langle\lambda_{1/3}\rangle\rangle^2}. \tag{42}$$

In Fig. 22 we see how the evolution of both significant wave height and $\lambda_{1/3}$ affects eq. (42). A comparative analysis between Figs. 22 and 23 suggest that $\tau$ is controlled by the variability of the $H_{1/3}$ time series, with $\tau$ growing as the storm reaches its peak while decreasing afterwards. The correlation coefficients for the three intervals (pre-peak, peak, post-peak) of storm 29 were respectively of 0.766, 0.556 and -0.049, whereas for the combined interval of peak and post-peak was of 0.689. Therefore, it is not surprising to have a very low storm geometry parameter $\tau$ in storm 146 and a relatively high one for storm 29. Nonetheless, the previous equation is based on the asymptotic large $N$ limit, such that the parameter $\tau$ might be a few percent smaller than its estimate. Fig. 24 shows how the mean $\tau$ estimate for is related to various variance measures of $H_{1/3}(t)$. Though a full expression is not available, these approximate measures show how the input $\tau$ is dependent on how large and quick are $H_{1/3}(t)$ variations. Qualitatively, Fig. 22 confirms the mean estimates in Table 5.

### 5.3.2. Model comparison

After all the constraints and algebra, we are able to finally find an expression for eqs. (24–25). We then use the model of Fig. 20 and eq. (32) to find:

$$\mathbb{E}(\,\|\,\alpha\,\|\,) = \frac{22}{105\langle s_s\rangle}\left[1-\sqrt{\frac{\pi\tau\langle\lambda_2\rangle}{180\,D}}.\frac{\langle T_z\rangle^{1/4}}{\eta_{1/3}^4}\right]\tanh\left(\frac{30\pi D}{22\langle\lambda_2\rangle}\right), \tag{43}$$

where one originally derived it in terms of $\lambda_{1/3}$ with coefficient reading $3\pi\tau/875$, from which we used the values of Table 3, but for the sake of consistency, we applied $\lambda_{1/3}\approx 1.6\lambda_2$ to obtain the previous equation. Notice that both the storm geometry parameter and the ratio among spectral periods $T_E$, $T_1$ and $T_2$ will affect the expectancy model (see Appendix B), accomplishing the goal of predicting ocean wave statistics from sea state parameters that evolve from the ocean spectrum, as opposed to the number of waves in the record alone. The correlation between theory and observation vary wildly (Fig. 25): For the model in eq. (43) with constant $\tau$ we have a correlation of nearly 0.47, with maximum correlation (using error bars) of 0.88 whereas a varying $\tau$ changes it to 0.61 and 0.92 respectively. In contrast, the alternative model in the second figure had an average correlation of -0.34 (blue circles), Longuet-Higgins [48] of 0.30 and Forristall [24] of 0.31 (with the last two expectancy models derived in Appendix D). Then, Longuet-Higgins's and Forristall's models demonstrate that their lack of variability for the exceeding probability in Section 3 is naturally reproduced in the expected maximum height. Put into perspective, the difference between the estimated maximum $\alpha$ for the storm with the super-rogue wave (172) to the one with no rogue waves (195) was of 64% ($\approx$3.19/1.95). However, according to both Longuet-Higgins [48] and Forristall [24] we should expect this difference to be of only 8% (respectively estimating $\approx$2.39/2.21 and $\approx$2.14/1.99), whilst eq. (43) returns 47% ($\approx$2.77/1.88). In fact, the models in equations (D.7) and (D.10) would make one believe that the expected maximum would be unusually high in large data sets with hundreds of millions of waves, however, Table 1 reports the exact opposite, that the largest rogue waves are not necessarily found in the largest data sets. Since Haring et al. [33], Tayfun [77] and Forristall [25] were derived for wave crests, we did not plot them in Fig. 25 as it would do an injustice to their predictive models. However, they will feature a slightly larger variability but not sufficient to explain the observed $\|\,\alpha\,\|$. In other words, any expectancy for the dimensionless heights that is constructed from distributions without physical variables will ultimately be unfit because it depends on the total number of waves in the storm alone. Even if a distribution depends on a physical variable, it may fail because it might not cover both super-Rayleigh and sub-Rayleigh regimes.

### 5.4. Additional expected maxima

In order to estimate the maximum effective steepness one shall apply the expectancy operator to eq. (16) and using the ratio in Fig. 12 while adding the error $\pm 0.09$ in Fig. 9:

$$\frac{\mathbb{E}(\,\|\,S_{\mathrm{Eff}}\,\|\,)}{\mathbb{E}(\,\|\,\alpha\,\|\,)\langle s_s\rangle} = \frac{0.93\cdot(0.96\pm 0.07)}{(0.88\pm 0.02)} \approx 1.02\pm 0.08, \tag{44}$$

hence, employing eq. (43) to eq. (44), one finds:

$$\mathbb{E}(\,\|\,S_{\mathrm{Eff}}\|\,) = \frac{3}{14}\left[1-\sqrt{\frac{\pi\tau\langle\lambda_2\rangle}{180\,D}}.\frac{\langle T_z\rangle^{1/4}}{\eta_{1/3}^4}\right]\tanh\left(\frac{30\pi D}{22\langle\lambda_2\rangle}\right). \tag{45}$$

Fig. 26 shows the accuracy of it with error bars that follows the same methodology described in Fig. 25. Likewise, the reader might use the same approach from the previous sections to obtain the expected maximum of the significant height-to-depth ratio. Contrary to this idea, Fig. 27 shows the tendency of the significant height-to-depth ratio to drop for the bulk of waves as the dimensionless height increases.

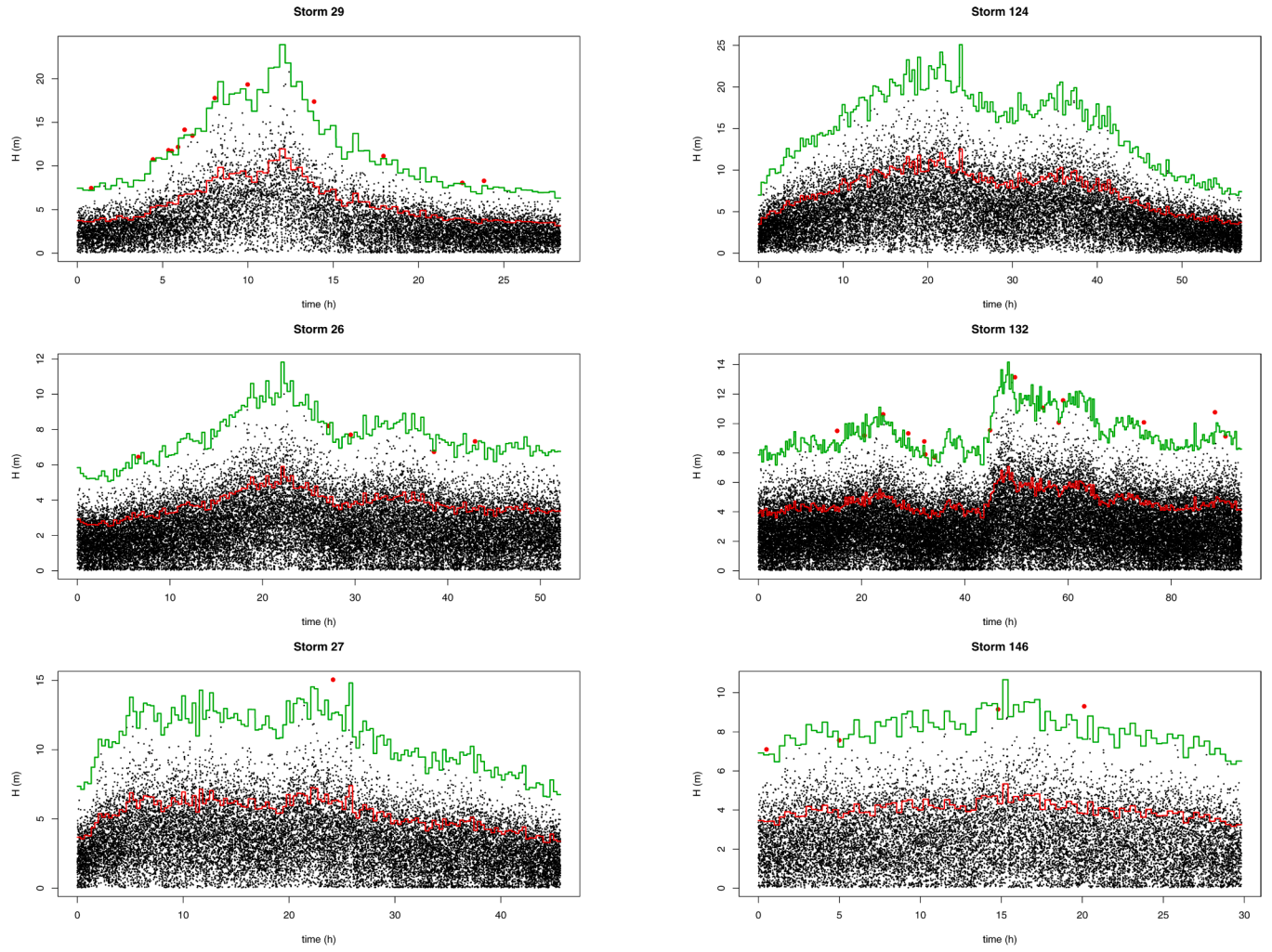


**Fig. 22.** Significant wave height time series of selected storms shown by red lines, whereas green lines depict the rogue wave threshold. Red dots are rogue waves and black dots otherwise. Storm parameters are estimated to be at a maximum $\langle \tau_{29} \rangle \approx 15$ and at a minimum $\langle \tau_{146} \rangle \approx 6$ (see Table 5). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

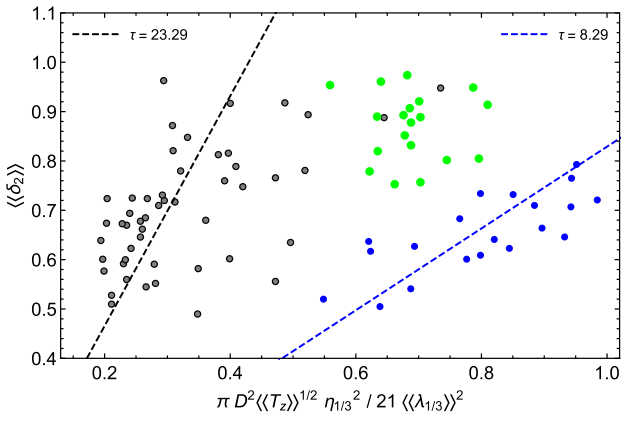


**Fig. 23.** Storm29's $\tau$ model of eq. (42) before the peak (blue dots), defined as $\langle\langle \lambda_2 \rangle\rangle > \langle \lambda_{1/3} \rangle$, during the peak (grey dots) and after the peak (green dots). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

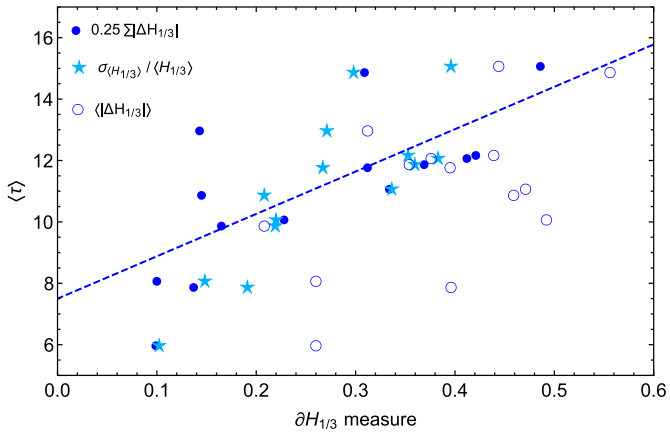


**Fig. 24.** Scatter plot among $\tau$ and two measures of for the evolution of $H_{1/3}(t)$. The dashed line shows the best fit between $\tau$ and $\sigma_{\langle H_{1/3} \rangle} / \langle H_{1/3} \rangle$, whose correlation is of 0.812. The other two measures estimate the curvature by the absolute value of interval variations of $H_{1/3}$: blue dots is calculated by splitting the entire storm into five intervals, whereas blue circles are calculated through $(n-1)$ intervals, with respective correlations of 0.926 and 0.433. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

**Table 5**
Comparison between observed maximum dimensionless height and our model of (43) adjusted to each storm geometry parameter $\tau$, estimated from the combination of the model of eq. (42) and a rough measure of significant wave height evolution $\sigma_{\langle H_{1/3}\rangle}/\langle H_{1/3}\rangle$, as well as based on Longuet-Higgins [48] and Forristall [24], whose estimates are found in Appendix D. In addition, we explicitly estimate values for each one of the three upper bounds for the dimensionless heights, where Miche [56] limit is calculated by eq. (47) .

| Storm ID | $\|\alpha\|$ | $\mathbb{E}(\|\alpha\|)$ | $\mathbb{E}(\mathscr{D}_{\|\alpha\|})$ | $\mathbb{E}(\mathscr{T}_{\|\alpha\|})$ | $\langle\tau\rangle$ | $\frac{\sigma_{\langle H_{1/3}\rangle}}{\langle H_{1/3}\rangle}$ | $\sum_{k=1}^{4}\frac{\lvert\Delta H_{1/3}\rvert_k}{4}$ | $\alpha_\infty$ | $\frac{\|\epsilon\alpha\|_\infty}{\langle\epsilon\rangle}$ | Miche's |
|---|---|---|---|---|---|---|---|---|---|---|
| 29 | **(2.30)** | $2.36 \pm 0.17$ | 2.25 | 2.02 | **15** | 0.396 | 0.486 | 4.79 | 3.31 | 2.47 |
| 149 | **(2.50)** | $2.59 \pm 0.19$ | 2.39 | 2.14 | **11** | 0.220 | 0.228 | 4.84 | 3.37 | 2.61 |
| 90 | **(2.65)** | $2.53 \pm 0.18$ | 2.38 | 2.13 | **13** | 0.383 | 0.412 | 4.60 | 3.29 | 2.45 |
| 172 | **(3.19)** | $2.57 \pm 0.19$ | 2.31 | 2.07 | **11** | 0.336 | 0.334 | 4.72 | 3.40 | 2.55 |
| 132 | **(2.30)** | $2.48 \pm 0.18$ | 2.38 | 2.13 | **8** | 0.148 | 0.100 | 4.69 | 3.59 | 2.62 |
| 28 | **(2.38)** | $2.34 \pm 0.17$ | 2.30 | 2.06 | **12** | 0.360 | 0.369 | 4.93 | 3.38 | 2.52 |
| 146 | **(2.46)** | $2.56 \pm 0.19$ | 2.26 | 2.03 | **6** | 0.102 | 0.099 | 4.95 | 3.86 | 2.78 |
| 23 | **(2.08)** | $2.13 \pm 0.16$ | 2.31 | 2.07 | **12** | 0.353 | 0.421 | 4.83 | 3.19 | 2.44 |
| 26 | **(2.16)** | $2.40 \pm 0.17$ | 2.33 | 2.08 | **10** | 0.219 | 0.165 | 4.99 | 3.79 | 2.69 |
| 127 | **(2.09)** | $2.38 \pm 0.17$ | 2.26 | 2.02 | **13** | 0.271 | 0.143 | 6.75 | 4.21 | 3.24 |
| 25 | **(2.59)** | $2.09 \pm 0.15$ | 2.27 | 2.04 | **12** | 0.267 | 0.312 | 4.70 | 3.24 | 2.40 |
| 27 | **(2.40)** | $2.53 \pm 0.18$ | 2.29 | 2.05 | **8** | 0.191 | 0.137 | 5.08 | 3.64 | 2.72 |
| 124 | **(1.97)** | $2.06 \pm 0.15$ | 2.30 | 2.06 | **15** | 0.298 | 0.309 | 5.71 | 3.40 | 2.70 |
| 195 | **(1.95)** | $2.03 \pm 0.15$ | 2.21 | 1.99 | **11** | 0.208 | 0.145 | 4.99 | 3.23 | 2.49 |

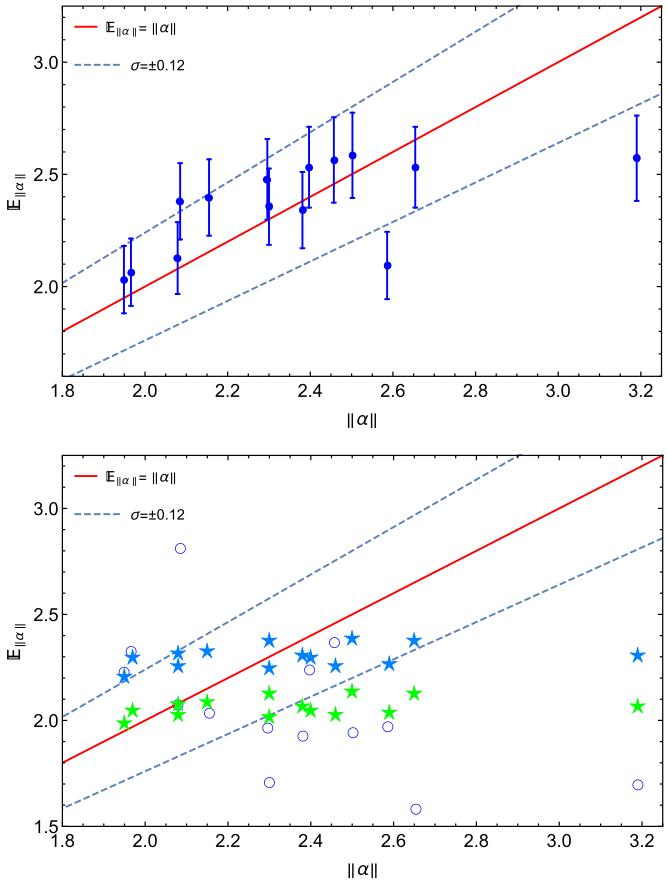


**Fig. 25.** Theoretical expected maximum $\alpha$ (vertical axis) of eq. (43) versus the observation (horizontal) described in Table 1: (top) Blue dots represent the midpoints of eq. (43) and the bars come from the error in $\alpha_\star$ with estimated $\tau$ for each storm (see Table 5) and (bottom) blue circles for the expectancy according to eq. (34) if our expression for the deviation was $\langle\delta_2\rangle \approx \eta_{1/3}^2/6\aleph_1$ and $\tau = 11$ (the average among all storms) and predictions by Longuet-Higgins [48] in blue stars and [24] in green stars (see Appendix D). Notice that besides a higher correlation, our model of eq. (43) follows an increasing trend while the Longuet-Higgins [48] and [24] models portray a rigid horizontal line. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

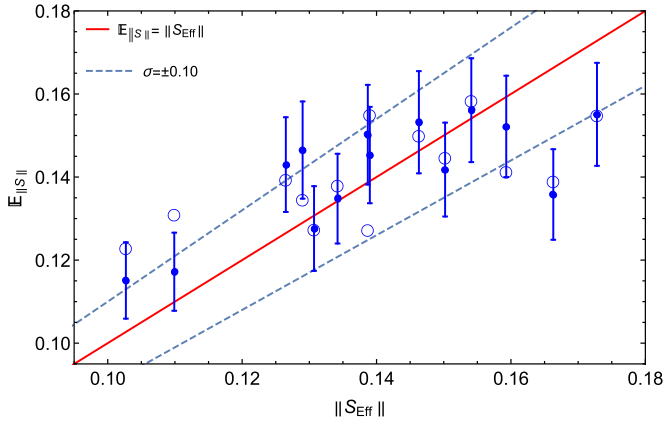


**Fig. 26.** Theoretical expected maximum effective steepness versus observation. Blue dots and bars are obtained from eq.(45) with varying $\tau$ from Table 5 and combined error from $\alpha_\star$ and eq. (44), becoming $0.08 \times 22/105 \approx \pm 0.017$, whereas blue circles are midpoints of the same model with $\tau = 11$. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

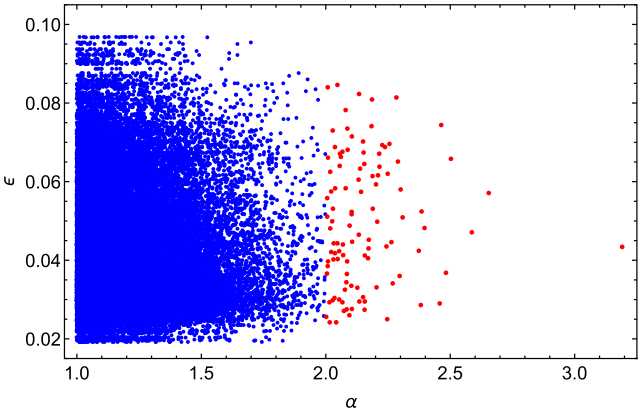


**Fig. 27.** The bulk of height-to-depth ratio with varying normalized heights.

Therefore, there is little use in finding the maximum $\epsilon$ because the typical waves associated with such values are not rogue waves, precisely as eq. (18) describes it.

### 5.5. Conjectured bounds

Furthermore, one can estimate the greater upper bound for the dimensionless height in the tail $(1 + \langle\langle \delta_2 \rangle\rangle)$ of the population:

$$\alpha_\infty = \frac{22}{105\langle s_s \rangle}\left[1 + \sqrt{\frac{\pi\tau\langle\lambda_2\rangle}{180\,D}.\frac{\langle T_z\rangle^{1/4}}{\eta_{1/3}^4}}\right]\tanh\left(\frac{30\pi D}{22\langle\lambda_2\rangle}\right). \tag{46}$$

In order to estimate its evolution from deep water waves to shallow water, we need to isolate the $\lambda_2/D$ ratio from the zero-crossing period. Fig. 28 shows the mathematical limit in detail, demonstrating a clear tendency of optimal rogue wave formation in intermediate or near deep water wave regimes, e.g. shallow water regime has an exceeding typically below the Rayleigh regime [3,40,54,95]. However, a higher steepness will likely diminished the chances of forming rogue waves, especially in shallow water. This framework provides fresh insights into rogue wave research: For one, it was already known that neither high individual steepness nor significant steepness is associated with rogue wave formation [11,74], however, its was not known that the significant steepness could have a negative effect on the rogue wave occurrence. Additionally, the bound eq. (46) flat out rejects the weakly nonlinear model as it does not depend on the number of waves. Nevertheless, as soon as the estimates derived in Section 4.1 and Appendix B lose validity, the accuracy of eq. (43) and eq. (46) will tend to diminish because their coefficients depend on the ocean spectrum and are not fixed.

On the other hand, let us analyze other problematic consequences of Miche's limit in the context of maximum normalized wave heights. One may be inclined to think that eq. (46) is a cumbersome approach and rewrite eq. (17) as:

$$\| \alpha \|_{M,\infty} \leqslant \frac{\langle\lambda_2\rangle}{7\langle H_{1/3}\rangle}\tanh\left(\frac{2\pi D}{\langle\lambda_2\rangle}\right) \tag{47}$$

This simplified version for the maximum dimensionless height is the one used in Fig. 28. Such naive derivation, however, is the qualitative equivalent of eq. (46) and not the expected maximum, which means it would lead to even lower expected maxima when compared to eq. (47). As shown in Table 5, the average expected $\| \alpha \|$ for all storms in Table 1 according to Miche [56] is of the order of 2.62 (with $\langle s_s \rangle \approx 0.0561$), however, the Miche expectancy model fails for storms 90, 172 and 25. To the keen reader, it is already clear the difference between the upper bounds in Table 5: $\alpha_\infty$ is the ceiling for rogue waves, whereas the other two are empirically derived, reproducing the physics of the data they were extracted from, thus, being vulnerable to unusual sea state conditions. Furthermore, an additional matter calls for attention: the upper bound $\alpha_\infty$ reaches values lower than unity in shallow water wave regime in Fig. 28, which by definition can not happen because the maximum wave height can not be smaller than the average of the tallest wave heights $H_{1/3}$. Therefore, is convenient to define an upper bound for the significant steepness, whose excess will prevent any rogue wave from being formed:

$$\langle s_s \rangle_\infty = \frac{11}{105}\left[1 + \sqrt{\frac{\pi\tau\langle\lambda_2\rangle}{180\,D}.\frac{\langle T_z\rangle^{1/4}}{\eta_{1/3}^4}}\right]\tanh\left(\frac{30\pi D}{22\langle\lambda_2\rangle}\right). \tag{48}$$

However, having the significant steepness smaller than its bound does not guarantee the formation of rogue waves. When dealing with shallow water wave regime in eqs. (46–48), the lack of knowledge on how the *nonlinearity* $\eta_{1/3}$ is affected by finite depth induces a degree of uncertainty in the conjectured models when $\lambda_2/D \to \infty$, as it is expected to have an increased nonlinear behavior in shallow water. Accordingly, one should expect the tails in Figs. 28–29 (whose full finite depth expression is found in Appendix C) to be lowered by 20–30%.

### 5.6. Discussion

It is worth mentioning that this study was mainly observational and the dataset did not include information on the directionality, which can indeed affect the results obtained here. The reader is referred to the most important advancements in respect to experimental [64,86,91] and numerical [83,84] analysis of rogue wave statistics sensibility to directionality and sea state parameters.In this regard, the validity of all equations, both empirical and conjectured, of Section 5 fulfilled the conditions $\|H_{1/3}\| \leq 2\langle H_{1/3}\rangle$ and $\|T_z\| \leq 2\langle T_z\rangle$. It is very unlikely that models such as the one described by eq. (43) will be valid for non-homogeneous wave records due to its strong sensibility to sea state parameters, while the empirical findings of Section 4.1 could be valid even when the conditions described above are not met. Is important to note that the discussion on the homogeneity of sea states and the effect of partitioning a storm into steady-state subgroups on the maximum (absolute) wave height and significant wave height is well-known in the literature [7,21,43,69,88]. However, there are two major distinctions between this approach and ours: First, the parameter $\tau$ is an extensive measure of "homogeneity" and has no clear equivalent in these works, whereas Borgman [7] tries to estimate what effect the variations of sea state has on the Longuet-Higgins [48] distribution, e.g. the integration of the Rayleigh cumulative distribution over a time-varying sea state. Moreover, Borgman [7] does not ascribe any bound to the duration of the sea, e.g. how large the deviations can be so that accuracy is preserved, whereas $\tau$ has an upper bound that defines what is considered "too heterogeneous" to create an accurate explanation of the "uneven distribution" phenomena [74]. Therefore, we do not define homogeneity in the same way, e.g. as a steady-state sea, rather as a particular low interval for $\tau$ (see Fig. 22), such that this work and Borgman [7] answer

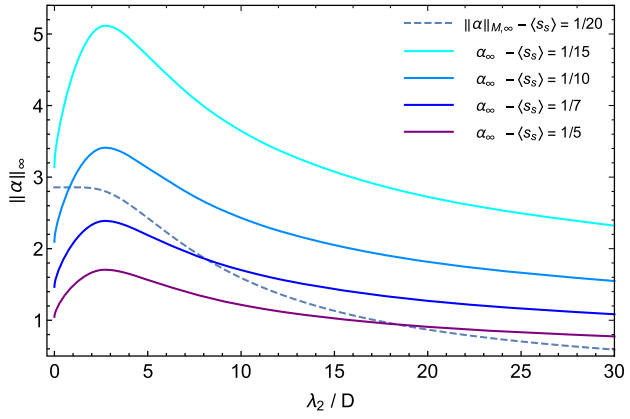


**Fig. 28.** Comparison between the models of eq. (46) and eq. (47) with fixed depth ($D = 100$m), nonlinearity ($\eta_{1/3} = 1.2$) and storm geometry ($\tau = 15$).

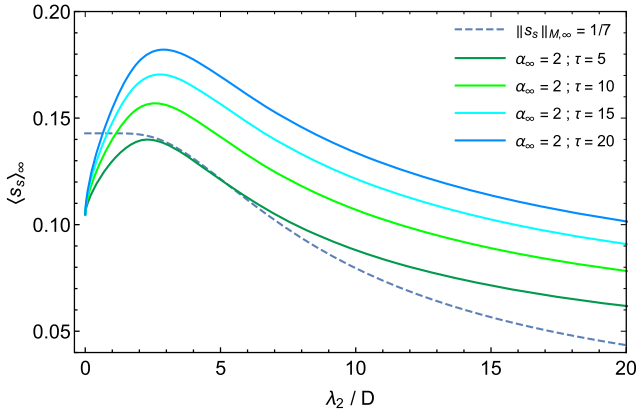


**Fig. 29.** Contrast between the models of eq. (48) and of eq. (17) as a function of $\tau$ and fixed depth ($D = 100$m) and nonlinearity ($\eta_{1/3} = 1.2$).

different questions. Furthermore, when the storm parameter is near $\tau \approx 40$ the model for normalized heights starts to violate the minimum allowed $\|\alpha\| \geq 1$. Further analysis is needed to check the correlation of an exact measure of the curvature of $H_{1/3}(t)$ and $\tau$, but assuming a near linear relation between them, the $\tau \approx 40$ (twice the highest value of the North Alwyn data) estimated limit would imply the bound $\|H_{1/3}\| < 5\langle H_{1/3}\rangle$ as the validity of our models and perhaps a rule for sampling partition. As seen in Fig. 22, a higher $\tau$ is generally associated with a higher variability in $H_{1/3}$ over time, and according to our conjecture, higher variability will affect the predictive models negatively, that is because $\langle\langle \delta_2 \rangle\rangle$ will oscillate vigorously with $\tau$ extremes. However, a very low $\tau$ can also have undesirable effects, as it will increase the maximum $\mathbb{E}_\alpha$ and decrease the bound $\alpha_\infty$, thus making it possible in principle to breach the upper bound. For prediction purposes then, very low or very high $\tau$ can lead to either overpredicting or severely underpredicting rogue wave maximum normalized heights. Lastly, the comparison between Figs. 25 and 26 suggest that the relation in eq. (15), though on average precise, starts to decay in accuracy past the threshold $\alpha > 2.6$, such that all conjectured bounds may need a small correction in a future work with data sets and numerical studies with a statistically significant amount of super-rogue waves. Finally, the reader might reach the conclusion that the values in Table 4 are on average (see Fig. 1) overestimated by 4% if $\alpha$ was normalized by the spectral significant wave height $H_{m0}$ instead of $H_{1/3}$. However, it can be proven that the ratio between the models and observation of Table 4 as well as the exceeding probabilities are left invariant under a corrected normalization (see D.2). Moreover, Massel [54] argues that when $H_{m0}$ differs from $H_{1/3}$, it cannot be used directly to estimate wave height, which is the case of Fig. 1, hence validating our normalization choice.

## 6. Rogue wave crests

As discussed in the first part of this work devoted to wave heights, Haring et al. [33], Forristall [24] and Tayfun [77] among many others have extended Longuet-Higgins's original work to more realistic sea states, some for wave heights and others for wave crests or wave troughs. Yet, there is no universal distribution that explains rogue wave occurrence for an arbitrary sea state, valid for wave heights and crest heights alike [39].

From very early on, it was apparent that RD often overpredicts rogue wave heights and yet severely underpredicts rogue wave crests, an effect that is qualitatively easy to ascribe to nonlinearity, but quantitatively hard to describe. Strikingly, except in the narrow-banded regime, most studies do not find a common structure for both distributions. Therefore, in addition to the challenges on prediction and characterization of storms by sea states in the first part, we also seek to address the discrepancy between crest and height statistics. Here, we build upon the approach described in eq. (1) (see Table 6):

$$\beta = \frac{\mathscr{Z}_c}{H_{1/3}}, \tag{49}$$

where $\mathscr{Z}_c$ is the wave crest. A wave height $H$ and a wave crest $\mathscr{Z}_c$ will be considered rogue if the conditions below are satisfied:

**Table 6**
Proposed definition for a category of waves according to the approximation in eq. (53) and to the wave height definitions in Table 2.

| | Ordinary Wave | Large Wave | Sub-Rogue Wave | Rogue Wave | Super-Rogue Wave |
|---|---|---|---|---|---|
| **Definition II** | $\beta \in (0, 0.6]$ | $\beta \in (0.6, 1]$ | $\beta \in (1, 1.2]$ | $\beta \in (1.2, 1.8]$ | $\beta > 1.8$ |
| **Total Count II** | 324,867 **(91.8%)** | 27,742 **(7.8%)** | 875 **(0.25%)** | 236 **(0.07%)** | 8 **(0.0022%)** |

$$H > \alpha_0 H_{1/3}, \quad \mathscr{Z}_c > \beta(\alpha_0) H_{1/3} \equiv \beta_0 H_{1/3}. \tag{50}$$

Where $\alpha_0$ is the standard rogue parameter introduced by Dean [12] and $\beta_0$ is the rogue crest counterpart. According to Petrova et al. [66], the most widely used coefficients are respectively 2.00 and 1.25. Following Cartwright and Longuet-Higgins [8], Longuet-Higgins [48], the RD exceeding probability for narrow-banded seas ($\alpha = 2\beta$) is expressed as,

$$\mathscr{R}_\beta \equiv \mathscr{R}\left(\mathscr{Z}_c > \beta H_{1/3}\right) = e^{-8\beta^2}, \tag{51}$$

However, crest heights can reach the rogue threshold while the wave height may not at the same time and vice-versa, this being the main source of discrepancy between crest and wave height distributions. Similarly to what we did in Table 1, we lay down the most crucial statistics of rogue wave crest occurrence for the North Alwyn data in Table 7. In fact, the underprediction of rogue wave crests based on Longuet-Higgins [48] is two orders of magnitude worse than for wave heights. Hence, one confronts the problem of explaining why the same distribution slightly overpredicts rogue wave heights but severely underpredicts wave crests. In this regard, nonlinear corrections and spatiotemporal effects have been analyzed in recent works and provide a good guidance on the subject Fedele et al. [23], Laface et al. [45], Romolo et al. [70], 71], Voermans et al. [90], however, a full expression relating $\alpha$ and $\beta$ without absolute heights is not entirely known, see Wolfram et al. [93] for a model where the ratio $\beta/\alpha$ is a function of the significant wave height. Moreover, care should be taken when trying to explain time series data (such as in Table 1) with spatiotemporal distributions [55].

### 6.1. Dimensionless heights conversion

In this section we show how to relate two rogue parameters in accordance to the nomenclature in the introduction. Towards that end, we make use of the *nonlinearity* parameter defined in eq. (11), so that eq. (1) can be rewritten as follows:

$$\left(\frac{H}{H_{1/3}}\right)_{H>0.74H_{1/3}} \approx \left(\frac{1+\eta_{1/3}}{\eta_{1/3}}\right)\left(\frac{\mathscr{Z}_c}{H_{1/3}}\right)_{H>0.74H_{1/3}}, \tag{52}$$

because Longuet-Higgins [48] expects the lowest normalized height of the 1/3 tallest waves to be $(0.5\ln 3)^{1/2} \approx 0.74$, suggesting that the dimensionless wave height can be converted into the dimensionless wave crest in the following manner:

$$\beta(\alpha) \xrightarrow{\alpha>1} \left[\frac{\eta_{1/3}}{1+\eta_{1/3}} + \mathscr{O}\left(\eta_{1/3}, \alpha\right)\right]\alpha. \tag{53}$$

This dependence of the ratio between wave crest and wave height on the *nonlinearity* may partially explain the disparity between Longuet-Higgins's prediction for rogue wave heights and rogue wave crests. This is important because the equivalent alternative would be to fix the probability distribution and make the rogue wave threshold flexible, but this is of course unreasonable. Moreover, we see that for a typical $\eta_{1/3} \approx 1.25$ we find $\beta_0/\alpha_0 \approx 1.11$, so that we choose $\beta_0 = 1.2$ as the threshold closest to the most common definition that is consistent to the above relation. Up to first order, using the inverse transformation the normalized crest obeys the *nonlinearity* adjustment:

$$\mathscr{R}^*_\beta \approx \exp\left\{-2\beta^2\left(\frac{1+\eta_{1/3}}{\eta_{1/3}}\right)^2\right\}. \tag{54}$$

Interestingly, Longuet-Higgins has been adjusted to the nonlinear behavior by including finite bandwidth effects [51], such that the adjusted distributions of Fig. 30 tend to oscillate around the original probability [48], but only the one adjusted by the *nonlinearity* can

**Table 7**
Summary of the North Alwyn data [74] with storms ordered from lower (top) to higher (bottom) $\mathcal{N}_{\beta=1.2}$. The last column shows each storm maximum observed dimensionless crest height $\|\beta\|$. Numbers in bold font show Longuet-Higgins's prediction, including the *nonlinearity*-adjusted version in eq. (56), for the total number of rogue waves, maximum $\beta$ and return period.

| Storm ID | $\beta>0$ | $\beta>1$ | $\beta>1.2$ | $\beta>1.35$ | $\beta>1.5$ | $\beta>1.8$ | $\|\beta\|$ | $\mathcal{N}_{\beta=1.2}$ |
|---|---|---|---|---|---|---|---|---|
| 172 | 23,591 | 112 | 35 | 17 | 5 | 1 | 2.46 | 674 |
| 29 | 13,610 | 59 | 17 | 10 | 7 | 1 | 1.86 | 801 |
| 90 | 44,867 | 207 | 55 | 27 | 13 | 2 | 2.11 | 816 |
| 149 | 52,766 | 199 | 63 | 33 | 14 | 2 | 1.87 | 838 |
| 28 | 22,155 | 57 | 15 | 5 | 3 | 2 | 2.03 | 1477 |
| 23 | 25,068 | 76 | 12 | 2 | 0 | 0 | 1.42 | 2089 |
| 146 | 15,109 | 33 | 6 | 2 | 0 | 0 | 1.42 | 2518 |
| 132 | 45,056 | 126 | 17 | 4 | 0 | 0 | 1.47 | 2650 |
| 25 | 16,896 | 44 | 6 | 2 | 0 | 0 | 1.46 | 2816 |
| 26 | 27,774 | 53 | 8 | 3 | 0 | 0 | 1.40 | 3472 |
| 127 | 14,845 | 29 | 3 | 1 | 0 | 0 | 1.36 | 4948 |
| 27 | 20,379 | 45 | 4 | 2 | 1 | 0 | 1.56 | 5095 |
| 195 | 9875 | 23 | 1 | 0 | 0 | 0 | 1.22 | 9875 |
| 124 | 21,737 | 56 | 2 | 0 | 0 | 0 | 1.33 | 10,869 |
| Total | 353,728 | 1119 | 244 | 108 | 43 | 8 | 2.46 | 1450 |
| Total × $\mathscr{R}_\beta$ | - | **119** | **4** | **0.16** | **0.005** | **0.000002** | **1.29** | **100,710** |
| Total × $\mathscr{R}^\star_\beta$ | - | **3,018** | **585** | **146** | **32** | **1** | **1.67** | **605** |

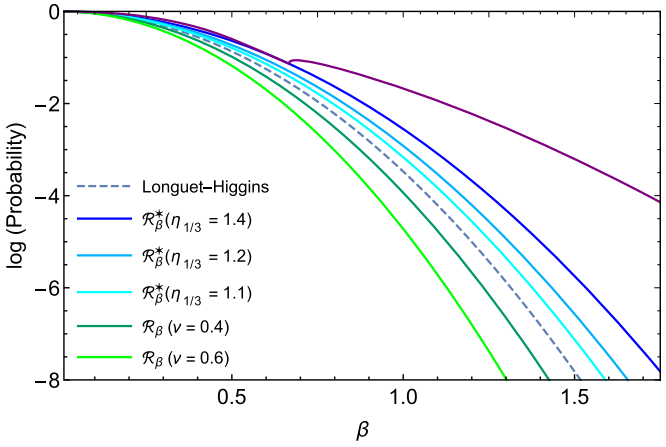


**Fig. 30.** Comparison between the standard wave crest Rayleigh distribution of eq. (51), the one adjusted by the *nonlinearity* $\eta_{1/3}$ and the bandwidth corrected due to [51]. The curve in purple shows the second-order correction to $\beta/\alpha$ in eqs. (53) and (55) for $\eta_{1/3}=1.4$. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

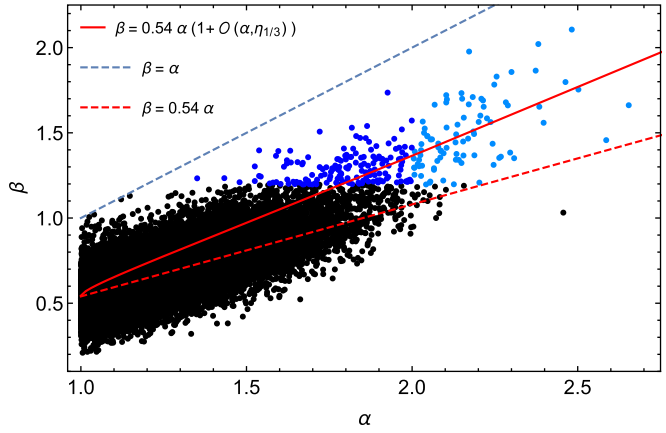


**Fig. 31.** How the ratio between crest height and wave height varies for all fourteen storms waves fulfilling the condition $\alpha\geq 1$. The lighter shade of blue denotes simultaneous rogue wave heights and crests, whereas the darker shade shows rogue wave crests that has $\alpha<2$ and black dots are the remaining waves. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

increase the standard probability (see Fig. 30). On the other hand, the data shows that the first order term in eq. (53) will underestimate how the nonlinear character of the waves grows, as shown in Fig. 31. However, for the average among tall waves the first order seems accurate: based on Table 3 we estimate that the fourteen-storm average is $\eta_{1/3}\approx 1.178$, such that we would have $\beta\approx 0.541\alpha$, whereas observation showed that for all waves with $\alpha\geq 1$ it reads $\beta\approx 0.539\alpha$. The correction for higher normalized heights necessary to fit the observation is approximately described by:

$$\left\langle\frac{\beta}{\alpha}\right\rangle \approx \frac{\eta_{1/3}}{1+\eta_{1/3}}\left[1+\frac{2\eta_{1/3}\sqrt{\alpha-1}}{7+2\sqrt{\alpha-1}}\right]\equiv \Im_\beta\ ,\ \forall\alpha>1\ . \tag{55}$$

Taking into account the above correction, we rewrite the distribution, finding (the value of 1.5 is an anticipation to the iteration, e.g. the measured value in Fig. 31, since we can no longer use $\alpha=2\beta$ in eq. (53)):

$$\mathscr{R}^\star_\beta \approx \exp\left\{-2\beta^2\left(\frac{1+\eta_{1/3}}{\eta_{1/3}}\right)^2\left[1+\frac{2\eta_{1/3}\sqrt{1.5\beta-1}}{7+2\sqrt{1.5\beta-1}}\right]^{-2}\right\}\ . \tag{56}$$

Notice that the ratio $\Im_{\beta\to\eta_{1/3}}/(1+\eta_{1/3})$ for $\alpha\leq 1$ (if we took the real part of $\Im_\beta$ for $\alpha\leq 1$, the typical value would be $\Im_\beta\approx 1.05$), or alternatively when $\beta\leq 2/3$. The keen reader will observe that when $\eta_{1/3}=1$ we recover the equivalency of eqs. (51) and (2), i.e. the perfectly narrow-banded sea has no vertical asymmetry. On the other hand, in the limit $(\alpha,\ \beta)\to\infty$ we reach the average $\Im_{\beta\to}1$. Notice, however, that some waves can reach $\beta/\alpha\approx 1$, which is not the same as the average of a large group of waves with $\Im_\beta=1$. Nonetheless, as demonstrated by eq. (46), this limit is not reachable because the maximum normalized height is finite and denoted by $\alpha_\infty$, so that the approximate model in eq. (53) is mathematically bounded and at the same time avoids unrealistic scenarios.

### 6.2. Physical bounds & rogue wave patterns

Besides the significant underprediction of wave crests in the range $\beta\geq 1$ by Longuet-Higgins [48], one outstanding distinction between Table 1 and Table 7 is highlighted: while a few storms displayed lower return period for rogue wave heights than predicted by RD and most storms had a higher return period, for wave crests all storms featured a

return period much lower than Longuet-Higgins [48]. However, as discussed by Karmpadakis et al. [39], splitting wave data into ocean state homogeneous groups provides a better insight into the usefulness of distributions. To this extent, we plotted the empirical distributions according to the selected groups of Table 8. The rogue wave likelihood variability is much greater for wave crests (see Fig. 32). Interestingly, the selected groups start to display diametrical statistics in the sub-rogue wave zone $\beta \geq 1$ and $\alpha \geq 1.8$, providing a stronger qualitative argument for the rogue wave thresholds than arbitrary statistical rarity alone [12], in addition to being the threshold where the RD considerably departs from observation. Following the proposed evolution of wave crest to wave height ratio in eq. (55), we can expect the approximate relation:

$$\langle \| \beta \| \Big/ \| \alpha \| \rangle \left( \frac{\eta_{1/3}}{1+\eta_{1/3}} \right) \left[ 1 + \frac{2\eta_{1/3}\sqrt{\| \alpha \| - 1}}{7 + 2\sqrt{\| \alpha \| - 1}} \right] \tag{57}$$

Such an approximation is in fact confirmed by observation, as shown in Fig. 33. Notice however, that the oscillation around the red curve representing eq. (57) is due to the oscillation in *nonlinearity* $\eta_{1/3}$, as we have used the fourteen storm mean instead of each storm value, such that making a distinction between storms with highest and lowest *nonlinearity* shows the qualitative precision of eq. (57), e.g. higher $\eta_{1/3}$ will increase the group average $\langle \|\beta\|/\|\alpha\| \rangle$. In fact, as shown in Fig. 34, the average ratio for Group I was $\langle \|\beta\|/\|\alpha\| \rangle \approx 0.751$ whereas the combined remaining storms had 0.666, whose predictions made by eq. (57) are respectively 0.732 and 0.686.

Furthermore, using the definitions of the introduction and comparing with Section 4 we introduce the crest effective steepness:

$$Z_{\text{Eff}} = \frac{\mathscr{Z}_{c,i}}{\langle \langle \lambda_2 \rangle \rangle} \equiv \left( \frac{\beta}{\alpha} \right) S_{\text{Eff}}. \tag{58}$$

Moreover, the main empirical finding for wave heights has its crest counterpart (see Fig. 35):

$$\| \beta \| \cdot \langle s_s \rangle \cong 0.96 \| Z_{\text{Eff}} \|, \tag{59}$$

To be precise, the wave height result had $0.958 \pm 0.061$ while the above returns $0.960 \pm 0.074$. Without prior knowledge of eq. (53), it would not be possible to anticipate the crest equivalent of eq. (15), such that the previous relation is only completely intelligible from the perspective of Fig. 33 which assures us that eq. (53) applies to storm maxima. The relationship between normalized effective steepness and the former is (Fig. 36):

$$\frac{Z_{\text{Eff}}}{\langle Z_{\text{Eff}} \rangle_{1/3}} \approx 1.55\,\beta, \tag{60}$$

confirming that wave crests, like wave heights, increase their effective steepness with higher normalized heights.

### 6.3. Statistical theory

The striking similarity between the relations in eq. (15–16) and its wave height counterparts is not the result of mere coincidence. In fact, the keen reader will notice that the last approximation comes from the fourteen storm average $H_{1/3} \approx \langle \alpha / \beta \rangle \mathscr{Z}_{c\,1/3} \approx 1.85$, such that having full knowledge of eq. (53) would suffice to make an accurate prediction for the r.h.s of the above relation, becoming $1.85 \times 0.836 \approx 1.547$. However, while the origin of the 1.546 coefficient is understood, is not clear from a first look where the 0.836 comes from. Quantitatively, this is related to Longuet-Higgins's estimate for the significant wave height. Let us first prove that the wave height counterpart of eq. (16) holds. We start with the basic statistical definition for the Longuet-Higgins's probability density:

$$f_{\mathscr{R}}(\alpha) = \frac{dF_{\mathscr{R}}(\alpha)}{d\alpha} = \frac{d}{d\alpha}\Big[ 1 - \mathscr{R}(\alpha) \Big] = 4\alpha\, e^{-2\alpha^2}, \tag{61}$$

where $F_{\mathscr{R}}(\alpha)$ denotes the RD cumulative distribution. Hence, the mean normalized wave height of all waves above $\alpha \geq 3/4$ reads:

$$\langle \alpha \rangle_0 = \frac{\int_{3/4}^{+\infty} 4\alpha^2 e^{-2\alpha^2} d\alpha}{\int_{3/4}^{+\infty} 4\alpha e^{-2\alpha^2} d\alpha} \approx 1.008, \tag{62}$$

where the subscript implies the measure is calculated for waves that consist in the 1/3 wave group, e.g. the set $\{H;\, \langle H \rangle = H_{1/3}\}$ normalized becomes $\{\alpha;\, \langle \alpha \rangle = 1\}$. The exact lower bound for the integral should have been $\alpha = 0.7397$, confirming the estimate of eq.(52). However, another useful estimate is the average normalized height for the group $\alpha \geq 1$ appearing in the empirical relations (see Appendix E):

$$\begin{aligned} \langle \alpha \rangle_{1/3} &= \frac{\int_1^{+\infty} 4\alpha^2 e^{-2\alpha^2} d\alpha}{\int_1^{+\infty} 4\alpha e^{-2\alpha^2} d\alpha} = 1 + e^2 \sqrt{\frac{\pi}{8}}\, erfc\left(\sqrt{2}\right), \\ &\cong 1 + \frac{5}{6\sqrt{2\pi}} \sqrt{\frac{\pi}{8}} \approx 1.208. \end{aligned} \tag{63}$$

Put into perspective, the above theoretical prediction is quite accurate, since for storm 29 this average was 1.203 and overall $1.197 \pm 0.005$. Therefore, we find the comparative relation:

$$\frac{\langle \alpha \rangle_{1/3}}{\langle \alpha \rangle_0} \equiv \frac{\langle H \rangle_{1/3}}{H_{1/3}} \approx \frac{6}{5}. \tag{64}$$

Consequently, one can easily show that the following holds:

$$\frac{S_{\text{Eff}}}{\langle S_{\text{Eff}} \rangle_{1/3}} = \frac{H}{\langle H \rangle_{1/3}} \approx \frac{5H}{6H_{1/3}} \equiv \frac{5}{6}\alpha, \tag{65}$$

Thus,

$$\begin{aligned} \frac{Z_{\text{Eff}}}{\langle Z_{\text{Eff}} \rangle_{1/3}} &= \frac{Z_{\text{Eff}} \cdot \langle S_{\text{Eff}} \rangle_{1/3}}{\langle S_{\text{Eff}} \rangle_{1/3} \cdot \langle Z_{\text{Eff}} \rangle_{1/3}} = \left[ \frac{\beta S_{\text{Eff}}}{\alpha \langle S_{\text{Eff}} \rangle_{1/3}} \right] \cdot \frac{\langle H \rangle_{1/3}}{\langle \mathscr{Z}_c \rangle_{1/3}}, \\ &\approx \frac{5}{6}\beta \cdot \left\langle \frac{\alpha}{\beta} \right\rangle_{1/3} \approx \frac{5(1+\eta_{1/3})}{6\eta_{1/3}}\beta, \end{aligned} \tag{66}$$

which for the fourteen storm with $\eta_{1/3} \approx 1.178$ becomes $1.541\beta$. Nevertheless, the theoretical explanation for the empirical findings through Longuet-Higgins [48] statistics by no means indicates the validity of its distribution and expectancy models for the upper tail. *Au contraire*, the above rational confirms that the exceeding statistics of the RD model works perfectly for $\alpha = 1$. Additionally, distributions such as

**Table 8**
North Alwyn data set main sea parameters arranged by groups of storms. Since we assumed narrow-banded sea for the conversion of crest to height distributions in eq. (11), for Forristall [25] the conversion to wave crests is $U_{00,\beta} = 2^{U_1} \cdot U_{00}$, such that $\mathscr{F}^*_\beta = \exp[-U_{00,\beta}\beta^{U_1}]$.

| Storm Group | | $\mathscr{N}_{\beta=1.2}$ | $\langle H_{1/3} \rangle$ | $\langle T_z \rangle$ | $\langle \lambda_{1/3} \rangle$ | $\eta_{1/3}$ | $\langle \nu \rangle$ | $S_1^{-1}$ | 1000 Ur | $U_{00}$ | $U_{00,\beta}$ | $U_1$ | $\aleph_1$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **172-29-90-149** | **(I)** | 782 | 6.06 | 8.28 | 161.6 | 1.237 | 0.736 | 25.94 | 1.442 | 1.850 | 7.050 | 1.930 | 1.005 |
| **14-Storm Mean** | **(II)** | 1450 | 5.71 | 8.14 | 166.2 | 1.178 | 0.729 | 28.06 | 1.486 | 1.859 | 7.103 | 1.934 | 1.085 |
| **28-23-146-132** | **(III)** | 2184 | 5.29 | 7.77 | 161.2 | 1.158 | 0.790 | 29.23 | 1.288 | 1.866 | 7.150 | 1.938 | 1.071 |
| **127-27-195-124** | **(IV)** | 7697 | 6.27 | 8.80 | 185.4 | 1.155 | 0.621 | 27.06 | 1.986 | 1.855 | 7.083 | 1.933 | 1.235 |

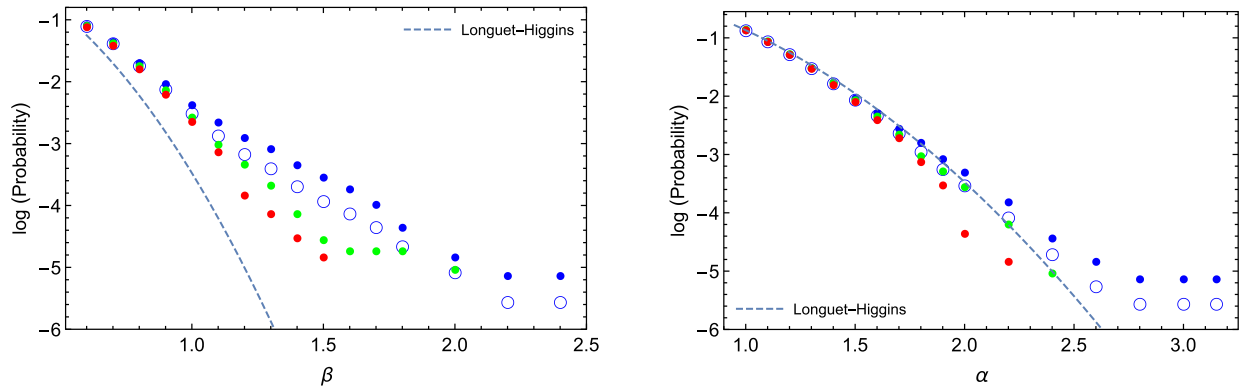


**Fig. 32.** Empirical exceeding probabilities of storm groups I (blue dots), II (blue circles), III (green dots) and IV (red dots) of Table 8. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

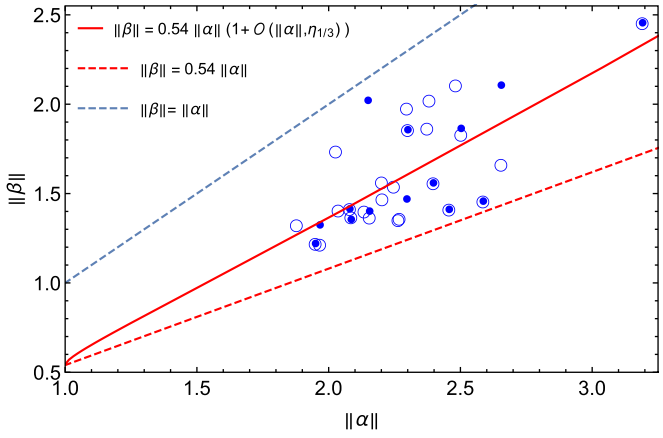


**Fig. 33.** Scatter plot between observed maximum wave crests and wave heights within a storm (blue dots) or within a 30-hour period (blue circles). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

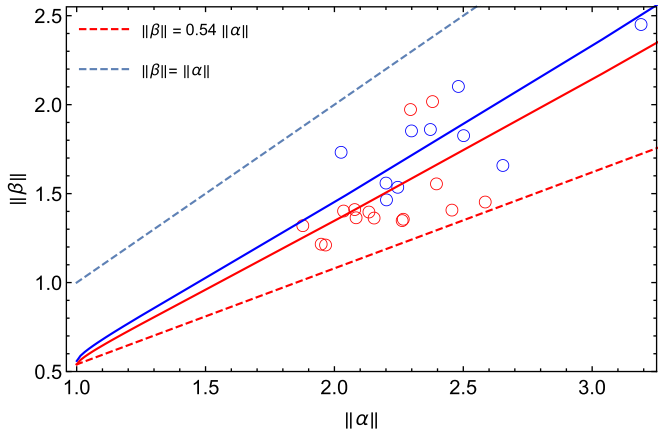


**Fig. 34.** Scatter plot between maximum wave crests and wave heights for Group I (blue circle and line) and the remaining storms combined (red). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

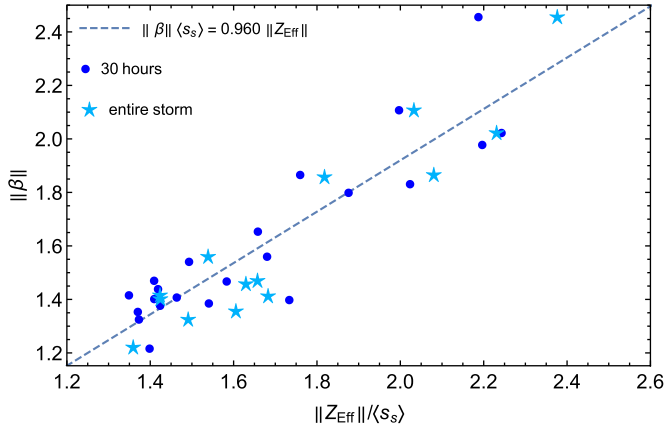


**Fig. 35.** The relationship between maximum normalized crest heights, maximum effective steepness and average significant steepness.

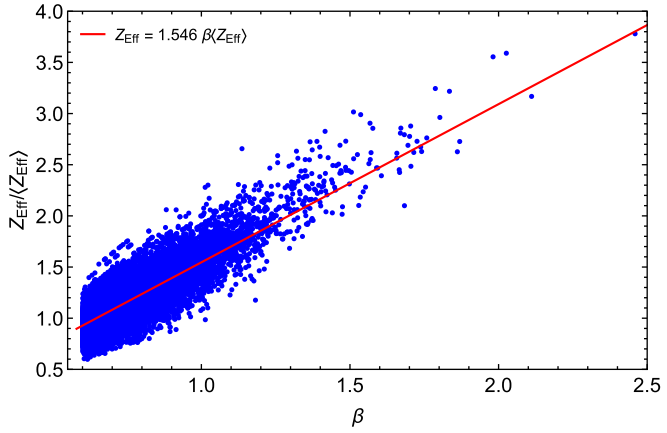


**Fig. 36.** Growth of relative effective crest steepness with normalized crest heights for all waves fulfilling both $\alpha > 1$ and $\beta > 0.5$.

Tayfun [77] or Haring et al. [33], despite providing a much higher rogue wave likelihood than RD, will provide very similar $\langle\alpha\rangle$ because their probability density deviation from Longuet-Higgins [48] is not large at $\alpha < 2$ (see Fig. 37 and Table 9). Though not of any qualitative advantage, Longuet-Higgins [48] has a clear numerical advantage by dealing with less cumbersome error integrals. Furthermore, we can generalize the previous equation:

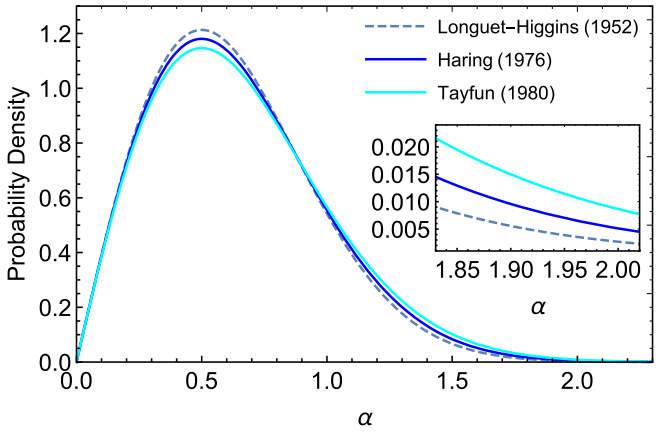


**Fig. 37.** Probability densities of the respective exceeding probability distributions discussed in Section 3 with up close look at rogue waves.

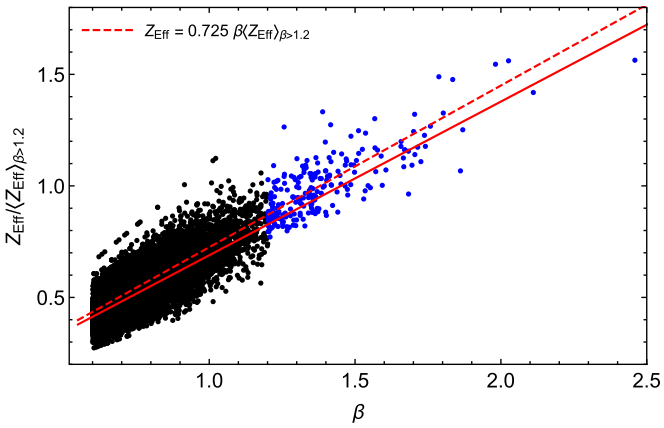


**Fig. 38.** Best fit (dashed line) versus the predicted relative effective steepness growth with normalized crests by Longuet-Higgins [48] with rogue crests highlighted (blue dots). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

**Table 9**
Major ensemble averages of normalized heights according to each probability density [33,48,77].

| | $\langle\alpha\rangle_{\alpha>0}$ | $\langle\alpha\rangle_{\alpha>0.74}$ | $\langle\alpha\rangle_{\alpha>1}$ | $\langle\alpha\rangle_{\alpha>2}$ | $\langle\alpha\rangle_{\alpha>3}$ |
|---|---|---|---|---|---|
| $\mathscr{R}(\alpha)$ | 0.627 | 1.000 | 1.211 | 2.118 | 3.081 |
| $\mathscr{H}(\alpha)$ | 0.641 | 1.018 | 1.229 | 2.139 | 3.102 |
| $\mathscr{T}(\alpha)$ | 0.656 | 1.036 | 1.249 | 2.159 | 3.122 |

$$\frac{Z_{Eff}}{\langle Z_{Eff}\rangle_{\alpha>\widetilde{\alpha}}} = \frac{\beta\int_{\alpha}^{\infty}\mathscr{R}(\alpha)\,d\alpha}{\int_{\alpha}^{\infty}\alpha\mathscr{R}(\alpha)\,d\alpha}\left(\frac{1+\eta_{1/3}}{\eta_{1/3}}\right)\left[1+\frac{2\eta_{1/3}\sqrt{\widetilde{\alpha}-1}}{7+2\sqrt{\widetilde{\alpha}-1}}\right]^{-1}. \tag{67}$$

Nonetheless, one may ask why choose the lower bound for the normalized wave height instead of the crest height. This is done because of the fluctuation in the ratio $\beta/\alpha$ for the thresholds. While we have used $\beta_0 = 0.6\alpha_0$ for the thresholds in Table 6, this choice is slightly higher than predicted by eq. (53) and starts to slightly underpredict it for higher thresholds. In other words, while for the set $\{\alpha;\,\langle\alpha\rangle = 2\}$ we have $\beta \approx 0.7\alpha$ (17% higher than the chosen threshold $\beta_0 = 0.6\alpha_0 = 1.2$), for the set $\{\alpha;\,\langle\alpha\rangle = 1\}$ one finds $\beta \approx 0.54\alpha$ and overall the relation $\beta = 0.6\alpha$ seems efficient, though suggesting that the rogue crest criteria should be $\beta > 1.4$, $\beta > 2.2$ for super-rogue waves and $\beta > 0.5$ for ordinary waves instead. Nevertheless, choosing the threshold with numerical accuracy does not change the predictive results of standard distributions, thus, the replacement of the fixed wave crest thresholds are not of uttermost importance provided they are good estimates for the $\alpha - \beta$ empirical relation. In fact, if we perform the integration of the RD crest density $16\beta e^{-8\beta^2}$ for the bound $\widetilde{\beta} = 0.42$, equivalent of 57% of the wave height counterpart $\widetilde{\alpha} = 0.74$, returns $\langle\beta\rangle_{0.42} \approx 0.539$, which is exactly the ratio between crest and wave heights at $\alpha = 1$. Therefore, for the rogue wave set and average *nonlinearity* of Group II, based on Longuet-Higgins [48] we anticipate (using the $\alpha = 1.5\beta$ as in eq. (56)):

$$\frac{Z_{Eff}}{\langle Z_{Eff}\rangle_{\alpha>2}} \approx \frac{13}{6\pi}\beta \quad ; \quad \frac{S_{Eff}}{\langle S_{Eff}\rangle_{\alpha>2}} \approx \frac{8}{17}\alpha. \tag{68}$$

Not surprisingly, the prediction has shown to be accurate up to a 3–6% deviation, as depicted by Figs. 38–39. However, a clear trend appears: when the exceeding probability strongly deviates from the empirical for rogue waves, the predicted effective steepness ratio error is much higher than 1%. For instance, for Group II rogue waves, we have that [48] is a better model than Haring et al. [33] or Tayfun [77], and their predicted effective ratios (calculated as $\langle\alpha\rangle_{\alpha>0.74}/\langle\alpha\rangle_{\alpha>2}$ from Table 9) are respectively 3.5%, 4.4% and 5.2% higher than observed (Fig. 39). For rogue wave crests, however, the picture is inverted because Tayfun [77]

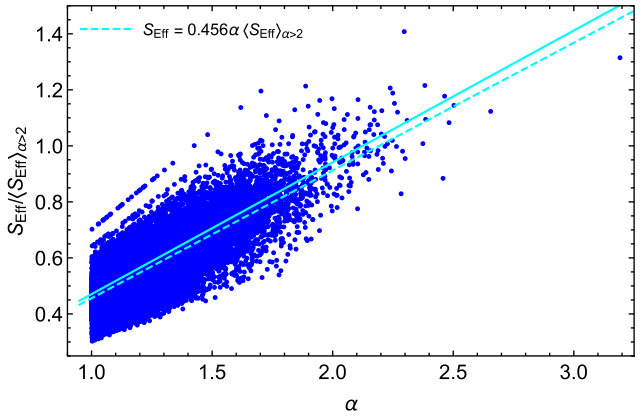


**Fig. 39.** Best fit (dashed line) versus the predicted relative effective steepness growth with normalized heights by Longuet-Higgins [48].

has the upper hand among the three distributions, such that their effective ratios are now respectively 4.6%, 3.8% and 3.0% lower than the best fit observed (Fig. 38), estimated as $1.4653\langle\alpha\rangle_{\alpha>0.74}/\langle\alpha\rangle_{\alpha>2}$. Notice also, in accordance with the ratio approach, a higher deviation from the empirical distribution in wave crests means a higher deviation from the best fit for the effective steepness ratio. Accordingly, these ratios for crest and wave heights might be an additional test for probability distributions in addition to the expected maximum, so that is suggested that an ideal exceeding probability will maintain an error smaller than 1%. In addition, this statistical exercise has showed that is possible to predict exactly how much "effectively" steeper the ensemble of rogue waves will be compared to any other ensemble of waves.

### 6.4. Rogue wave patterns

In Section 5 we discussed what combination or single sea state parameters helped understand the occurrence likelihood of rogue wave heights. Here we perform a similar analysis (see Fig. 40) for wave crests. Table 8 highlights an increase in return period from groups I to IV, with a especially large gap between groups III and IV. The extensive sea parameters $\langle H_{1/3}\rangle$ and $\langle T_z\rangle$ show a negative correlation as it shows a decreasing trend, but the correlation is small because for Group IV it increases whereas it should have decreased even further in order to control the wave statistics. The significant wavelength $\langle\lambda_{1/3}\rangle$, however, shows an oscillation, as it shows a constant change between increase and decrease. This oscillation pattern is also observed for the significant

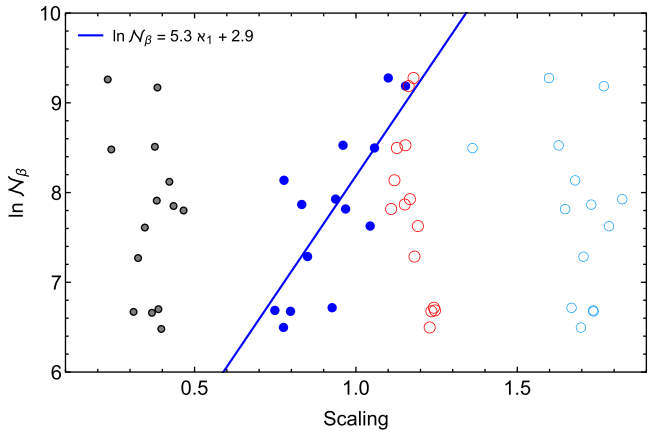


**Fig. 40.** Scatter plot between the growing wave crest return period of Table 1 and candidate sea parameters as in Section 5: the correlation coefficients between $\mathcal{N}_\beta$ and spectral parameters $30\langle s_s\rangle$ (blue circles), $\langle\nu\rangle$ (gray dots), $\eta_{1/3}$ (red circles), $\aleph_1$ (blue dots) are respectively -0.256, -0.386, -0.402 and 0.707. As for done earlier for wave heights the combined parameter $\aleph_1$ shows the highest correlation and is slightly higher for wave crests. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

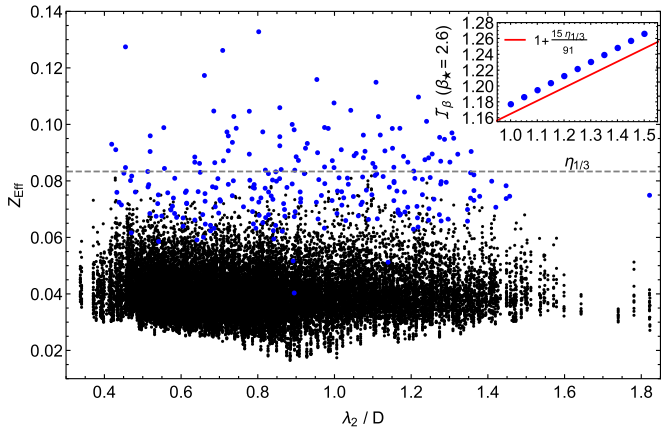


**Fig. 41.** Effective crest steepness of all waves taller than $0.6H_{1/3}$ (black dots) and rogue crest waves (blue dots) versus the crest adjusted Miche's limit (grey dashed line) and the North Alwyn estimated limit of 0.14 from eq. (70). Inner figure shows the accuracy of the approximation (red line) for $\mathcal{I}_\beta$ integral (blue dots) computed numerically, with a deviation growing from 1.1% at narrow-banded seas ($\eta_{1/3} = 1$) to 1.5% at highly nonlinear seas ($\eta_{1/3} = 1.5$). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

steepness and the Ursell number. Thus, according to Table 8, only the parameters $\langle\eta_{1/3}\rangle$ and $\aleph_1 = \langle\lambda_{1/3}\rangle/\eta_{1/3}D$ can explain the observation. However, Fig. 40 rules out the former mathematically, in addition to the undesirable feature of inducing a higher likelihood for rogue wave formation in shallow water where the *nonlinearity* is expected to grow.

### 6.5. Conjectured bounds

Since the building blocks towards the suggestion of the conjectured bounds have been confirmed for wave crests, especially the strong correlation between an increasing return period $\mathcal{N}_\beta$ and $\aleph_1$, we can repeat most steps of sections 4 and 5 by either applying the term $\mathfrak{S}_{\alpha,\beta}$ or a more iterative model depending on the same term. In fact, using eq. (57), we can write the lower bound:

$$Z_{\text{Eff}} \leqslant \frac{22}{105}\,\mathfrak{S}_\beta\left[1-\langle\langle\delta_2\rangle\rangle\right]\tanh\left(\frac{30\,\pi D}{22\langle\langle\lambda_2\rangle\rangle}\right). \tag{69}$$

Notice that the above lower bound, which applies as well for wave heights, is the equivalent of Miche [56] limit for individual steepness, as it covers nearly 99% of all waves and the expression for the expected maximum $\mathbb{E}(S_{\text{Eff}})$ that varies from storm to storm is typically 2% higher than the lower bound for $S_{\text{Eff}}$. In fact, using eq. (53) we can estimate Miche's limit for crests to be 1/12, whereas the full expression in eq. (69) is of the order of $22/105\cdot 3/5\cdot 2/3 \approx 1/11.9$ in deep water. Moreover, the above expression is problematic because both l.h.s and r.h.s depend on $\beta$ and we want to model a limit for that. Therefore, to avoid recurrent iterations we integrate the comparative term,

$$\begin{aligned}\mathfrak{S}_\beta &\approx \frac{1}{\beta_\star}\left(\frac{\eta_{1/3}}{1+\eta_{1/3}}\right)\int_0^{\beta_\star}\left[1+\frac{2\eta_{1/3}\sqrt{1.5\beta-1}}{7+2\sqrt{1.5\beta-1}}\right]d\beta\,,\\ &\approx \left(1+\frac{3\eta_{1/3}}{7\beta_\star}\right)\left(\frac{\eta_{1/3}}{1+\eta_{1/3}}\right)\equiv\mathcal{I}_\beta\left(\frac{\eta_{1/3}}{1+\eta_{1/3}}\right),\end{aligned}$$

with a 1–2% precision for a typical $\beta_\star \approx 2.6$ (this is related to the typical $\alpha_\star \approx 3\alpha_\infty/4 \sim 3.5$, see Table 5). The limit for the effective crest steepness becomes:

$$Z_{\text{Eff}} < \frac{22}{105}\left(1+\frac{\eta_{1/3}}{6}\right)\left(\frac{\eta_{1/3}}{1+\eta_{1/3}}\right)\tanh\left(\frac{30\,\pi D}{22\langle\langle\lambda_2\rangle\rangle}\right). \tag{70}$$

Applied to the North Alwyn dataset, it bounds the *nonlinearity* $Z_{\text{Eff}} < 0.14$. Fig. 41 shows that more crest rogue waves exceed Miche's limit than for wave heights while being bounded by the above result, an indication of nonlinearity and a possible consequence from Longuet-Higgins's severe crest underprediction. Nevertheless, the reader should have understood by now that the $Z_{\text{Eff}}$ bound is a result of an iteration over $\beta_\star$, which differs from the wave height counterpart that has no such procedure. In other words, in the same manner that rogue wave crest empirical distribution is more "volatile" than wave heights, so is their estimated bound for effective steepness. In addition, the previous equation is valid only for $\eta_{1/3} \leqslant \sqrt{6}$, in order to not imply $\beta \geq \alpha$. Realistically, however, such level of *nonlinearity* is not achievable by the majority of ocean states.

Naturally, the next step would consist of gathering all results in this section their counterparts for wave heights and work out the expected maximum model. Following the exact same steps and bearing in mind eqs. (55) and (70), as well as using $\beta_\star \approx 3\alpha_\star/4$ in accordance with eq.

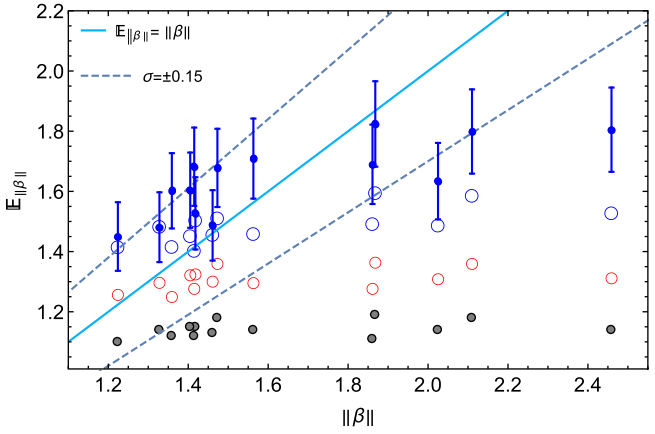


**Fig. 42.** Scatter plot between the $\beta$ model for the expected maximum normalized height (blue dots and bars) versus the observed maxima in Table 1 and the crest height predictions based on RD as grey circles, red circles for Tayfun [77] and blue ones for the *nonlinearity*-adjusted RD (see Appendix D). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

(55), we obtain:

$$\mathbb{E}(\,\|\,\beta\,\|\,) \equiv \mathbb{E}_\beta \cong \mathbb{E}_\alpha\left(1+\frac{4\eta_{1/3}}{7\mathbb{E}_\alpha}\right)\left(\frac{\eta_{1/3}}{1+\eta_{1/3}}\right). \tag{71}$$

As it could have been anticipated by the combination of the trend in Fig. 31 and the wave height equivalent of Fig. 42, the expected maximum crest height follows the same trend as for heights, following the growing observed maximum $\beta$ at lower values with a growing deviation at much larger values bordering super-rogue waves, which again suggests the necessity of adding a higher-order term effective for this ensemble. However, the success of the previous equation, besides doubling the correlation coefficient of the discussed distributions, is its shape $\mathbb{E}_\beta \sim \sqrt{\|\,\beta\,\|}$, albeit not perfectly following $\mathbb{E}_\beta = \|\,\beta\,\|$, consists of a much better regime than the typical straight line $\mathbb{E}_\beta \sim k \in \mathbb{R}_+$ of Tayfun [77] and Longuet-Higgins [48]. In other words, the latter models have a mid-range correlation but of random origin, e.g. there is no causality behind some of its accurate predictions, because it assigns maxima to the number of waves in a storm. The higher-order term is even more needed in Fig. 43 when we use the results of $\mathbb{E}_\alpha$ transformed to crests by the fixed $\mathfrak{S}_\beta = 0.6$. At lower observed maximum $\beta$ we see that the rogue wave definition fixed rate $\beta = 0.6\alpha$ works well (notice that this case is similar to set $\mathscr{I}_\beta = 1$) while dropping considerable accuracy for the five super-rogue wave crests. Interestingly, however, the accuracy of the $\beta$ model is significantly higher for the super-rogue wave crest ensemble in comparison to wave heights, which points to the necessity of a term similar to $\mathscr{I}_\beta$ as a function of the significant steepness for $\mathbb{E}_\alpha$.

Though one could attempt to find the second-order missing term in $\mathfrak{S}_\beta$, its more likely that a correction in $\langle\langle\delta_2\rangle\rangle$ would fix the deviation for large super-rogue wave heights and crests at once, as they both seem to follow the same trend. Likewise, the same is expected for the effective crest steepness in comparison with the wave height counterpart. This is so because of the accuracy of $\mathfrak{S}_\beta$. Likewise, for the effective steepness we have $\mathbb{E}_Z = \mathbb{E}_S\mathscr{I}_\beta\cdot\eta_{1/3}/(1+\eta_{1/3})$, shown in Fig. 44.

Lastly, we can compute the upper bound for the wave crests. However, the term $\mathscr{I}_\beta$ as in eq. (71) will rapidly increase in the vicinity of the shallow water wave regime, such that when plotting over a wide range of $\lambda_2/D$ we choose $6\mathscr{I}_\beta = 6+\eta_{1/3}$. Plotted in Fig. 45, we see that contrary to wave heights (green curves in the inner figure), there is an overlap between curves with varying *nonlinearity*, which goes back to the "volatility" feature of wave crests in comparison with wave heights. Moreover, such behavior is likely the reason for the swapping of storm IDs in both order of lower return period and maximum observed crest heights.

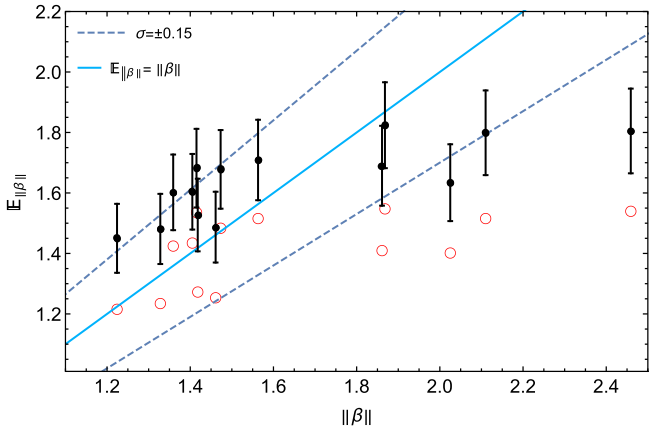


**Fig. 43.** Scatter plot between the $\beta$ model for the expected maximum normalized height (blue dots and bars) and the fixed $\mathbb{E}_\beta = 0.6\mathbb{E}_\alpha$ (red circles) versus observation with deviation centered at $\mathbb{E}_\beta = \|\,\beta\,\|$. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

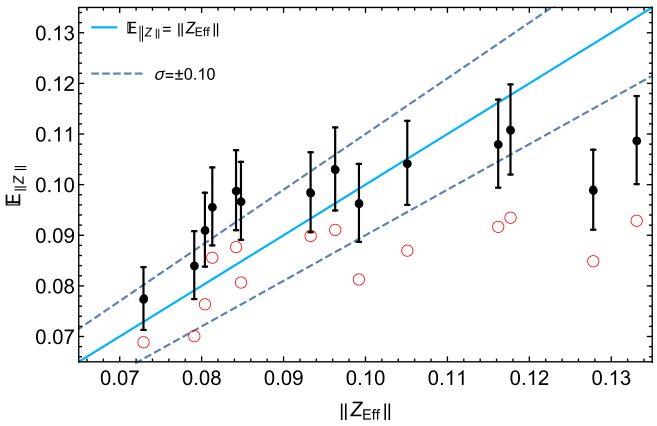


**Fig. 44.** Effective crest steepness equivalent of Fig. 44 with red circles denoting the typical relation $\mathfrak{S}_\beta = 0.6$ of Table 2. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

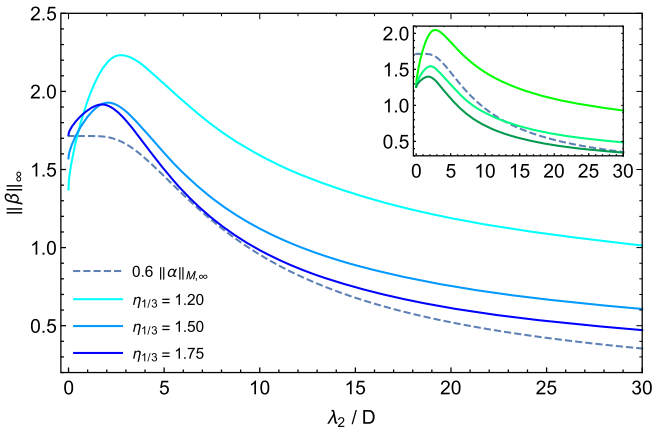


**Fig. 45.** Upper bound for wave crests with fixed steepness ($\langle s_s\rangle = 1/10$), storm geometry ($\tau = 15$) and depth ($D = 100$m). The Miche [56] limit is calculated with half the steepness and the inner figure shows the bound of wave heights adjusted by $\mathfrak{S}_\beta = 0.6$.

## 7. Conclusions

For a long time, rogue waves have been described as waves with individual steepness much higher than ordinary ones, though it was not known how much steeper they could be. The data shows that in seas with very high significant steepness the steepest waves are not extreme. Rogue waves are found to be among the steepest waves, which include several hundreds of ordinary waves that are as steep or steeper, a feature described by Christou and Ewans [11]. Therefore, we conclude that waves with exceptionally high individual steepness will not necessarily become rogue waves. Surprisingly, however, waves with high effective steepness tend to have high normalized wave heights. Furthermore, the likelihood of appearance of rogue waves have not be found to depend individually on any of sea state parameters, also in agreement with the literature on the subject. Yet, according to Table 4, there seems to be a trend where rogue wave occurrence is related to a combination of the nonlinearity $\eta_{1/3}$, the height-to-depth ratio $\langle\epsilon\rangle$ and the significant steepness $\varepsilon$. The advantage of these variables is the possibility of extracting them from both hindcast and spectral shape, which is a desirable feature Gibson et al. [27].

The data has also pointed to the existence of several empirical rules

that may come to aid rogue wave forecasting, establishing bounds for the sea state parameters and providing a framework to assess the validity of exceeding probability distributions. One of the main concerns for these distributions is the generation of finite probabilities for physical scenarios where sea state parameters breached their upper bounds. In fact, the distributions discussed in Section 3 have no mention of any range of validity for their variables. Based on the observed empirical rules and without the support of any exceeding probability distribution, we were able to assess and obtain expressions for the expected highest normalized wave height and effective steepness. Combining the empirical rules with the models for expected maxima, an upper bound for the normalized wave height was obtained. Unfortunately, the data analysis did not provide a clear bound for the nonlinearity $\eta_{1/3}$, though it supplied a connection between the former and the storm averaged skewness, opening the possibility of obtaining $\eta_{1/3}$ from the spectral shape [1].

In this reassessment of Stansell's North Sea storm data we have further demonstrated the inability of Longuet-Higgins's model to explain observation. Longuet-Higgins [48] deviates from the upper part of the bulk waves ($1.5 < \alpha \leq 2$) significantly. This is of particular interest because Stansell's data demonstrates that sub-rogue waves of large absolute wave height appear as often as rogue waves of equivalent height. On the other hand, we confirmed that Longuet-Higgins [48] can provide good statistics for $\alpha < 1.5$, especially if contianing a large number of waves. However, when dealing with wave records that are equivalent to a couple of days, it performs poorly with only 40% of entries in Table 1 being predicted accurately. Using a weather analogy, Longuet-Higgins [48] can be a relatively good model for the climate while it is no longer useful for the daily forecast. On the other hand, some empirical distributions may perform well for the daily forecast, but not for a longer forecast, or a forecast with diametrical sea parameters. As another example, Christou and Ewans [11] concluded that Forristall [25] overpredicted rogue wave crests by 165%, whereas studies with smaller data sets have attested this model accuracy for wave crests [27,29,44]. This suggests that the manner one chooses the time series size or how one combines different wave records will affect the resulting return period. In this regard, the dimensionless sea state tracker $\aleph_1$ can be used as a grouping criteria, as it showed relatively strong correlation with the return period of rogue waves. Moreover, we have found that the storm sea state evolution (somewhat related to $\tau$) in time and its variability around its storm mean (or combination) could be related to robust overprediction or underprediction by most models such that the likely source of disagreement between several studies is that they analyzed the wave statistics of million of waves from several different locations, with diametrical sea state parameters and water depths, resulting in a "climate approach" statistics instead of one based on homogeneous storm records. When analyzed by homogeneity of sea states, we expect the return period to vary wildly in each group of wave records. This was confirmed by a new study based on an even larger data set analyzed accordingly [39], concluding that no unified probability model can explain the widest range of sea states properly.

Regarding the maximum height in a storm, Longuet-Higgins [48] predicts expected values for $\alpha$ that necessarily grows with the number of waves within the storm record. Such standard model of extreme wave statistics has proven to be a double-sword that leads to a twofold deficiency: First, it requires a very large number of waves for super-rogue wave formation, challenging recent observations [63]. Secondly, it can greatly underpredict or overpredict the maximum normalized wave height by assigning an extreme value to a unique number of waves. The combination of the two issues jeopardizes the exceeding probability by not allowing it to increase the assigned likelihood of a specific rogue wave without increasing the storm maximum expected normalized wave height. In this matter, the expected maximum for the normalized wave height obtained from empirical relations was found to be a function of several sea state variables. Our model shows very good variability for the maxima while the narrow-band formulation features incredibly small degree of variation. Therefore, our expectancy model of extreme dimensionless heights also explains the observed uneven distribution among storms that couldn't be understood in terms of Longuet-Higgins [48] or any second-order model. A possible setback, however, is the finite depth extension: the expressions that contain hyperbolic functions were obtained from deep and lower intermediate water regimes and they need to be validated for upper intermediate and shallow water wave regimes in a future work. Additionally, the spectral shape might change the ratio between mean wave period and the zero-crossing period, slightly changing the coefficients in equations (23), (43), (45–46).

Regarding crest heights, it is well-known that exceeding probability distributions for wave heights are typically not reliable for wave crests and vice-versa, hence, authors typically analyze them separately. In view of this problem, we have found empirically how wave crests heights measure up to wave heights with growing $\alpha$ and observed that this relation is remarkably accurate even for the maximum of both heights and crest heights, a feature not expected at the end of the tail. The finding that the average of this ratio, denoted by $\mathfrak{S}_\beta$, holds for both lower and maximum values of the normalized crests assures us of its own strong reliability in deep water, thus, explaining why Longuet-Higgins [48] prediction for wave crests are much worse than for wave heights. Accordingly, we applied such ratio to the most important results for wave heights and explained the origin of the ratio between rogue waves effective steepness and the remaining waves. It became clear that up to small deviation of a few percent, most distributions will be in agreement, however, the less successful the distribution in comparison to the empirical exceeding probability, the bigger the deviation from the observed ratio. This effect could be implemented as a third test for distributions, in addition to its comparison with the empirical distribution and the calculation of expected maxima. Moreover, we have obtained further justification for the sub-rogue wave classification, which seems to be the threshold for the "uneven distribution of rogue waves" among selected groups in Table 3. Therefore, a possible physical definition of a rogue wave can be related to the minimum normalized height (and crest height) where the stratification of empirical distributions such as in Fig. 32 is triggered, e.g. when the sea state parameters grow its influence on the exceeding probability. Not surprisingly, we confirmed that most distributions, whether for wave heights or crest heights, have very small variability of their exceeding distributions, such that one-parameter distributions are most likely to fail at least one of the crest or wave height distributions and at their expected maximum. In fact, Tayfun's predicted maxima were higher than Longuet-Higgins's and yet far below observation and with nearly indistinguishable values for $\mathbb{E}_\beta$. Remarkably, the ratio $\mathfrak{S}_\beta$ has shown to be very valuable in extending the bounds conjectured for wave heights to wave crests, maintaining a high correlation with observation, of the order of 0.78, whereas Longuet-Higgins [48] displayed a smaller correlation of 0.38 and Tayfun [77] of 0.40. Although the *nonlinearity*-adjusted RD shows much better agreement with observation, its variability is not sufficient to explain the "uneven distribution of rogue waves" described in Stansell [74] due to the failure of Longuet-Higgins [48] for wave heights, implying that the success or failure of explaining wave height statistics are reflected in the crest statistics through $\mathfrak{S}_\beta$.

A possible direction for future work is the unknown effect of the significant wave height time evolution on the rogue wave occurrence and expected maximum height, as well as reshaping all formulas, constraints and bounds obtained here in respect to directionality. As discussed by Gibson et al. [27] and Tayfun and Alkhalidi [81], is undesirable to rely too much on the spectral shape modelling, so that we intend to find a way to relate $\tau$ and $\eta_{1/3}$ to variables with higher accuracy when obtained by the spectral shape in a future work. Also, the conjectured model for heights starts to flatten out at the very end of the scatter plot, e.g. the error becomes of the order of the standard deviation for very high normalized crests or wave heights (with better accuracy for

the former), which is attributed to a lack of a second-order term $\langle\langle\delta_2\rangle\rangle^2$ in $\mathbb{E}_\alpha$. However, since the present data has very small variation in $k_2D$, or conversely $\lambda_2/D$, this is left for a future work with broader data.

## CRediT authorship contribution statement

**S. Mendes:** Conceptualization, Formal analysis, Data curation, Writing - original draft. **A. Scotti:** Formal analysis, Funding acquisition. **P. Stansell:** Data curation, Formal analysis, Methodology.

## 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.

## Acknowledgements

S.M. and A.S. were supported by the National Science Foundation under grant OCE-0729636.

## Appendix A. Crest Height Distributions

In order to convert crest height distributions to wave heights, we shall review the basics of statistics. Given the probability density $f_\alpha$ of water waves in narrow-banded seas [54], we can obtain the exceeding probability as follows:

$$\mathbb{P}\big(H>\alpha H_{1/3}\big)=\int_\alpha^{+\infty} 4\alpha\, e^{-2\alpha^2}\, d\alpha = e^{-2\alpha^2}. \tag{A.1}$$

The Tayfun [77] distribution is only one among several different variations of crest height distribution based on the Stokes second-order model [68]. For instance, Forristall [25] describes how Tung and Huang [89] differs from Tayfun [77] only on the definition of steepness and some coefficient, keeping the same structure of eq.(3). Following Tucker and Pitt [88], starting from the classical Stokes second-order expression for the water surface one can show that the probability density of the crest height is:

$$f(\xi)=\frac{\sqrt{1+2\xi\sigma k_1}-1}{\sigma k_1\sqrt{1+2\xi\sigma k_1}}\exp\left\{-\frac{\left[\sqrt{1+2\xi\sigma k_1}-1\right]^2}{2\sigma^2k_1^2}\right\}, \tag{A.2}$$

with $\xi=\mathscr{Z}_c/\sqrt{m_0}\equiv\mathscr{Z}_c/\sigma$. Then, by means of eq. (A.1) we arrive the equivalent exceeding probability:

$$\mathbb{P}(\mathscr{Z}_c>\xi\sigma)=\int_\xi^{+\infty} f(\xi)\,d\xi=\exp\left\{-\frac{\left[\sqrt{1+2\xi\sigma k_1}-1\right]^2}{2\sigma^2k_1^2}\right\}. \tag{A.3}$$

However, it is more convenient to write the exceeding probability in terms of significant wave height instead of the spectrum variance $\sigma$. Noting that the approximation $H_{1/3}=4\sqrt{m_0}$ holds [54], we rewrite the above expression as:

$$\mathbb{P}\big(\mathscr{Z}_c>\beta H_{1/3}\big)=\exp\left\{-\frac{\left[\sqrt{1+8\beta\mu}-1\right]^2}{2\mu^2}\right\}, \tag{A.4}$$

where $\mu=\sigma k_1=k_1H_{1/3}/4$ is the significant steepness, matching the original crest height distribution in Fedele et al. [22], Tayfun [77]. Rewriting the significant steepness as $k_1H_{1/3}=\breve{s}$, we rewrite Tayfun's distribution assuming a narrow-banded sea state ($H\approx 2\mathscr{Z}_c$):

$$\mathbb{P}\big(2\mathscr{Z}_c>2\beta H_{1/3}\big)=\exp\left\{-\frac{\left[\sqrt{1+4\cdot(2\beta)\cdot(\breve{s}/4)}-1\right]^2}{2(\breve{s}/4)^2}\right\}\Longrightarrow\mathbb{P}\big(H>\alpha H_{1/3}\big)\equiv\exp\left\{-\frac{8\left[\sqrt{1+\alpha\breve{s}}-1\right]^2}{\breve{s}^2}\right\}. \tag{A.5}$$

## Appendix B. Ocean Spectra and Wave Ratios

Since the available data does not contain information about the ocean spectrum [74,75], it is paramount to obtain theoretical and experimental relations among several different definitions of wave periods based on a given sea state. Given a spectrum $S(\omega)$, where $\omega$ is the wave frequency, the $k$-th moment is defined as [54,88]

$$m_k=\int_0^\infty \omega^k\, S(\omega)\,d\omega,\quad k\in\mathbb{Z}. \tag{B.1}$$

Following Longuet-Higgins [50], we define the spectral bandwidth as follows:

$$\nu=\sqrt{\frac{m_0m_2}{m_1^2}-1}. \tag{B.2}$$

The spectral significant wave height is obtained from the total variance of the spectrum $m_0$ [88]:

$$H_{m0}\approx 4\sqrt{m_0}, \tag{B.3}$$

begin approximately equal to the in situ measured significant wave height $H_{m0} \approx 1.04 H_{1/3}$ [54]. The major periods are defined as [88]:

$$T_E := \frac{2\pi m_{-1}}{m_0} \;\; ; \;\; T_1 := \frac{2\pi m_0}{m_1} \;\; ; \;\; T_2 := 2\pi\sqrt{\frac{m_0}{m_2}} \;, \tag{B.4}$$

being respectively known as the energy, mean and zero-crossing periods. As for the significant wave height, there is a distinction between their expected values based on the spectrum and the actual measurement. For the zero-crossing period, storm 29 showed that $T_2 \approx 0.9 T_z$ during the peak of the storm and almost perfect equivalency otherwise. Though the absence of any theoretical connection, the energy period shows remarkable correspondence with $T_{1/3}$ (the mean period of the tallest 1/3 waves), with an average $T_E \approx 1.01 T_{1/3}$ in storm 29. The analysis of the spectral moments of all storms leads to the average relation:

$$T_1^2 \approx \frac{22}{15} T_2^2 \quad \therefore \quad S_1 \approx \frac{15}{22} \langle s_s \rangle. \tag{B.5}$$

Lastly, is of great importance to know how the 1/3 tallest waves period compares to the zero-crossing one. From the spectral moments one reaches $T_E^2 \approx 2T_2^2$, whereas observation has $T_{1/3}^2 \approx 1.6 T_z^2$. Using the above relation, we can alternatively calculate the Ursell number from the significant steepness:

$$\mathrm{Ur} = \frac{\langle H_{1/3} \rangle \cdot \frac{15}{7} \langle \lambda_2 \rangle^2}{4\pi^2 D^3} \approx \frac{\langle H_{1/3} \rangle^3}{2\pi^2 D^3 \langle s_s \rangle^2}. \tag{B.6}$$

## Appendix C. Statistical Notation

Following the nomenclature where $p$ is the number of 20-min records within a storm containing $N$ waves and $n$ is the number of waves within a given 20-min record so that $\langle n \rangle p = N$, the averaging is described as:

| | |
|---|---|
| $X_i$ | $X$ measured at the $i$-th wave |
| $X_{ij}$ | $X_i$ measured at the $j$-th 20-min record |
| $X_j$ | $X$ measured at the $j$-th 20-min record |
| $\langle\langle X \rangle\rangle$ | 20-min record average: $\sum_{i=1}^{n} \frac{X_i}{n}$ |
| $\langle X \rangle$ | Average of 20-min record averages: $\sum_{j=1}^{p} \frac{\langle\langle X \rangle\rangle_j}{p}$ |
| $\overline{X}$ | Storm mean: $\sum_{i=1}^{N} \frac{X_i}{N}$ |
| $\| X \|$ | Storm maximum of variable $X$ |

Such distinction is necessary as some variables can only be defined for ensembles, such as $H_{1/3}$. As for the functions of averages, we have one example of distinction:

$$\langle \eta_{1/3} \rangle := \left\langle \frac{\langle\langle \mathscr{Z}_c \rangle\rangle}{\langle\langle \mathscr{Z}_t \rangle\rangle} \right\rangle_{\alpha > 1} \neq \eta_{1/3}. \tag{C.1}$$

Likewise, wavelengths, which are a function of every $T_{i,\ell}$ are calculated as ($\ell$ is the type of spectrum derived period):

$$\lambda_{i,\ell} := \frac{g T_{i,\ell}^2}{2\pi} \tanh\left(\frac{2\pi D}{\lambda_{i,\ell}}\right) \xrightarrow{\lambda_\ell / D \to 0} \langle \lambda_\ell \rangle = \frac{g}{2\pi} \langle T_\ell \rangle^2 \;, \tag{C.2}$$

with the associated wave number $k_{i,\ell} = 2\pi/\lambda_{i,\ell}$. Given the above definition and $\lambda_{1/3} \approx 8\lambda_2/5$, the conjectured models of Section 5 can be fully converted to a finite depth model:

$$\frac{\alpha_\infty}{\alpha_\star} = 1 + \left(\frac{2\pi}{g}\right)^{1/8} \left(\frac{24\pi\tau}{4375}\right)^{1/2} \left(\frac{\langle \lambda_2 \rangle}{D}\right)^{5/8} \frac{D^{1/8}}{\eta_{1/3}^4 \tanh^{1/8}\left(\frac{2\pi D}{\langle \lambda_2 \rangle}\right)} \approx 1 + \frac{19}{153} \left(\frac{\langle \lambda_2 \rangle}{D}\right)^{5/8} \frac{D^{1/8}\sqrt{\tau}}{\eta_{1/3}^4 \tanh^{1/8}\left(\frac{2\pi D}{\langle \lambda_2 \rangle}\right)}. \tag{C.3}$$

Moreover, the *unbiased* skewness [38] in eq. (14) of the surface elevation $H_i$ is measured as:

$$\langle\langle \mu_3 \rangle\rangle = \frac{\sqrt{n(n-1)}}{(n-2)} \frac{\left[\frac{1}{n}\sum_{i=1}^{n} (H_i - \langle\langle H \rangle\rangle)^3\right]}{\left[\frac{1}{n}\sum_{i=1}^{n} (H_i - \langle\langle H \rangle\rangle)^2\right]^{3/2}}. \tag{C.4}$$

Since the fourteen storm average of $\langle n \rangle$ is of the order of 120 waves per 20-min record, the *unbiased* skewness is 1.3% higher than the *biased* skewness (e.g. the central moment approach).

## Appendix D. Gumbel's Extreme Value Theory

In order to compute the maximum height of a narrow-banded sea we need to find the mode of the distribution and the intensity function. According to Gumbel et al. [31], one can obtain the latter in following manner:

$$\widetilde{\mu}_N = \frac{f_\alpha}{1-F_\alpha} \equiv -\frac{1}{\mathbb{P}_\alpha}\frac{d\mathbb{P}_\alpha}{d\alpha} = \left(\frac{4\alpha e^{-2\alpha^2}}{e^{-2\alpha^2}}\right)_{\alpha=\widetilde{\alpha}_N} = 4\widetilde{\alpha}_N, \tag{D.1}$$

With $\widetilde{\alpha}_N$ being the mode and $N$ the total number of waves anf $F_\alpha = 1 - \mathbb{P}_\alpha$ is the cumulative probability distribution. The mode is found by solving the equation below [31]:

$$F\left(\widetilde{\alpha}_N\right) = 1 - \frac{1}{N} \quad \therefore \quad N\cdot\mathbb{P}_{\widetilde{\alpha}_N} = 1. \tag{D.2}$$

Then, Gumbel et al.'s generalized formula for the maximum is:

$$\mathbb{E}_\alpha = \widetilde{\alpha}_N + \frac{\gamma_E}{\widetilde{\mu}_N} \approx \widetilde{\alpha}_N + \frac{\sqrt{3}}{3\widetilde{\mu}_N}, \tag{D.3}$$

where $\gamma_E$ is the Euler-Mascheroni constant. Applied to Longuet-Higgins's narrow-banded sea distribution, Gumbel et al.'s expected maximum dimensionless height reads:

$$\mathbb{E}_\alpha = \sqrt{\frac{\ln N}{2}} + \frac{\sqrt{3}}{12\sqrt{\frac{\ln N}{2}}} = \sqrt{\frac{\ln N}{2}}\left[1 + \frac{\sqrt{3}}{6\ln N}\right]. \tag{D.4}$$

Following the same rationale for crest heights, one can obtain:

$$\widetilde{\mu}_N = -\frac{1}{\mathbb{P}_\beta}\frac{d\mathbb{P}_\beta}{d\beta} = \left(\frac{16\beta e^{-8\beta^2}}{e^{-8\beta^2}}\right)_{\beta=\widetilde{\beta}_N} = 16\widetilde{\beta}_N, \tag{D.5}$$

With $\widetilde{\beta}_N$ being the mode and $N$ the total number of waves. The mode is found by solving the equation below [31]:

$$N\cdot\mathbb{P}_{\widetilde{\beta}_N} = 1, \tag{D.6}$$

such that applied to Longuet-Higgins's narrow-banded sea distribution, Gumbel et al.'s expected maximum dimensionless height reads:

$$\mathbb{E}_\beta = \widetilde{\beta}_N\left[1 + \frac{\sqrt{3}}{48\widetilde{\beta}_N^2}\right] = \sqrt{\frac{\ln N}{8}}\left[1 + \frac{\sqrt{3}}{6\ln N}\right]. \tag{D.7}$$

For the Forristall [24] model, however, we apply the same procedure and find its intensity:

$$\widetilde{\mu}_N = -\frac{1}{\mathbb{P}_\alpha}\frac{d\mathbb{P}_\alpha}{d\alpha} = 4.811\widetilde{\alpha}_N^{1.126}, \tag{D.8}$$

while the mode reads:

$$N = e^{2.263\widetilde{\alpha}_N^{2.126}} \quad \therefore \quad \widetilde{\alpha}_N \approx (0.44\ln N)^{0.47}. \tag{D.9}$$

Hence, combining the two last equations, we calculate Forristall's expected maximum normalized wave height:

$$\mathbb{E}_{\alpha,\mathscr{F}} = (0.44\ln N)^{0.47} + \frac{\sqrt{3}}{3\cdot 4.811\cdot(0.44\ln N)^{0.53}} \approx (0.44\ln N)^{0.47}\left[1 + \frac{3}{11\ln N}\right]. \tag{D.10}$$

### D1. Tayfun's expected maximum

We start with the full expression for the Tayfun [77] distribution:

$$\mathbb{P}\left(\mathscr{Z}_c > \beta H_{1/3}\right) \equiv \exp\left\{-\frac{8}{\breve{s}^2}\left[\sqrt{1+2\beta\breve{s}} - 1\right]^2\right\}, \tag{11}$$

whose mode can be found through,

$$N = \exp\left\{\frac{8}{\breve{s}^2}\left[\sqrt{1+2\widetilde{\beta}_N\breve{s}}-1\right]^2\right\}, \tag{12}$$

leading to the relation:

$$\sqrt{1+2\widetilde{\beta}_N\breve{s}} = 1+\breve{s}\,\widetilde{\beta}_0, \quad \widetilde{\beta}_0 = \sqrt{\frac{\ln N}{8}}, \tag{13}$$

and solution,

$$\widetilde{\beta}_N = \widetilde{\beta}_0\left(1+\frac{\breve{s}\,\widetilde{\beta}_0}{2}\right). \tag{14}$$

On the other hand, the intensity reads:

$$\widetilde{\mu}_N = \frac{16\left(\sqrt{1+2\widetilde{\beta}_N\breve{s}}-1\right)}{\breve{s}\sqrt{1+2\widetilde{\beta}_N\breve{s}}} \equiv \frac{16\,\widetilde{\beta}_0}{\left(1+\widetilde{\beta}_0\breve{s}\right)}. \tag{15}$$

In the same manner that Tayfun [77] recovers Longuet-Higgins [48] in the limit $\breve{s}\to 0$, likewise, the mode and intensity of Tayfun's distribution shall recover the mode and intensity of the previous section, satisfied by the above expressions. Therefore, it is straightforward to arrive at:

$$\mathbb{E}_\beta = \sqrt{\frac{\ln N}{8}}\left[1+\frac{\sqrt{3}}{6\ln N}+\breve{s}\sqrt{\frac{\ln N}{8}}\left(\frac{1}{2}+\frac{\sqrt{3}}{6\ln N}\right)\right], \tag{16}$$

which can be approximated by up to 1% error as ($\ln N \sim 10$):

$$\mathbb{E}_{\beta,\mathscr{T}} \approx \sqrt{\frac{\ln N}{8}}\left[1+\frac{\sqrt{3}}{6\ln N}\right]\left(1+\frac{4\breve{s}}{7}\right) \approx \sqrt{\frac{\ln N}{8}}\left[1+\frac{\sqrt{3}}{6\ln N}\right]\left(1+\frac{22}{9}\langle s_s\rangle\right). \tag{17}$$

### 4.2. normalization through $H_{m0}$

In order to prove the assertion that the normalization whether through $H_{m0}$ or $H_{1/3}$ leaves the accuracy of eq. (43) relative to the observed $\|\alpha\|$ invariant, let us define $\alpha^* = H/H_{m0}$, so that according to Fig. 1 we have on average $\alpha = 1.04\alpha^*$. The Rayleigh distribution (RD) would read instead:

$$\mathscr{R}_{\alpha^*} \equiv \mathscr{R}\left(H > \alpha^* H_{1/3}\right) = e^{-2.1632\alpha^{*\,2}}. \tag{18}$$

Without loss of generality, the intensity now looks instead:

$$\mu_N^* = \left(\frac{4.32\alpha^* e^{-2.1632\alpha^{*\,2}}}{e^{-2.1632\alpha^{*\,2}}}\right)_{\alpha=\widetilde{\alpha}_N} = 4.3264\alpha_N^*, \tag{19}$$

whereas the mode has:

$$\alpha_N^* = \sqrt{\frac{\ln N}{2.1632}} = \frac{1}{1.04}\sqrt{\frac{\ln N}{2}} = \frac{\widetilde{\alpha}_N}{1.04}, \tag{20}$$

such that we find:

$$\mu_N^* = 4.3264\times\frac{\widetilde{\alpha}_N}{1.04} = 4.16\widetilde{\alpha}_N = 1.04\widetilde{\mu}_N. \tag{21}$$

Accordingly, the maximum expected normalized height through $H_{m0}$ instead of $H_{1/3}$ is:

$$\mathbb{E}_\alpha^* = \alpha_N^* + \frac{\sqrt{3}}{3\mu_N^*} = \frac{\widetilde{\alpha}_N}{1.04} + \frac{\sqrt{3}}{3\times 1.04\widetilde{\mu}_N} = \frac{\mathbb{E}_\alpha}{1.04}, \tag{22}$$

ultimately leading to the conclusion of the proof:

$$\frac{\mathbb{E}(\|\alpha\|)}{\|\alpha\|} = \frac{1.04\times\mathbb{E}(\|\alpha^*\|)}{1.04\times\|\alpha^*\|} = \frac{\mathbb{E}(\|\alpha^*\|)}{\|\alpha^*\|}. \tag{23}$$

Likewise, the discrepancy between observed statistics and $\mathscr{R}_{\alpha}$ will not change when we use $\mathscr{R}_{\alpha^*}$, as the observed statistics will also change due to the redefinition in Table 1.

### 4.3. Nonlinearity-adjusted Higgins (1952)

Following Eqs. (–D.3), the adjusted Longuet-Higgins's distribution in eq. (56) will lead to the following intensity (using the approximation of eq. (70) for $\mathscr{I}_{\beta}$):

$$\widetilde{\mu}_N = 4\widetilde{\beta}_N \left(\frac{\eta_{1/3}}{1+\eta_{1/3}}\right)^{-2} \left(1+\frac{\eta_{1/3}}{6}\right)^{-2}, \tag{24}$$

and respective mode:

$$\widetilde{\beta}_N = \left(\frac{\eta_{1/3}}{1+\eta_{1/3}}\right)\left(1+\frac{\eta_{1/3}}{6}\right)\sqrt{\frac{\ln N}{2}}. \tag{25}$$

Ultimately, the expected maximum becomes:

$$\mathbb{E}_{\beta,\star} \approx \sqrt{\frac{\ln N}{8}}\left[1+\frac{\sqrt{3}}{6\ln N}\right]\left(\frac{2\eta_{1/3}}{1+\eta_{1/3}}\right)\left(1+\frac{\eta_{1/3}}{6}\right). \tag{26}$$

## 5. Error Function

While evaluating average heights and crest heights for the RD, one integral is of particular interest:

$$\widetilde{\mathscr{G}}_0^+(x) = \frac{2}{\sqrt{\pi}}\int_x^{+\infty} e^{-t^2}\,dt := erfc(x) \equiv 1 - erf(x)\ . \tag{E.1}$$

Obviously, we have $\widetilde{\mathscr{G}}_0^+(0) = erfc(0) = 1$. Integration by parts leads to [28]:

$$\widetilde{\mathscr{G}}_0^+(x) = \frac{e^{-x^2}}{x\sqrt{\pi}}\sum_{k=0}^{+\infty}(-1)^k\frac{(2k-1)!!}{(2x^2)^k} \sim \frac{e^{-x^2}}{x\sqrt{\pi}}\ ,\ x>1\ . \tag{E.2}$$

In fact, up to a 5% deviation, we have that $erfc(1) \approx 1/6$. In the vicinity of $x = 1$ the leading order term is, then:

$$erfc(x) \sim \frac{5e^{-x^2}}{6x\sqrt{\pi}}. \tag{E.3}$$

## 6. Bandwidth adjusted Rayleigh Distribution in (0+1) dimensions

[48] described the following Rayleigh distribution generalization:

$$\mathbb{P}(\mathscr{Z}_c > u) = e^{-u^2/\overline{a}^2}, \tag{F.1}$$

where $\overline{a}$ is the rms amplitude of the sea surface [54]. For narrow-banded seas the rms amplite reads $\overline{a} = \sqrt{2m_{000}}$, such that the ratio becomes $(u/\overline{a})^2 = \beta^2 H_{1/3}/2m_{000}$. Since $H_{1/3} \approx 4\sqrt{m_{000}}$ [54], we find $(u/\overline{a})^2 = 8\beta^2$. Longuet-Higgins [51] showed that given a finite spectral bandwidth $\nu$, the rms amplitude can be rewritten as:

$$\overline{a}^2 \approx 2m_{000}\left(1-\frac{11\nu^2}{15}\right) \equiv \frac{2m_{000}}{\nu_{\star}} = \frac{H_{1/3}^2}{8\nu_{\star}}. \tag{F.2}$$

Therefore, one finds the corrected crest height probability:

$$\mathscr{R}\left(\mathscr{Z}_c > \beta H_{1/3}\right) = e^{-\beta^2 H_{1/3}^2/\overline{a}^2} = e^{-8\nu_{\star}\beta^2} \equiv \mathscr{R}_{\beta,\nu}. \tag{F.3}$$

## Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.apor.2020.102402

## References


[1] Annenkov, S.Y., Shrira, V.I., 2014. Evaluation of skewness and kurtosis of wind waves parameterized by JONSWAP spectra. J. Phys. Oceanogr. 44 (6), 1582–1594.
[2] Barbariol, F., Benetazzo, A., Carniel, S., Sclavo, M., 2015. Space-time wave extremes: the role of metocean forcings. J. Phys. Oceanogr. 45 (7), 1897–1916.
[3] Battjes, J.A., Groenendijk, H.W., 2000. Wave height distributions on shallow offshores. Coastal Eng. 40, 61–82.
[4] Bitner-Gregersen, E.M., Fernandez, L., Lefévre, J.M., Monbaliu, J., Toffoli, A., 2014. The north sea andrea storm and numerical simulations. Nat. Hazards Earth Syst. Sci. 14 (6), 1407–1415.
[5] Boccotti, P., 1989. On mechanics of irregular gravity waves. In: Atti della AScad. naz. dei Lincei A386 19.
[6] Boccotti, P., 2000. Wave mechanics for ocean engineering. Elsevier Oceanography Series.
[7] Borgman, L.E., 1973. Probabilities for highest wave in hurricane. ASCE J. Waterw Harbors Coastal Eng Div 99 (WW2), 185–207.
[8] Cartwright, D.E., Longuet-Higgins, M.S., 1956. The statistical distribution of the maxima of a random function. Proc. R. Soc. A 237 (1209), 212–232.
[9] Cherneva, Z., Petrova, P., Andreeva, N., Guedes Soares, C., 2005. Probability distributions of peaks, troughs and heights of wind waves measured in the black sea coastal zone. Coastal Eng. 52 (7), 599–615.
[10] Chien, H., Kao, C.C., Chuang, L.Z., 2002. On the characteristics of observed coastal freak waves. Coastal Eng. 44, 301–319.
[11] Christou, M., Ewans, K., 2014. Field measurements of rogue water waves. J. Phys. Oceanogr. 9, 2317–2335.
[12] Dean, R.G., 1990. Freak waves: a possible explanation. In: Tørum A. and Gudmestad D.T.(eds) Water wave kinematics. NATO ASI Series (E:Applied Sciences) 178.
[13] Didenkulova, I., Anderson, C., 2010. Freak waves of different types in the coastal zone of the baltic sea. Nat. Hazards Earth Syst. Sci. 10, 2021–2029.
[14] Didenkulova, I., Rodin, A., 2012. Statistics of shallow water rogue waves in baltic sea conditions: the case of tallinn bay. in: Proc. of the 2012 IEEE/OES Baltic Int. Symposium 1–6.
[15] Draper, L., 1964. Freak ocean waves. Oceanus 10, 13–15.
[16] Draper, L., 1971. Severe wave conditions at sea. J. Inst. Navig. 24 (3), 273–277.
[17] Dysthe, K., Krogstad, H.E., Muller, P., 2008. Oceanic rogue waves. Annu. Rev. Fluid. Mech. 40, 287–310.
[18] Earle, M.D., 1975. Extreme wave conditions during hurricane camille. J. Geophys. Res. 80 (3), 377–379.
[19] Faukner, D., 2002. Shipping safety: a matter of concern. Ingenia, The Royal Academy of Engineering, Marine Matters 13–20.
[20] Fedele, F., 2015. On the kurtosis of deep-water gravity waves. J. Fluid Mech. 782, 25–36.
[21] Fedele, F., Arena, F., 2010. Long-term statistics and extreme waves of sea storms. J. Phys. Oceanogr. 40 (5), 1106–1117.
[22] Fedele, F., Brennan, J., De Leon, S.P., Dudley, J., Dias, F., 2016. Real world ocean rogue waves explained without the modulational instability. Sci. Rep. 6, 27715.
[23] Fedele, F., Lugni, C., Chawla, A., 2017. The sinking of the el faro: predicting real world rogue waves during hurricane joaquin. Sci. Rep. 7 (1).
[24] Forristall, G.Z., 1978. On the distributions of wave heights in a storm. J. Geophys. Res. 83, 2353–2358.
[25] Forristall, G.Z., 2000. Wave crest distributions: observations and second order theory. J. Phys. Ocean. 30, 1931–1943.
[26] Forristall, G.Z., 2005. Understanding rogue waves - are new physics necessary? Proc. 14th Aha Huliko'a Winter Workshop.
[27] Gibson, R., Christou, M., Feld, G., 2014. The statistics of wave height and crest elevation during the december 2012 storm in the north sea. Ocean Dyn. 64, 1305–1317.
[28] Gradshteyn, I.S., Ryzhik, I.M., 2007. Table of integrals, series and products. Academic Press.
[29] Gramstad, O., Bitner-Gregersen, E., Breivik, O., Magnusson, A.K., Reistad, M., Aarnes, O.J., 2018. Analysis of rogue waves in north-sea in-situ surface wave data. Proc. Int. Conf. on Offshore Mechanics and Arctic Eng. 3.
[30] Green, M.O., 1994. Wave-height distribution in storm sea: effect of wave breaking. J. Waterway, Port, Coastal Ocean Eng. 120, 283–301.
[31] Gumbel, E., J., 1958. Statistics of extremes.New York
[32] Hallowell, S., Myers, A.T., Arwade, S.R., 2015. Variability of breaking wave characteristics and impact loads on offshore wind turbines supported by monopiles. Wind Energy 19, 301–312.
[33] Haring, R.E., Osborne, A.R., Spencer, L.P., 1976. Extreme wave parameters based on continental shelf storm wave records. Proc. 15th Int. Conf. on Coastal Engineering, Honolulu, HI 151–170.
[34] Haver, S., 2004. A possible freak wave event measured at the draupner jacket january 1, 1995. Proc. Rogue Waves 20–22 October IFREMER.
[35] Haver, S., Andersen, O.J., 2000. Freak waves: rare realizations of a typical population of typical realizations of a rare population? Proc. Int. Offshore Polar Eng. Conf. 3, 123–130.
[36] Jahns, H.O., Wheeler, J.D., 1973. Long-term wave probabilities based on hindcasting of severe storms. J. Petrol. Technol. 25, 473–486.
[37] Janssen, P.A.E.M., 2003. Nonlinear four-wave interactions and freak waves. J. Phys. Oceanogr. 33 (4), 863–884.
[38] Joanes, D.N., Gill, C.A., 1998. Comparing measures of sample skewness and kurtosis. J. Royal Stat. Soc. D 47 (1), 183–189.
[39] Karmpadakis, I., Swan, C., Christou, M., 2020. Assessment of wave height distributions using an extensive field database. Coastal Eng. 157.
[40] Katsardi, V., de Lutio, L., Swan, C., 2013. An experimental study of large waves in intermediate and shallow water depths. part i: wave height and crest height statistics. Coastal Eng. 73, 43–57.
[41] Kharif, C., Pelinovsky, E., 2003. Physical mechanisms of the rogue wave formation. Eur. J. Mech. B Fluids 22, 603–634.
[42] Kjeldsen, S.P., 1984. Dangerous wave groups. Nor. Marit. Res. 12, 4–16.
[43] Krogstad, H.E., 1985. Height and period distributions of extreme waves. Appl. Ocean Res. 7 (3), 158–165.
[44] Kvingedal, B., Bruserud, K., Nygaard, E., 2018. Individual wave height and wave crest distributions based on field measurements from the northern north sea. Ocean Dyn. 68 (12), 1727–1738.
[45] Laface, V., Malara, G., Kougioumtzoglou, I.A., Romolo, A., Arena, F., 2018. Nonlinear wave crest distribution on a vertical breakwater. Coastal Eng. 138, 227–234.
[46] Liu, P.C., MacHutchon, K.R., 2006. Are there different kinds of rogue waves? Proc. Conf. on Offshore Mech. and Arctic Eng. 865–870.
[47] Liu, P.C., Pinho, U.F., 2004. Freak waves - more frequent than rare! Ann. Geophys. 5, 1839–1842.
[48] Longuet-Higgins, M.S., 1952. On the statistical distribution of the heights of sea waves. J. Mar. Res. 11, 245–265.
[49] Longuet-Higgins, M.S., 1963. The effect of non-linearities on statistical distributions in the theory of sea waves. J. Fluid Mech. 17, 459–480.
[50] Longuet-Higgins, M.S., 1975. On the joint distribution of the periods and amplitudes of sea waves. J. Geophys. Res. 80 (18), 2688–2694.
[51] Longuet-Higgins, M.S., 1980. On the distribution of the heights of sea waves: some effects of nonlinearity and finite band width. J. Geophys. Res. 85 (C3), 1519–1523.
[52] Lu, W., Li, J., Li, X., Tian, X., Wu, X., Zhang, X., 2019. Experimental investigation on the statistics of rogue waves under a random wave background. Ocean Eng. 186.
[53] Mallory, J.K., 1974. Abnormal waves in the south-east coast of south africa. Int. Hydrog. Rev. 51, 89–129.
[54] Massel, S.R., 2017. Ocean surface waves: Their physics and prediction, 3rd. World Scientific, Singapore.
[55] Mendes, S., Scotti, A., 2020. Rogue wave statistics in (2+1) gaussian seas i: narrow-banded distribution. Appl. Ocean Res. 99, 102043.
[56] Miche, R., 1944. Mouvements ondulatoires de la mer en profendeur constante ou decroissante forme limite de la houle lors de son deferlement. Ann. Ponts Chaussees 121, 285–319.
[57] Mori, N., 2003. Occurrence probability of freak waves in nonlinear wave field. Ocean Eng. 31, 165–175.
[58] Mori, N., Janssen, P.A.E.M., 2005. On kurtosis and occurrence probability of freak waves. J. Phys. Oceanogr. 36 (7), 1471–1483.
[59] Mori, N., Liu, P.C., Yasuda, T., 2002. Analysis of freak waves measurements in the sea of japan. Ocean Eng. 29 (11), 1399–1414.
[60] Mori, N., Onorato, M., Janssen, P.A.E.M., 2011. On the estimation of the kurtosis in directional sea states for freak wave forecasting. J. Phys. Oceanogr. 41 (8), 1484–1497.
[61] Mori, N., Yasuda, T., 2002. A weakly non-gaussian model of wave height distribution random wave train. Ocean Eng. 29 (10), 1219–1231.
[62] Naess, A., 1985. On the distribution of crest to trough wave heights. Ocean Eng. 12 (3), 221–234.
[63] Nikolina, I., Didenkulova, I., 2011. Rogue waves in 2006–2010. Nat. Hazards Earth Syst. Sci. 11, 2913–2924.
[64] Onorato, M., Cavaleri, L., Fouques, S., Gramstad, O., Janssen, P.A.E.M., Monbaliu, J., Osborne, A.R., Pakozdi, C., Serio, M., Stansberg, C.T., Toffoli, A., Trulsen, K., 2009. Statistical properties of mechanically generated surface gravity waves: a laboratory experiment in a three-dimensional wave basin. J. Fluid Mech. 627, 235–257.
[65] Pelinovsky, E., Kharif, C., 2008. Extreme ocean waves. World Scientific, New York.
[66] Petrova, P., Cherneva, Z., Guedes Soares, C., 2007. On the adequacy of second-order models to predict abnormal waves. Ocean Eng. 22, 603–634.
[67] de Pinho, U.F., Liu, P.C., Parente Ribeiro, C.E., 2004. Freak waves at campos basin, brazil. Geofizika 21, 53–67.
[68] Prevosto, M., Bouffandeau, B., 2002. Probability of occurrence of a "giant" wave crest. Proc. Int. Conf. on Offshore Mechanics and Arctic Eng. 2, 483–490.
[69] Romolo, A., Arena, F., 2015. On adler space-time extremes during ocean storms. Journal of Geophysical Research: Oceans 120 (4), 3022–3042.
[70] Romolo, A., Arena, F., Laface, V., 2014. A generalized approach to the mechanics of three-dimensional nonlinear ocean waves. Probab. Eng. Mech. 35, 96–107.
[71] Romolo, A., Malara, G., Laface, V., Arena, F., 2016. Space-time long-term statistics of ocean storms. Probab. Eng. Mech. 44, 150–162.
[72] Rosenthal, W., Lehner, S., 2007. Individual wave height from SAR. ESA Special Publication: Envisat Symposium SP–636.
[73] Rosenthal, W., Lehner, S., 2008. Rogue waves: results of the maxwave project. J. Offshore Mech. Arctic Eng. 130, 021006.
[74] Stansell, P., 2004. Distribution of freak wave heights measured in the north sea. Appl. Ocean Res. 26, 35–48.
[75] Stansell, P., 2005. Distributions of extreme wave, crest and trough heights measured in the north sea. Ocean Eng. 32, 1015–1036.
[76] Stokes, G.G., 1847. On the theory of oscillatory waves. Trans. Camb. Phil. Soc. 8, 441–455.
[77] Tayfun, M.A., 1980. Narrow-band nonlinear sea waves. J. Geophys. Res. 85, 1548–1552.
[78] Tayfun, M.A., 1990. Distribution of large wave heights. J. Waterway, Port, Coastal Ocean Eng. 116 (6), 686–707.
[79] Tayfun, M.A., 2006. Distributions of wave steepness and surf parameter. J. Waterway, Port, Coastal Ocean Eng. 132 (1), 1–9.

[80] Tayfun, M.A., 2006. Statistics of nonlinear wave crests and groups. Ocean Eng. 33 (11), 1589–1622.
[81] Tayfun, M.A., Alkhalidi, M.A., 2020. Distribution of sea-surface elevations in intermediate and shallow water depths. Coastal Eng. 157.
[82] Tayfun, M.A., Fedele, F., 2007. Wave-height distributions and nonlinear effects. Ocean Eng. 34 (11), 1631–1649.
[83] Toffoli, A., Benoit, M., Onorato, M., Bitner-Gregersen, E.M., 2009. The effect of third-order nonlinearity on statistical properties of random directional waves in finite depth. Nonlinear Process. Geophys. 16 (1), 131–139.
[84] Toffoli, A., Bitner-Gregersen, E., Onorato, M., Babanin, A.V., 2008. Wave crest and trough distributions in a broad-banded directional wave field. Ocean Eng. 35 (17–18), 1784–1792.
[85] Toffoli, A., Lefevre, J.M., Bitner-Gregersen, E., Monbaliu, J., 2005. Towards the identification of warning criteria: analysis of a ship accident database. Appl. Ocean Res. 27, 281–291.
[86] Toffoli, A., Proment, D., Salman, H., Monbaliu, J., Frascoli, F., Dafilis, M., Stramignoni, E., Forza, R., Manfrin, M., Onorato, M., 2017. Wind generated rogue waves in an annular wave flume. Phys. Rev. Lett. 118 (14).
[87] Trulsen, K., Dysthe, K.B., 1997. Freak waves - a three-dimensional wave simulation. in: Proc. 21st Symposium on Naval Hydrodynamics.
[88] Tucker, M.J., Pitt, E.G., 2001. Waves in ocean engineering. Elsevier, Amsterdam.
[89] Tung, C.C., Huang, N.E., 1985. Peak and trough distributions of nonlinear waves. Ocean Eng. 12, 201–209.
[90] Voermans, J.J., Laface, V., Babanin, A.V., Romolo, A., Arena, F., 2020. Standing wave field observations at a vertical wall. Coastal Eng. 160.
[91] Waseda, T., Kinoshita, T., Tamura, H., 2009. Evolution of a random directional wave and freak wave occurrence. J Phys Oceanogr 39 (3), 621–639.
[92] Weggel, J.R., 1972. Maximum breaker height for design. In: Proceedings 13th International Conference Coastal Engineering.
[93] Wolfram, J., Linfoot, B., Stansell, P., 2000. Long- and short-term extreme wave statistics in the north sea: 1994–1998. In: M. Olagnon and G.A. Athanassoulis (Eds.), Rogue Waves 2000 32, 341–347.
[94] Wu, H.J., 1973. A discussion of the wave records obtained at weather station papa, pacific ocean. Proc. of the Second Int. Conf. on Port and Ocean Eng. under Arctic Conditions 49.
[95] Wu, Y., Randell, D., Christou, M., Ewans, K., Jonathan, P., 2016. On the distribution of wave height in shallow water. Coastal Eng. 111, 39–49.