# Statistical distributions of free surface elevation and wave height for out-of-equilibrium sea-states provoked by strong depth variations

Jie Zhang, Yuxiang Ma, Michel Benoit

# Statistical distributions of free surface elevation and wave height for out-of-equilibrium sea-states provoked by strong depth variations

Jie Zhang[a], Yuxiang Ma[a,*] and Michel Benoit[b,c]

[a]State Key Laboratory of Coastal and Offshore Engineering, Dalian University of Technology, Dalian 116023, PR China
[b]EDF R&D, Laboratoire National d'Hydraulique et Environnement (LNHE), Chatou 78400, France
[c]LHSV, Ecole des Ponts, EDF R&D, Chatou 78400, France

ARTICLE INFO

*Keywords*:
Coastal extreme waves
Wave-bottom interaction
Out-of-equilibrium dynamics
Statistical analysis
Wave statistical distributions

ABSTRACT

In the recent work of Zhang *et al.* [Coastal Eng. **174**, 104099 (2022)], it is shown that, as unidirectional irregular waves propagate over an uneven bottom with a rapid reduction of water depth, the sea-state becomes out-of-equilibrium and non-equilibrium dynamics (NED) manifests. As the sea-state gradually adapts to the new (reduced) water depth, the NED results in significant modulation of wave spectrum over a relatively long distance. The present work further analyzes the numerical results of Zhang *et al.* (2022) and investigates the NED effects on wave statistics in both near-field and far-field regions after the water depth transition, including the probability distributions of free surface elevation and wave height, the statistical moments and maximum wave statistics. The primary contribution of the present work is to assess the applicability and limitations of several popular statistical distribution models that are either frequently studied in the scientific literature or used in engineering practices. Furthermore, a new distribution of the free surface elevation in a lognormal shape is proposed. It predicts the non-equilibrium free surface statistics with satisfactory performance and characterizes well the relationship between skewness and kurtosis. This work also shows that the statistics in the far-field region are significantly influenced by the near-field wave-wave interaction, and beyond the capability of all statistical models considered here. Despite this complexity, the sea-state in the far-field region exhibits lower freak wave probability than a Gaussian sea-state. Implications of these findings for engineering practices are discussed.

## Contents



*Corresponding author
shuidong0618@126.com (J. Zhang); yuxma@dlut.edu.cn (Y. Ma); michel.benoit@edf.fr (M. Benoit)
ORCID(s): 0000-0003-0794-2335 (J. Zhang); 0000-0003-4314-0428 (Y. Ma); 0000-0003-4195-2983 (M. Benoit)

## 1. Introduction

Wave statistics represents a key input for the design of maritime structures (see e.g. Chakrabarti, 2005; Goda, 2010). Typically, a design sea-state that corresponds to a 50∼100 years return period is obtained with extreme value analysis

of the long-term sea-state data from either in-situ measurements or hindcast simulations. In a prescribed design sea-state, the maximum wave load, run-up, and overtopping on fixed structures are determined based on the tail parts of the short-term statistical distributions. The design wave is obtained from the wave height (or crest height) statistical distribution with a prescribed occurrence probability. From the engineering perspective, improving the accuracy and reducing the uncertainty in the tail part of the statistical distributions is of great interest, as this tail part determines the extreme wave occurrence probability and plays a vital role in the balance between safety and economy in the structure design. However, from a theoretical point of view, modelling the tail of the distribution remains a challenging task, as it is sensitive to various factors like wave nonlinearity, water depth, wave breaking intensity, etc.

In recent decades, freak (or rogue) waves, defined as waves with crest-to-trough heights higher than twice the significant wave height, are recognized as hazardous events that occur all over the world with a startling frequency (see e.g. Godoi et al., 2017; O'Brien et al., 2017; Didenkulova and Pelinovsky, 2020; Ma et al., 2022; Didenkulova et al., 2023). They appear as abnormally heavy tails in probability distributions, with occurrence probability considerably higher than the prediction of a linear model (Longuet-Higgins, 1952; Dysthe et al., 2008). Characterizing freak wave statistics in coastal areas is of great importance since human activity is more involved and still rapidly growing in these areas. However, the wave evolution in coastal areas is complicated due to the complexity of the surrounding conditions. Although many hypotheses have been put forward to explain the formation of freak waves (see e.g. Kharif et al., 2009; Onorato et al., 2013; Dematteis et al., 2019) from either a linear or a nonlinear perspective, yet no consensus is achieved so far on the dominant mechanism of freak waves in the real sea. The modulation (or Benjamin-Feir) instability (MI) is probably the most well-known nonlinear mechanism for freak waves (Benjamin and Feir, 1967; Janssen, 2003). However, it assumes deep-water narrow-banded waves. The role played by MI in the formation of freak waves in real sea is still under debate (e.g. Chabchoub, 2016; Fedele et al., 2016), with the main argument being that MI is ineffective in coastal areas with sufficiently shallow water depth.

One decade ago, it was shown that the non-equilibrium dynamics (NED) induced by rapid depth variations could enhance the occurrence probability of coastal freak waves (Trulsen et al., 2012; Zeng and Trulsen, 2012; Viotti and Dias, 2014). Since then, the NED, as a new perspective of nonlinear focusing, has attracted considerable attention. It renders some generality in explaining freak waves since the NED could be provoked not only by rapid depth change but also by changes of other environmental conditions (Trulsen, 2018). For example, an evident non-equilibrium phenomenon could be observed for the sea-states that encounter a significant change of ambient flow velocity (Zhang et al., 2023) or wind speed (Annenkov and Shrira, 2015). As the environmental condition changes rapidly, the sea-state could be forced to leave the equilibrium it had before and gradually adapt to a modified equilibrium state associated with the new conditions. During the equilibration process of an out-of-equilibrium sea-state, waves are characterized by evident non-Gaussian features.

Many works have been devoted to the investigation of NED induced by depth change (see recent updates in Zhang et al., 2019; Bolles et al., 2019; Kashima and Mori, 2019; Trulsen et al., 2020; Zheng et al., 2020; Zhang and Benoit, 2021; Lawrence et al., 2021; Li et al., 2021; Zhang et al., 2022, and the references therein). In summary, previous studies showed that the non-equilibrium sea-state after a rapid depth variation is characterized by:

(a) local variations of skewness and kurtosis of the free surface elevation (FSE) $\eta$ (see their definitions in eqs. (1) and (2)) and of the kinematics beneath waves;
(b) a rapid exchange of wave energy within the wave spectrum, including excitation of both free and bound high-order harmonics and enhancement of wave energy level in the low-frequency (infra-gravity) range;
(c) a heavy tail in the wave height distribution, i.e., higher occurrence probability of extreme waves.

Zhang et al. (2022) (hereafter abbreviated as ZBM2022) investigated the equilibration process of an out-of-equilibrium sea-state with refined numerical simulations. They showed that the sea-state equilibration process after the depth transition consists of two stages: a fast modulation stage in a relatively short scale (within a few wavelengths) in which the NED mainly affects the wave statistics; and a slow modulation stage in a longer scale (within dozens of wavelengths) in which considerable wave spectral changes develop. The cross-spectral analysis indicates that the spectral evolution in the longer scale mainly results from the interaction among free wave harmonics. It is evidently different from that in the shorter scale, where the bound and free super-harmonics play a dominant role. The incident wave nonlinearity affects the magnitude of NED, but not the length of the scales. However, the detailed investigation of the statistical distributions of wave height and FSE during the two-scale non-equilibrium wave evolution is neither covered in ZBM2022 nor studied in any existing literature. Thus, the present work is dedicated to studying the wave statistical parameters and the distributions during the stabilization process of the out-of-equilibrium sea-state.

The statistical parameters are essential for describing the sea-state and building probability distributions. Evaluation of statistical parameters from the results of phase-averaged models is therefore of great practical interest. In particular, the skewness and kurtosis of FSE, as measures of the sea-state non-Gaussianity, received considerable attention. Based on the Zakharov equation, Janssen (2003) and Mori and Janssen (2006) elaborated the computation of the dynamic component of kurtosis resulting from nonlinear four-wave interaction. The formulation is valid for weakly nonlinear deep-water waves with narrow spectral bandwidth. Subsequently, the computation has been extended to consider the

effects of finite water depth, and finite spreading angle (see e.g. Janssen and Onorato, 2007; Janssen and Bidlot, 2009; Mori et al., 2011; Fedele, 2015). The computation of skewness and the bound component of kurtosis which are caused by the asymmetric wave profile (bound harmonics) are elaborated in Tayfun and Lo (1990) and Janssen (2009). The "total" kurtosis is then considered to be the algebraic sum of the bound and the dynamic kurtosis components. Ducrozet and Gouin (2017) compared the measured skewness and kurtosis with the theoretical computations following the formulations of Fedele et al. (2016) near the depth transition area (namely, the so-called short scale). Here, we extend the discussion to the long scale of the non-equilibrium phenomenon and focus on the 1DH case. In addition, the relationship between skewness and kurtosis will be discussed.

Considering a sea-state to be stationary, ergodic, and a Gaussian random process, the FSE $\eta$ then follows the normal distribution. By further assuming the wave spectrum to be narrow-banded, the wave height $H$ (assumed to be twice the crest height) follow the Rayleigh distribution (Longuet-Higgins, 1952). These assumptions allow for straightforward estimations of all statistical properties of the sea-state, thus are of great convenience in the design practices (Longuet-Higgins, 1957; Ochi, 1998). However, they are sometimes inaccurate in real seas, since the wave nonlinearity, the finite spectral bandwidth, and the sea-state non-stationarity significantly influence the distribution functions, in particular regarding the shape of their tails (Nieto-Reyes, 2021). For the extreme waves in severe sea-states, wave nonlinearity is prominent. Both second-order (Longuet-Higgins, 1963; Tayfun and Alkhalidi, 2016) and third-order nonlinear effects (Mori and Yasuda, 2002; Janssen, 2003; Onorato et al., 2006; Mori and Janssen, 2006; Cherneva et al., 2009) could result in strong non-Gaussian statistics. In coastal zones, the finite water depth effects could also lead to non-Gaussianity of the wave statistics (Bitner, 1980; Ochi and Ahn, 1995; Cherneva et al., 2005). The stationary assumption is often taken for granted in theoretical statistical models, but when the environment undergoes rapid changes, like strong depth variations in the present study, the non-stationarity affects the wave statistics significantly (Shemer et al., 2010; Mendes et al., 2022; Mendes and Kasparian, 2023). Currently, the study of the non-stationary effects on anomalous wave statistics is still limited. In addition to putting forward a new model that requires a vast amount of validation work, it is also of great interest to examine how well the existing distribution models perform in describing non-equilibrium wave statistics. The applicability and limitations of several popular statistical models that are either frequently studied in the scientific literature or used in engineering practices will be analyzed and discussed.

The remainder of this paper is organized as follows: in Section 2, the adopted statistical models are listed, including the theoretical or semi-empirical statistical distributions of FSE $\eta$ and wave height $H$ (whose formulations are gathered in appendices A and B, respectively), the theoretical formulations for statistical moments are introduced. In Section 3, the configurations and characteristics of the numerical simulations are briefly introduced. In Section 4, the theoretical statistical models are assessed regarding their validity and accuracy in describing the out-of-equilibrium sea-state. The analysis of the relation between skewness and kurtosis, and the statistics of the maximum wave heights, are discussed in Section 5. Conclusions are summarized in Section 6.

## 2. Statistical moments and models of probability distributions

### 2.1. Statistical moments

The first four statistical moments of the FSE play crucial roles in characterizing a sea-state. The first- and second-order moments are the mean $\langle\eta\rangle$ ($\langle\cdot\rangle$ denotes a mean operation) and variance $\sigma^2$, respectively, which are sufficient for fully describing the statistics of a Gaussian sea-state (Longuet-Higgins, 1952). The non-Gaussianity can be characterized by the third- and fourth-order moments, respectively skewness and kurtosis, which are defined as:

$$\lambda_3(\eta) \equiv \frac{\left\langle (\eta - \langle\eta\rangle)^3 \right\rangle}{\sigma^3} = \left\langle \bar{\eta}^3 \right\rangle, \tag{1}$$

$$\lambda_4(\eta) \equiv \frac{\left\langle (\eta - \langle\eta\rangle)^4 \right\rangle}{\sigma^4} = \left\langle \bar{\eta}^4 \right\rangle, \tag{2}$$

with $\bar{\eta} \equiv (\eta - \langle\eta\rangle)/\sigma$ the normalized FSE.

In a Gaussian sea-state, $\lambda_3 = 0$ and $\lambda_4 = 3$ are expected. The skewness $\lambda_3$ offers a measure of the mean asymmetry of the wave profile in the vertical direction. A positive value of $\lambda_3$ indicates sharp wave crests associated with second-order bound harmonics. $\lambda_4$ is a measure of the "tailedness" of a probability distribution in comparison to the normal distribution. Its value is positively correlated with the probability of extreme events. The excess kurtosis, defined as $\lambda_{40} \equiv \lambda_4 - 3$, is also frequently used.

### 2.2. FSE distributions

For uncorrelated random waves, the probability density function (PDF) of the FSE, $p(\eta)$, follows the Gaussian distribution. As the relative water depth decreases or the wave steepness increases, the wave nonlinearity is enhanced, and the non-Gaussianity develops in the random process of $\eta$. Various FSE distributions have been put forward to model wave nonlinearity with different approaches. As summarized by Machado (2003), the statistical models of FSE are typically built based on: (a) statistical moments; (b) process characteristic function; (c) transform Gaussian method. In addition, (d) empirical fitting method and (e) asymptotic (Stokes) representation of FSE are commonly used as well. The list of the PDF models is long and will not be elaborated here, the readers are suggested to refer to Ochi (1998); Tayfun and Alkhalidi (2020) for a more detailed introduction to various FSE distributions.

In this work, seven PDF models of FSE are assessed:

1. the classical Gaussian distribution (denoted as $p_G$) is firstly selected for comparative purposes.

2. the extension of $p_G$ with the Gram-Charlier series introduced by Bitner (1980) (denoted as $p_B$). Compared to $p_G$, the $p_B$ model is built using $\lambda_3(\eta)$ and $\lambda_4(\eta)$ as inputs, and shows good performance in describing the non-Gaussian statistics in coastal areas with varying bathymetry.
3. the second-order deep-water model by Socquet-Juglard et al. (2005) (denoted as $p_{SJ}$), which has been applied to study non-stationary wave statistics.
4. the second-order finite-depth model in its simplified form, recently introduced by Tayfun and Alkhalidi (2020) (denoted as $p_{TA}$). It shows plausible performance in predicting statistics of shoaling waves.
5. the exponential Gamma model used by Herrman et al. (1997) (denoted as $p_H$).
6. the Gamma model recently proposed by Bolles et al. (2019) (denoted as $p_\Gamma$). This model successfully predicts the increased occurrence of extreme waves encountering a vertical step of the seabed in experiments with irregular long-crested waves.
7. a new model proposed in this study in lognormal form (denoted as $p_{LN}$).

These models belong to three different types: $p_G$ corresponds to a linear model; both $p_{SJ}$ and $p_{TA}$ models are based on Stokes second-order wave theory, while $p_B$, $p_\Gamma$, $p_H$ and $p_{LN}$ models are all 3-parameters models involving location, scale, and shape parameters, which are obtained by fitting the statistical moments. Detailed formulations of the above-mentioned statistical models are given in appendix A.

## 2.3. Wave height distributions

In a Gaussian sea-state, the distribution of the normalized wave height $P(\bar{H})$ (with $\bar{H} \equiv H/\sigma$) follows the Rayleigh distribution, with $H$ assumed to be twice the wave crest height. This approximation is valid only for linear waves with narrow spectral bandwidth in deep water. For nonlinear waves with arbitrary spectral width in real seas, the Rayleigh distribution is known to over-predict the occurrence probability of large waves (Forristall, 1978).

The development of theoretical nonlinear wave height distribution models mainly relies on the narrow-band assumption. Complicated physical processes, like wave breaking, finite depth, or broad-band effects, are often incorporated into statistical models by means of empirical fitting. Again, a comprehensive review of all existing wave height distributions is beyond the scope of the present work. Many of the existing models are listed and commented on, for instance in Katsardi et al. (2013); Karmpadakis et al. (2020) and references therein.

We choose to focus on the following five models among the ones presented in the literature:

1. a linear quasi-deterministic (Q-D) model by Boccotti (2000) (denoted as $P_B$) which has been validated using measured oceanic waves (Tayfun and Fedele, 2007). We choose $P_B$ instead of the Rayleigh model to represent the expected distribution of linear waves with finite spectral bandwidth;
2. a well-known distribution of Weibull shape by Forristall (1978) (denoted as $P_F$), which is an empirical model fitted for the field measurement in the Gulf of Mexico;
3. an expansion of the Q-D model by Alkhalidi and Tayfun (2013) (denoted as $P_{AT}$) with Gram-Charlier series that incorporates both third-order nonlinearity and finite spectral bandwidth;
4. a variation of the empirical model by Glukhovskiy (1966) as proposed by Van Vledder (1991) (denoted as $P_{GK}$), this kind of model is considered because it takes the local water depth into account.
5. a composed Weibull distribution introduced in Wu et al. (2016) (denoted as $P_W$), which is fitted for shallow water waves with wave height up to the breaking limit.

The formulations of these models are given in appendix B. It should be pointed out that the recent work of Mendes et al. (2022) is not included for comparison here, despite its ability in predicting the non-stationary wave statistics of waves over an uneven bottom. Because the target of this work is to evaluate statistical models for practical purposes, the simplicity of the formulation is an important criterion when choosing candidate models.

## 2.4. Computation methods of skewness and kurtosis

### *2.4.1. Evaluation of skewness and kurtosis based on PDF*

Given a continuous PDF of the FSE, $p(\bar{\eta})$, the skewness $\lambda_3$ and kurtosis $\lambda_4$ can be evaluated as

$$\lambda_3(\bar{\eta}) = \int_{-\infty}^{\infty} \bar{\eta}^3 p(\bar{\eta}) \mathrm{d}\bar{\eta}, \tag{3}$$

$$\lambda_4(\bar{\eta}) = \lambda_{40} + 3 = \int_{-\infty}^{\infty} \bar{\eta}^4 p(\bar{\eta}) \mathrm{d}\bar{\eta}. \tag{4}$$

For an empirical PDF built from a FSE time record, with the integrals evaluated in a discrete fashion, the empirical skewness and kurtosis are obtained. If the input $p(\bar{\eta})$ corresponds to a theoretical distribution, then the outputs of eqs. (3) and (4) represent the model predictions of skewness and kurtosis.

Typically, a theoretical PDF of FSE requires statistical moments as inputs to determine its characteristics, like position, scale, and shape parameters. The capability of a PDF model of this type can be illustrated by the Skewness–Kurtosis (S–K) plot, expressing kurtosis as a function of skewness. In particular, the S–K plot shows the range of non-Gaussianity that a PDF model could describe. In the S–K plot, it is known that a 2-parameter distribution is represented by a single point (e.g. the Gaussian distribution corresponds to $(\lambda_3, \lambda_{40}) = (0, 0)$), a three-parameter model corresponds to a curve, and a model of higher dimension covers an area. Particular interest is paid here to 3-parameter models, not only because they are commonly used in practical applications, but also because they can provide a simple and general (sea-state independent) S–K relation. Furthermore, in some

of the models, physical properties like wave steepness, relative water depth or bottom gradient are required as inputs for better describing the wave characteristics. Such models predict S–K relations which are dependent on the sea-state, and the model capability largely depends on how well the physical properties describe the wave characteristics.

Among the seven PDF models considered in this study (see section 2.2 and appendix A), Fig. 1 shows the S–K relationships for $p_G$, $p_H$, $p_{TA}$, $p_\Gamma$ and $p_{LN}$. These five models are also compared with the result by Mori and Kobayashi (1998) who derived a theoretical expression of the S–K relationship for waves of second-order in steepness, having a quadratic form, as a nonlinear reference:

$$\lambda_{40} = \frac{16}{9}\lambda_3^2. \tag{5}$$

As already mentioned, the Gaussian distribution $p_G$ is the linear expectation of the S–K relationship, as the single point $(0, 0)$. Model $p_B$ is not included because it is a four-parameter model, requiring both $\lambda_3$ and $\lambda_{40}$ as inputs, therefore the results of eqs. (3) and (4) are analytically equal to the input $\lambda_3$ and $\lambda_{40}$. The second-order model $p_{SJ}$ is also excluded because it is a three-parameter model (requiring first- and second-order moments and wave steepness as inputs). With wave steepness required as an input, the S–K relationship predicted by $p_{SJ}$ is not universal but dependent upon the sea-state.

The S–K relation given by $p_H$ is implicit, but can be evaluated numerically by varying $a_0$ from $-\infty$ to 0 in eqs. (A.6–A.7). The S–K relation of $p_{LN}$ model is obtained numerically from eqs. (A.25–A.26), and can be approximated as $\lambda_{40} \approx 1.85\lambda_3^2$. The simplified model $p_{TA}$ considers $\lambda_3$ as the only input, resulting in a one-parameter model. In this model, $\lambda_3$ is used to determine other parameters, including $\lambda_{40}$. The simplified model $p_{TA}$ predicts a cubic relationship between $\lambda_3$ and $\lambda_{40}$, but the contribution of the cubic term is rather low in comparison to the quadratic term, see eq. (A.14). According to Bolles et al. (2019), the Gamma distribution is robust and provides "near-complete" statistical description of the surface displacement for the out-of-equilibrium sea-state after abrupt depth change. The S–K relation predicted by $p_\Gamma$ is of a simple quadratic form, as provided in eq. (A.21).

It is noted from Fig. 1 that, in essence, the S–K relationships predicted by $p_H$, $p_{TA}$, $p_\Gamma$, and $p_{LN}$ follow a similar general trend. Remarkably, models $p_{LN}$ and $p_H$ predict a quite similar S–K relationship, also very close to the second-order analytical result eq. (5) by Mori and Kobayashi (1998). When compared with these three models, $p_\Gamma$ and $p_{TA}$, which are very close to each other, give lower $\lambda_{40}$ for the same $\lambda_3$ in the range $\lambda_3 > 0.3$. This first assessment of models already provides valuable insight into their modelling capabilities. In particular, $p_{LN}$ and $p_H$ on the one hand, and $p_\Gamma$ and $p_{TA}$ on the other, can be expected to perform almost identically when applied to FSE records.

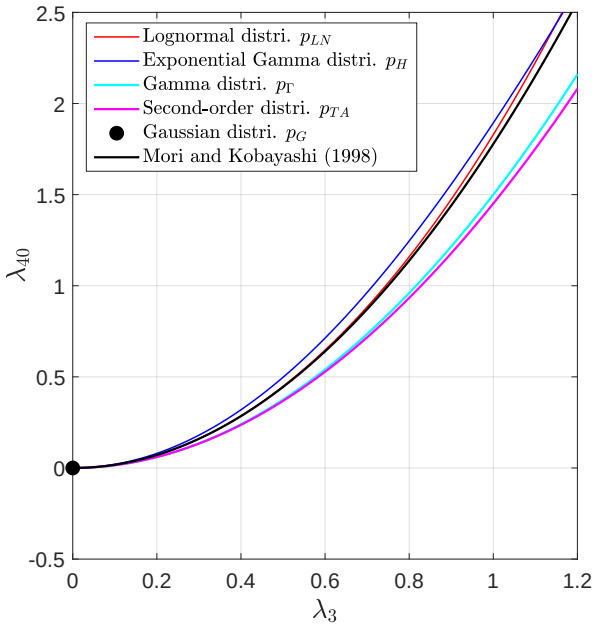


**Figure 1:** Skewness–Kurtosis (S–K) relationships associated with five theoretical probability distributions of FSE, plus the second-order relation (5) from Mori and Kobayashi (1998)

### 2.4.2. *Evaluation of skewness and kurtosis based on wave spectrum*

The theoretical $\lambda_3$ and $\lambda_{40}$ are derived for unidirectional, weakly-nonlinear and narrow-banded waves in the flat bottom case with finite water depth. The wave field is assumed to be homogeneous (i.e. the wave spectrum keeps its shape unchanged during wave evolution). The theoretical skewness $\lambda_{3,NB}$ (with subscript NB denoting narrow band) reads (Fedele et al., 2016)

$$\lambda_{3,NB} = 6\epsilon_m \left(\alpha_s + \Delta\right), \tag{6}$$

where

$$\alpha_s = \frac{3 - Q_m^2}{4Q_m^3}, \Delta = -\frac{1}{4}\frac{gh}{gh - c_{g,m}^2}\left[2\frac{1 - Q_m^2}{Q_m} + \frac{1}{\mu_m}\right], \tag{7}$$

where $k_m$ denotes the wave number corresponding to the mean frequency $\omega_m = m_1/m_0$ via linear dispersion relationship $\omega_m^2 h = g\mu_m Q_m$ (with $m_1$ being the first-order moment of wave spectrum, $\mu_m = k_m h$ and $Q_m = \tanh \mu_m$), $\epsilon_m = k_m\sigma$, $c_{g,m} = \mathrm{d}\omega_m/\mathrm{d}k$.

The theoretical excess kurtosis $\lambda_{40,NB}$ comprises a bound component $\lambda_{40,NB}^b$ due to non-resonant interactions and a dynamic component $\lambda_{40,NB}^d$ due to (quasi-) resonant interactions. It is formulated as

$$\lambda_{40,NB} = \lambda_{40,NB}^b + \lambda_{40,NB}^d. \tag{8}$$

The expression of the bound component $\lambda_{40,NB}^b$ for narrow-banded waves in intermediate water reads (Janssen, 2009;

Fedele et al., 2016)

$$\lambda^{b}_{40,NB} = \frac{4}{3}\lambda^{2}_{3,NB}\left[1 + \frac{\beta_s + \gamma_s}{2\left(\alpha_s + \Delta\right)^2}\right], \tag{9}$$

where

$$\beta_s = \frac{24 + 3\left(1 - Q_m^2\right)^3}{64 Q_m^6}, \quad \gamma_s = -\frac{\alpha_s^2}{2}. \tag{10}$$

Under deep water condition, the dynamic component $\lambda^{d}_{40,NB}$ is expressed as a six-dimensional integral (Janssen, 2003). Assuming a Gaussian-shape spectrum, $\lambda^{d}_{40,NB}$ is proportional to the square of the Benjamin-Feir index (BFI) in the long-time limit (Mori and Janssen, 2006)

$$\lambda^{d}_{40,NB} = \frac{\pi}{\sqrt{3}}\mathrm{BFI}^2, \tag{11}$$

where $\mathrm{BFI} = \sqrt{2\pi}\epsilon_m Q_p$, with $Q_p = 2\int_0^\infty f S^2 \mathrm{d}f / m_0^2$ being a measure of spectral width. The extension of dynamic excess kurtosis to intermediate water depth is achieved by replacing the BFI in eq. (11) with its shallow water version, $B_s$ (Janssen and Onorato, 2007). The expression of $B_s$ is

$$B_s^2 = -\mathrm{BFI}^2 \left(\frac{c_{g,0}}{c_0}\right)^2 \frac{g X_{nl}}{k_0 \omega_0 \omega_0''}, \tag{12}$$

where the subscript 0 denotes the carrier wave's property, $\omega_0''$ denotes the second-order derivative of $\omega_0(k_0)$, and $X_{nl}$ the nonlinear interaction coefficient, see their expressions in Janssen and Onorato (2007).

The previous formulations of $\lambda_{3,NB}$ and $\lambda_{40,NB}$ are subject to several assumptions, including weakly nonlinear and narrow-banded waves, constant water depth, and homogeneous wave field. As was pointed out by Ducrozet and Gouin (2017), applying the formulations of $\lambda_{3,NB}$ and $\lambda_{40,NB}$ in the study of depth-induced NED would more or less violate the above-mentioned assumptions. The parameters $\lambda_{3,NB}$ and $\lambda_{40,NB}$ are key parameters for operational freak wave warning systems, as they can be evaluated on the basis of the phase-averaged output spectrum. Therefore, the performance of these theoretical formulations is worth investigating in the depth-changing areas and for the non-equilibrium sea-states, to update our knowledge of the freak wave warning in such regions.

## 3. Dataset from numerical simulations

In this work, the statistical analysis is carried out on the numerical data obtained recently, whose spectral properties and the corresponding implications on the non-equilibrium wave evolution have been reported in ZBM2022.

The sketch of the bottom profile in the numerical flume is shown in Fig. 2. The bathymetry consists of a deeper flat zone with uniform water depth $h_1 = 0.53$ m and a shallower zone with $h_2 = 0.11$ m, connected by a steep (1/3.81) plane slope. The origin of the $x$-axis is set at the end of the slope. Relaxation zones are set on both sides of the flume for wave generation and absorption, as indicated in Fig. 2. The waves are imposed at $x = -2.7$ m (1.1 m before the up-slope) and damped after $x = 43.6$ m. The incident wave trains are described by the JONSWAP spectrum. The spectral peak period $T_p = 1.1$ s and spectral shape parameter $\gamma = 3.3$ are fixed throughout the study. The significant wave height $H_s$ varies from 0.01 m to 0.035 m for different levels of nonlinearity. The key wave parameters, together with the relative water depth $\mu$, wave steepness $\epsilon \equiv k_p H_s/(2\sqrt{2})$ and Ursell number $U_r = \epsilon/\mu^3$ in both the deeper and the shallower regions are listed in Table 1. For each case, a simulation lasts for 500 s and is repeated 10 times with different sets of random phases.

The numerical investigation is carried out with a fully nonlinear potential flow model, Whispers3D (W3D). Its mathematical formulations, numerical methods, wave generation and damping techniques have been presented in existing literature, (see Raoult et al., 2016; Benoit et al., 2017; Simon et al., 2019; Zhang and Benoit, 2021, for instance), and are not duplicated here. Furthermore, the validation against the experimental results of Trulsen et al. (2020) and the choices of numerical parameters are given in ZBM2022.

**Table 1**
Summary of wave field parameters.

| Case | $H_s$ [m] | $T_p$ [s] | Deeper/shallower area | | |
|---|---|---|---|---|---|
| | | | $\mu$ | $\epsilon$ | $Ur$ |
| 1 | 0.010 | 1.1 | 1.85/0.64 | 0.012/0.020 | 0.0019/0.077 |
| 2 | 0.015 | 1.1 | 1.85/0.64 | 0.019/0.031 | 0.0029/0.116 |
| 3 | 0.020 | 1.1 | 1.85/0.64 | 0.025/0.041 | 0.0039/0.155 |
| 4 | 0.025 | 1.1 | 1.85/0.64 | 0.031/0.052 | 0.0049/0.194 |
| 5 | 0.030 | 1.1 | 1.85/0.64 | 0.037/0.062 | 0.0058/0.232 |
| 6 | 0.035 | 1.1 | 1.85/0.64 | 0.043/0.073 | 0.0068/0.271 |

[a] For all cases, $\gamma = 3.3$, deeper region depth $h_1 = 0.53$ m, shallower region depth $h_2 = 0.11$ m.
[b] Each case is repeated 10 times with different sets of phases.

In the simulations, the NED develops as the incident waves pass over a slope and gradually adapt to the shallower depth after the slope. As was shown in the work of ZBM2022, the NED induced by depth variations comprises two spatial scales. In the relatively shorter scale ($0 \sim 5L_p$ after depth transition), skewness and kurtosis are locally enhanced, whereas the spectral width and peak frequency show limited variations. However, the situation is inverse in the longer scale ($5 \sim 30L_p$ after depth transition): The statistical moments nearly converge to steady levels, meanwhile, the spectral shape significantly modulates in the range around the spectral peak.

The FSE is computed with a constant time step $\Delta t = 0.01$ s (i.e., with sampling frequency 100 Hz) at 116 locations along the flume during the simulations. In each of the two areas close to the relaxation zones, $[-2.7, -2.0]$ m and $[43.0, 43.6]$ m, 5 probes with uneven spacing are set for evaluation of reflection. In the area $[-2, 2.8]$ m, 25 probes with

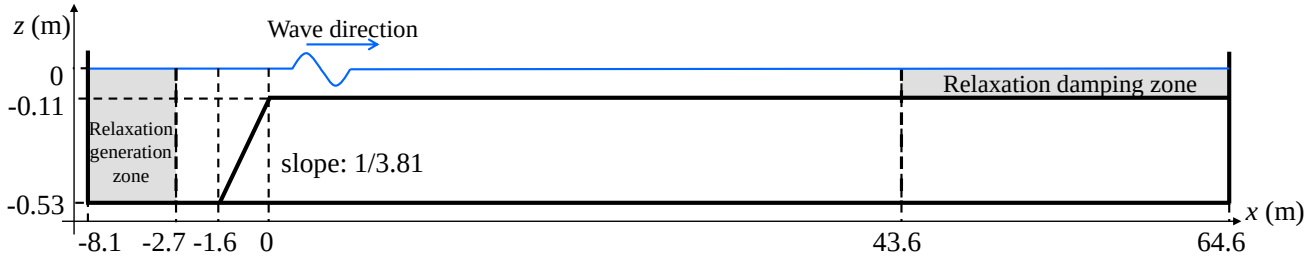


**Figure 2:** Sketch of the bathymetry adopted in the numerical simulations, with $L = 43.6$ m denoting the length of the shallower region, and gray zones representing the relaxation zones.

0.2 m interval, and in the area [3.0, 43.0] m, 81 probes with 0.5 m interval are set to capture the wave evolution in short and long scales, respectively. The results of the validity assessment of the theoretical statistical distributions and moments are displayed at eight representing positions, four located in the short scale and four in the long scale. In the short scale, we choose $x_1 = -2.7$ m (probe #1) located at the wavemaker position, $x_{16} = 0$ m (probe #16) located at the end of the slope, $x_{20} = 0.8$ m (probe #20) where local maximum values of skewness and kurtosis are achieved and $x_{24} = 1.6$ m (probe #24) where skewness and kurtosis decrease significantly. In the long scale, we choose $x_{45} = 10$ m (probe #45) where an evident peak frequency downshift occurs (see Fig. 6 and Fig. 9 in ZBM2022), $x_{65} = 20$ m (probe #65) is still in the range of long-scale non-equilibrium sea-state evolution, $x_{85} = 30$ m (probe #85) where the shallow-water steady states are achieved, and $x_{105} = 40$ m (probe #105) where the sea-states is in a new equilibrium state.

The statistical analyses are performed for all cases and discussed for cases 1, 3, and 6 with increasing $\epsilon$ to illustrate the influence of wave nonlinearity (see Table 1).

## 4. Probability distribution analysis

### 4.1. Assessment of FSE distributions

Fig. 3 shows the FSE distributions in the short scale. We see in Figs. 3(a.1-a.3), the generated waves leaving the wavemaker are Gaussian, this is mainly because the incident sea-states are characterized by relatively small Ursell numbers, and are generated with the linear superposition method in the relaxation zone. The theoretical and empirical distributions are almost superimposed. Comparing probes #16 and #1, it is noticed from Figs. 3(b.1-b.3) that, as waves propagate and pass over the slope, small deviations from the Gaussian distribution develop. The limited deviations from Gaussian can be described by all non-Gaussian models, with only minor differences among them.

As waves propagate over the shallower flat region ($x > 0$ m) to probe #20, the sea-states in all three cases depart significantly from Gaussian distribution. From Figs. 3(c.1-c.3), we notice the sea-states are evidently non-Gaussian, even in case 1 with the mildest wave steepness $\epsilon$. At this location, the level of sea-state non-Gaussianity increases with stronger incident wave nonlinearity. The distributions $p_{LN}$ (green curves), $p_H$ (blue curves) and $p_{TA}$ (cyan curves) all provide quite accurate predictions of the empirical FSE distribution. Notice that the G–C model $p_B$ becomes negative (singular points in log scale) for some negative values of $\bar{\eta}$ in all three cases. Such a non-physical behavior is because negative values in the G–C series could be achieved for large $\lambda_3$ and $\lambda_{40}$, limiting the valid range of G–C type models for strongly nonlinear scenarios. In Figs. 3(d.1-d.3), where the effects of NED decline considerably, the deviations of empirical FSE distributions from $p_G$ decrease correspondingly in comparison to those at probe #20. At this location, $p_H$, $p_{TA}$ and $p_{LN}$ still provide good estimates of the FSE distribution. $p_B$ breaks again for prediction of negative $\bar{\eta}$ in case 6 where the wave nonlinearity is the strongest. The models $p_B$ and $p_{SJ}$ underestimate the probability of large values of $\bar{\eta}$ in all three cases at all positions.

Fig. 4 shows the FSE distributions in the long scale. As a general remark, when waves propagate in the area relatively far from the depth transition, the deviations of the empirical FSE distributions from Gaussian for positive $\bar{\eta}$ are trivial. However, they are still non-Gaussian, in the sense that the probability density of negative $\bar{\eta}$ is clearly lower than Gaussian prediction. It is anticipated that the non-Gaussianity is attributed to the finite water depth effects, as the deviations do not exhibit noticeable change with increasing $\epsilon$ in three cases. Among all models, $p_B$ agrees slightly better with the empirical distribution, and no anomalous negative values appear since the values of $\lambda_3$ and $\lambda_{40}$ remain at low levels. The models $p_H$, $p_{TA}$, and $p_{LN}$ also provide reasonable predictions of the empirical FSE distribution, despite that they overestimated the positive tails in case 6 by about 10%.

As waves propagate over significant depth change, the NED affects the sea-state in both short and long spatial scales. $p_B$ shows the best performance in predicting the long-scale evolution of the empirical PDF, however, it is subject to unrealistic negative PDF in the short scale. $p_{SJ}$ is not able to capture the non-Gaussian behavior in the short scale. Both are not the best choice in the current scenario. The models of Herrman et al. (1997) ($p_H$), Tayfun and Alkhalidi (2020) ($p_{TA}$) and the lognormal distribution ($p_{LN}$) provide excellent predictions of the empirical PDF evolution in the short scale, and the predictions remain acceptable in the long scale. $p_{TA}$ has slightly better performance in the short scale where $\lambda_3$ and $\lambda_{40}$ achieve their maximum values. However, the empirical fitting of Tayfun and Alkhalidi (2020) is valid for $0 < \lambda_3 < 1.5$, the performance of this model may de-

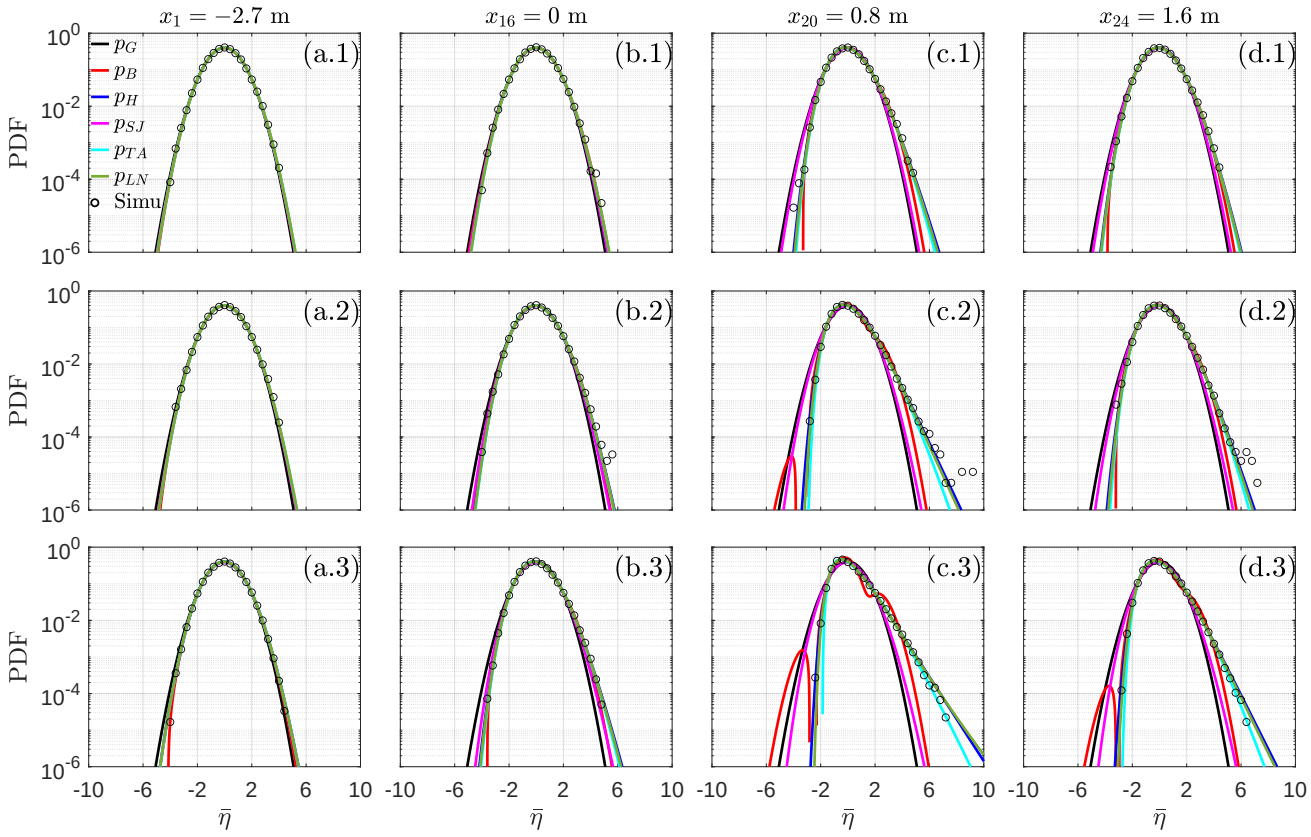


**Figure 3:** Short-scale evolution of PDF of $\bar{\eta}$ in cases 1, 3, and 6 shown in panels (n.1), (n.2) and (n.3) respectively, with n assuming a, b, c, or d, representing the four probe positions. The locations are indicated on top of the upper row of panels.

teriorate when applied out of this range. Both $p_{LN}$ and $p_H$ require $\lambda_3$ being non-negative, and provide very similar predictions. The parameters in lognormal distribution are relatively more straightforward to compute than those in Herrman et al. (1997) (comparing eq. (A.6) to eqs. (A.23-A.25)). We recall that all theoretical models are used here out of their nominal validity range due to the inhomogeneity of the media and the non-stationarity of the random process. Therefore, the good performances of $p_H$, $p_{TA}$ and $p_{LN}$ indicate the robustness of these three models.

It should be noted that the Gamma distribution $p_\Gamma$ is not included in Figs. 3 and 4 only for the sake of clarity. The prediction of $p_\Gamma$ is nearly identical with that of $p_{TA}$. So the comments to $p_{TA}$ also apply to $p_\Gamma$. Besides, the predictions of $p_H$ and $p_{LN}$ are very similar to each other as well. Such observations are in line with the indication of the S–K relationship provided in Fig. 1: the S–K relationships predicted by $p_H$ and $p_{LN}$ being very close, the same for $p_\Gamma$ and $p_{TA}$. The S–K relationship is considered a robust indicator of the capability of the FSE distribution model.

## 4.2. Assessment of wave height distributions

The distribution of crest-to-trough height of linear waves is dependent on the spectral width. The quasi-determinism (Q-D) model of Boccotti (2000) ($P_B$) is adopted as the linear expectation for finite-banded waves, instead of the Rayleigh distribution. At the wave maker, the sea-states are quasi-Gaussian according to previous analyses of the FSE distributions. Correspondingly, we observe in Figs. 5(a.1-a.3) that the Q-D model $P_B$ describes well the empirical wave height distributions in all three cases at the wavemaker. The nonlinear models (except for $P_W$ shown in cyan lines) show similar results as $P_B$. In Figs. 5(a.1-d.1), the agreement between $P_B$ and empirical wave height distributions remains good at four positions in case 1, because of the low level of sea-state nonlinearity.

In Figs. 5(c.2-c.3), the empirical distributions deviate significantly from $P_B$. It indicates more freak waves manifest than expected at the position where $\lambda_3$ and $\Lambda_{40}$ achieve their maximum values and NED takes the most pronounced effects. The deviation is more or less captured by $P_{AT}$, although $P_{AT}$ underestimates the probability of freak waves. In Figs. 5(d.2-d.3), as the NED effects decline, the probability of freak waves decreases correspondingly. The model $P_{AT}$ shows better performance compared with the others. Overall, for the short-scale evolution of a non-equilibrium sea-state, $P_{AT}$ seems to be suitable for describing the wave height distribution. It is noticed that, for sea-states with non-trivial nonlinearity, the model $P_W$ performs better (see panels c.2 and c.3 for instance) than for the (quasi-) linear waves.

The wave height distribution evolution in the long scale is shown in Fig. 6. We see that nonlinear models $P_F$ and $P_{GK}$ provide similar predictions as the linear model $P_B$ in all panels. These three models overestimate the tail part of the

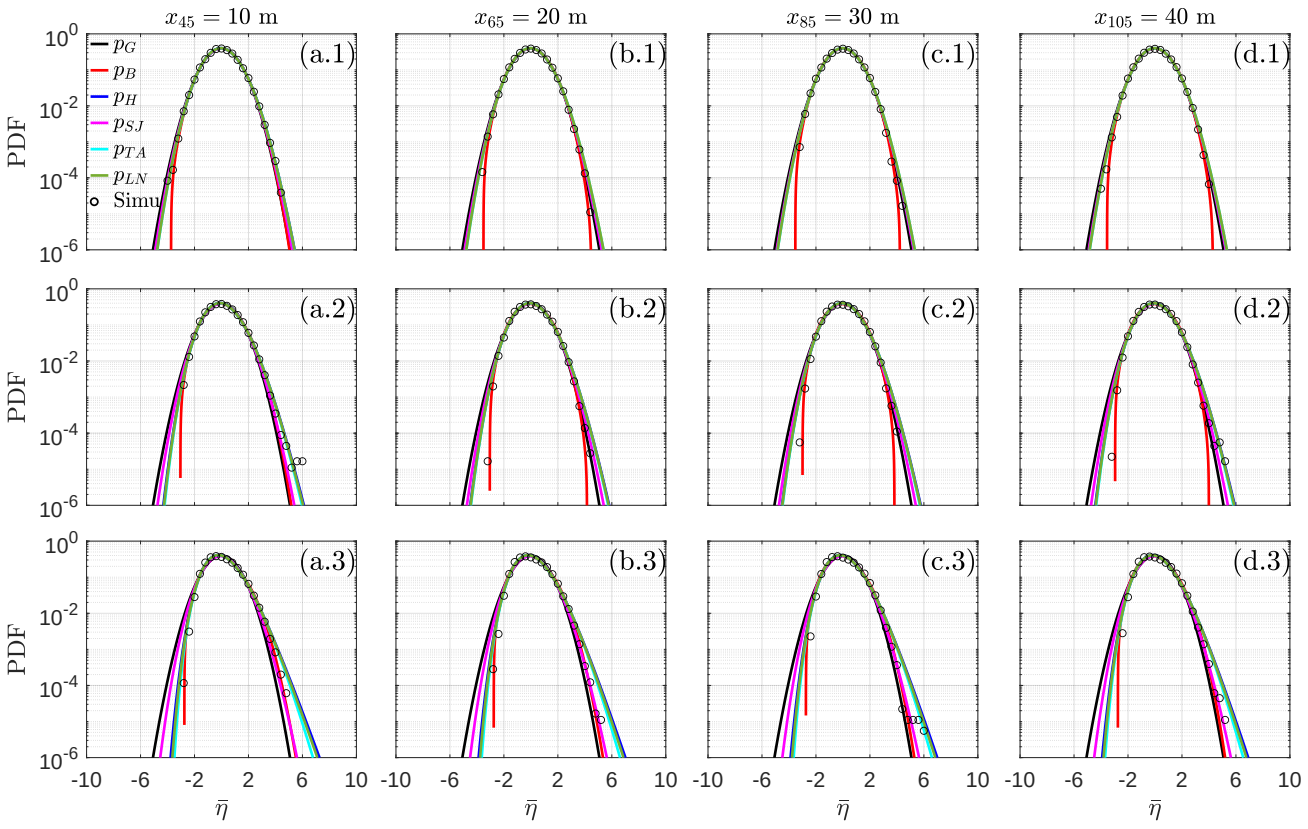


**Figure 4:** Long-scale evolution of PDF of $\bar{\eta}$ in cases 1, 3, and 6 shown in panels (n.1), (n.2) and (n.3) respectively, with n assuming a, b, c, or d, representing the four probe positions. The locations are indicated on top of the upper row of panels.

empirical distribution to some extent. In the meantime, $P_{AT}$ captures well the evolution of empirical distribution in cases 1 and 3, shown in Figs. 6(a.1-d.1) and (a.2-d.2). In case 6, the empirical distribution for large waves cannot be predicted by any of the tested models, and the empirical wave height probability falls between the predictions of $P_{AT}$ and $P_B$. In the long-scale evolution, the NED results in lower freak wave probability than the linear expectation, therefore it plays a role in "protecting" the engineering structures. However, a wave height distribution model that is suitable for the description of the sea-states with negative excess kurtosis is still needed.

It is worth mentioning that an accurate prediction of wave statistics in the long scale after the depth transition is very challenging. This is because the depth-induced NED would considerably enhance the nonlinearity of high waves, resulting in more active and rapid energy transfer among free and bound super-harmonics. As a result, the re-established sea-state in the long scale is clearly different from what we expect for a similar sea-state but in a steady regime. In other words, the sea-state keeps a memory about the short-scale NED evolution, and successful prediction of long-scale statistical evolution would require the short-scale wave dynamics as prior information.

In both Figs. 5 and 6, when the wave nonlinearity is relatively low, $P_W$ considerably overestimates the empirical wave height distribution for the largest waves. However, when waves are of relatively high nonlinearity, $P_W$ provides relatively reasonable predictions, although not the best when compared to other models. This is expected if one notices that, given $\sigma \to 0$ (i.e. infinitesimal wave height), $\bar{H}_{max} \to \infty$, $K \to 2$, $\mu_0 \to 1/2\alpha$ ($\mu_0 \approx 2.273$ for $\alpha = 0.22$), we have $\xi \to 0$ (for the definitions of these parameters, the readers are referred to appendix B). As a result, the Pareto part of eq. (B.41) would take the following limit:

$$\lim_{\xi\to 0} P_W(\bar{H}) = \lim_{\xi\to 0} \left[1 + \xi\left(\bar{H} - 4\right)\right]^{-\frac{1}{\xi}} \exp\left(-\mu_0\right) \tag{13}$$

$$= \lim_{\xi\to 0} \exp\left(4 - \frac{1}{2\alpha} - \bar{H}\right), \tag{14}$$

where $\bar{H} \in [4, \infty]$. Being proportional to $\exp\left(-\bar{H}\right)$, $P_W(\bar{H})$ significantly overpredicts the probability of large waves (with $\bar{H} > 4$) in sea-states with low nonlinearity. Thus, it should be noticed that $P_W(\bar{H})$ is more suitable for sea-states with relatively strong nonlinearity. This feature was not mentioned in Wu et al. (2016).

## 5. Analysis of statistical moments and freak wave statistics

### 5.1. Spatial evolution of skewness and kurtosis

We first show the comparison between the empirical $\lambda_3$ and $\lambda_{40}$ and those predicted by the narrow-band weakly non-

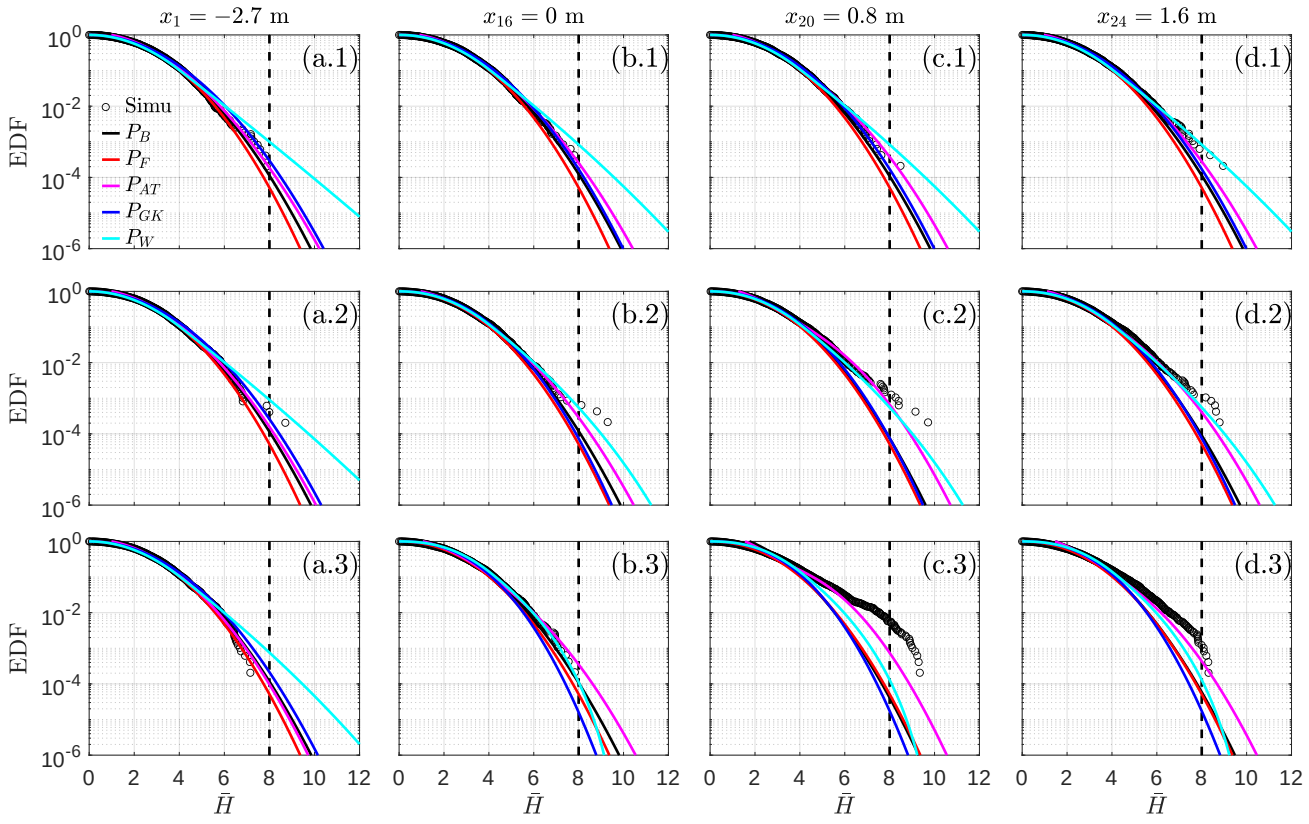


**Figure 5:** Short-scale evolution of EDF of $\bar{H}$ in cases 1, 3, and 6 shown in panels (n.1), (n.2) and (n.3) respectively, with n assuming a, b, c, or d, representing the four probe positions. The locations are indicated on top of the upper row of panels.

linear theory in cases 1, 3, and 6. Although the underlying assumptions of the theory are more or less violated in the current scenario, this comparison is still considered valuable because (1) Judging from the spectral peak wave steepness $\varepsilon = 0.062 \sim 0.073$ given in Tab. 1, waves are of intermediate nonlinearity in an averaged sense, despite some highly nonlinear extreme waves occurred locally over the submerged step; (2) The bottom slope changes the water depth, yet our main interest lies in the deeper and the shallower flat regions; (3) Waves evolve in an inhomogeneous medium, with spectral shape modulating in space. The computations of $\lambda_{3,NB}$ and $\lambda_{40,NB}$ adopt the local spectral parameters as inputs, such that the inhomogeneity is incorporated to some extent.

Fig. 7 shows the spatial evolution of skewness and kurtosis: the evolution in the shorter scale is shown in panels (a.1) and (b.1) and in the longer scale in panels (a.2) and (b.2). It is noticed in Fig. 7(a.1) that $\lambda_{3,NB}$ agrees well with the empirical $\lambda_3$ before the end of the slope ($x < 0$ m). As waves enter the shallower region ($x > 0$ m), the empirical $\lambda_3$ evidently modulates due to the development and then attenuation of NED during the two-stage non-equilibrium wave evolution. However, $\lambda_{3,NB}$ remains nearly unchanged and is always lower than the empirical $\lambda_3$. This is because, as indicated in eq. (6), $\lambda_{3,NB}$ is linearly dependent on $\epsilon_m = k_m\sigma$ for a given relative water depth. Within the shallower flat region, neither the mean steepness $\epsilon_m$ nor the mean relative water depth $k_m h_2$ changes remarkably. As a result, $\lambda_{3,NB}$ varies very mildly over the shallower region and is incapable of capturing the local enhancement of skewness due to NED.

In Figs. 7(b.1–b.2), the total theoretical kurtosis $\lambda_{40,NB}$ (plotted with solid lines) and its bound component $\lambda^b_{40,NB}$ (plotted with dash lines) are shown, the differences between the two curves indicate the dynamic contributions of kurtosis $\lambda^d_{40,NB}$. The evolution trend of $\lambda^b_{40,NB}$ is very similar to that of $\lambda_{3,NB}$, capturing the evolution of $\lambda_{40}$ only before the shallower region where the NED effects are small. This is again because $\lambda^b_{40,NB}$ is proportional to $\epsilon_m^2$ for a given relative water depth. The evolution of $\lambda_{40,NB}$ is, however, significant due to the negative contributions of the dynamic $\lambda^d_{40,NB}$. The negative $\lambda^d_{40,NB}$ is achieved resulting from the interaction factor $X_{nl}$, which is 0 for $k_m h = 1.363$ and decreases monotonically to $-\infty$ as $k_m h \to 0$. It means that the wave nonlinear interaction that stabilizes the sea-state rapidly becomes strong in shallow water, soon the weakly nonlinear assumption is violated and eq. (11) becomes ineffective. For this reason, $\lambda_{40,NB}$ overestimates the stabilizing effects due to nonlinear interactions in sufficiently shallow water (after $x = 0$ in the present cases), and the evolution of $\lambda_{40,NB}$ over the submerged step is not shown.

Then, $\lambda_3$ and $\lambda_{40}$ are computed by integrating the theoretical FSE distribution models $p_H$, $p_{TA}$, and $p_{LN}$ for cases 1, 3, and 6. The short-scale evolutions of $\lambda_3$ and $\lambda_{40}$ are presented in Figs. 8(a.1–b.1) and the long-scale evolution in Figs. 8(a.2–b.2). Only the legends for case 1 is shown (due to

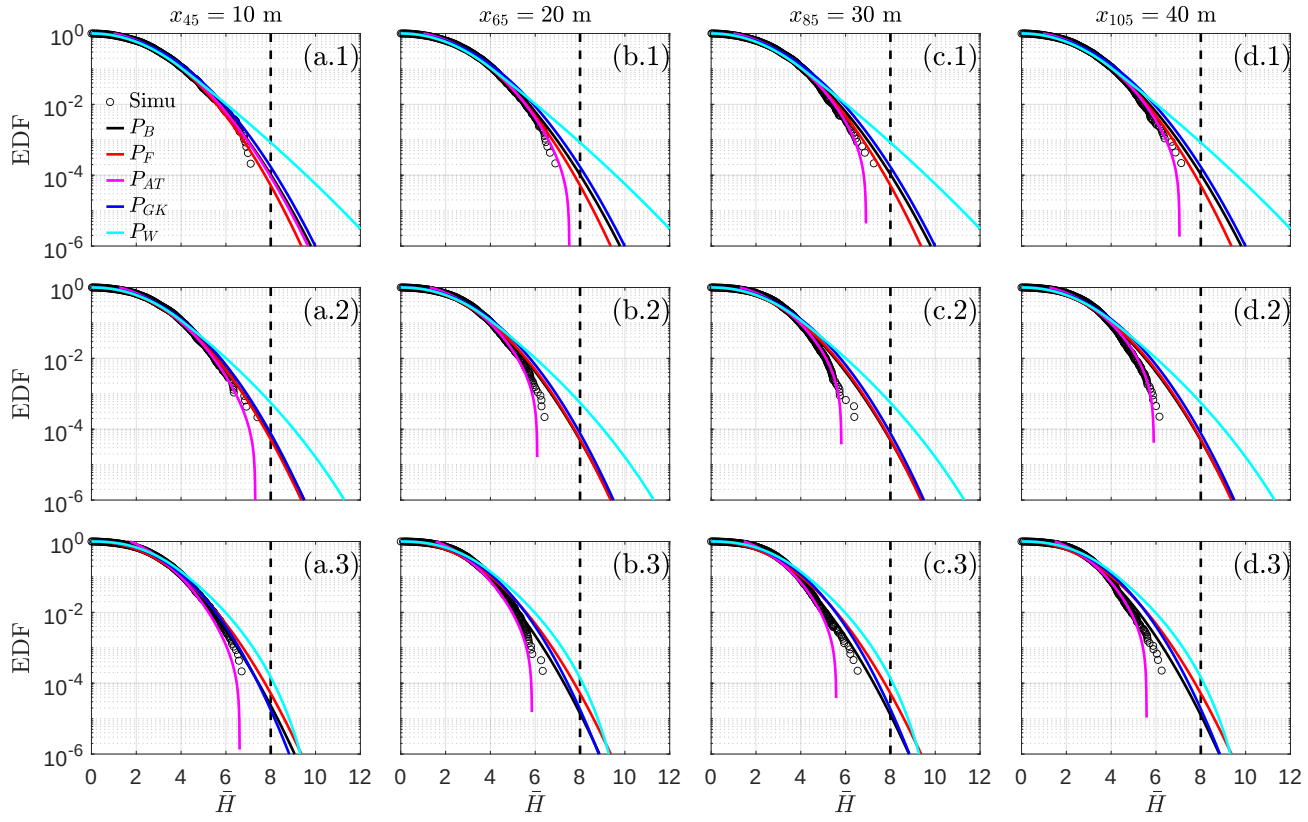


**Figure 6:** Long-scale evolution of EDF of $\bar{H}$ in cases 1, 3, and 6 shown in panels (n.1), (n.2) and (n.3) respectively, with n assuming a, b, c, or d, representing the four probe positions. The locations are indicated on top of the upper row of panels.

space limitations), the curves of other cases follow the same configuration of line styles but in different colors (case 3 in red; and case 6 in blue). It is seen in Figs. 8(a.1) and (a.2) that, the curves of $\lambda_3$ obtained by integrating $p_H$ and $p_{LN}$ are superimposed with the empirical $\lambda_3$ in both scales of all three cases. This is simply because both models require the empirical $\lambda_3$ as an input. The computed skewness $\lambda_3$ based on $p_{TA}$ is very similar to the empirical $\lambda_3$, only some minor differences are noticed in case 6. It also indicates that fitted expressions (A.11–A.13) are quite accurate for relatively large skewness. In Fig. 8(b.1), it is seen that, in all three cases, the predictions of $\lambda_{40}$ by $p_H$ and $p_{LN}$ are very close to each other and are slightly higher than the empirical values of $\lambda_{40}$. Meanwhile, $\lambda_{40}$ evaluated with $p_{TA}$ considerably underestimates the empirical kurtosis, indicating that the regression formula (A.14) is relatively less accurate in comparison to $p_H$ and $p_{LN}$. In Fig. 8(b.2), it is interesting to see that all three models provide positive predictions of $\lambda_{40}$, and none of them can capture the negative empirical $\lambda_{40}$ in the simulations.

From Fig. 7, it is seen that the skewness and kurtosis are significantly underestimated by the weakly nonlinear theory after significant depth change, especially in the short scale of a few wavelengths. This is because the second-order moment $\sigma^2(\eta)$ and peak frequency $f_p$ do not respond significantly to the NED after depth change, it is therefore challenging to compute $\lambda_3$ and $\lambda_4$ on the basis of the local estimates of steepness and relative water depth. Therefore, applying the theoretical formulations of kurtosis without any further treatment in a freak wave warning system could result in a considerable underestimation of freak wave risks in coastal areas with uneven bathymetry. From Fig. 8, it seems that evaluating kurtosis based on skewness is a promising alternative.

## 5.2. Relationship between skewness and kurtosis

As discussed in sec. 2.4.1 and illustrated in Fig. 1, the S–K plot characterizes the non-Gaussianity of a sea-state. It has been used in the analysis of non-equilibrium statistics, for instance in Kashima and Mori (2019). With eqs. (3) and (4), the S–K relation of any FSE PDF is readily obtained (either analytically or numerically). Here, the sea-state independent S–K relations predicted by models $p_H$, $p_{TA}$, $p_\Gamma$, and $p_{LN}$ are compared with the empirical S–K relations at all probes in the simulations of cases 1, 3 and 6.

The corresponding results are shown in Fig. 9(a–c). Although the spatial information is lacking in the S–K plot, it can still be recognized by tracing the evolution of $\lambda_{40}$: the kurtosis $\lambda_{40}$ starts from approximately 0 at $x = -2.7$ m, achieves its global maximum at $x = 0.8$ m and a secondary maximum at $x = 3.5$ m, and decreases to the steady level in the end of the flume at $x = 40L_{p,2}$. To distinguish the short-scale and the long-scale evolution, the simulated empirical S–K curves in the long scale ($x > 5L_{p,2}$) are plotted with

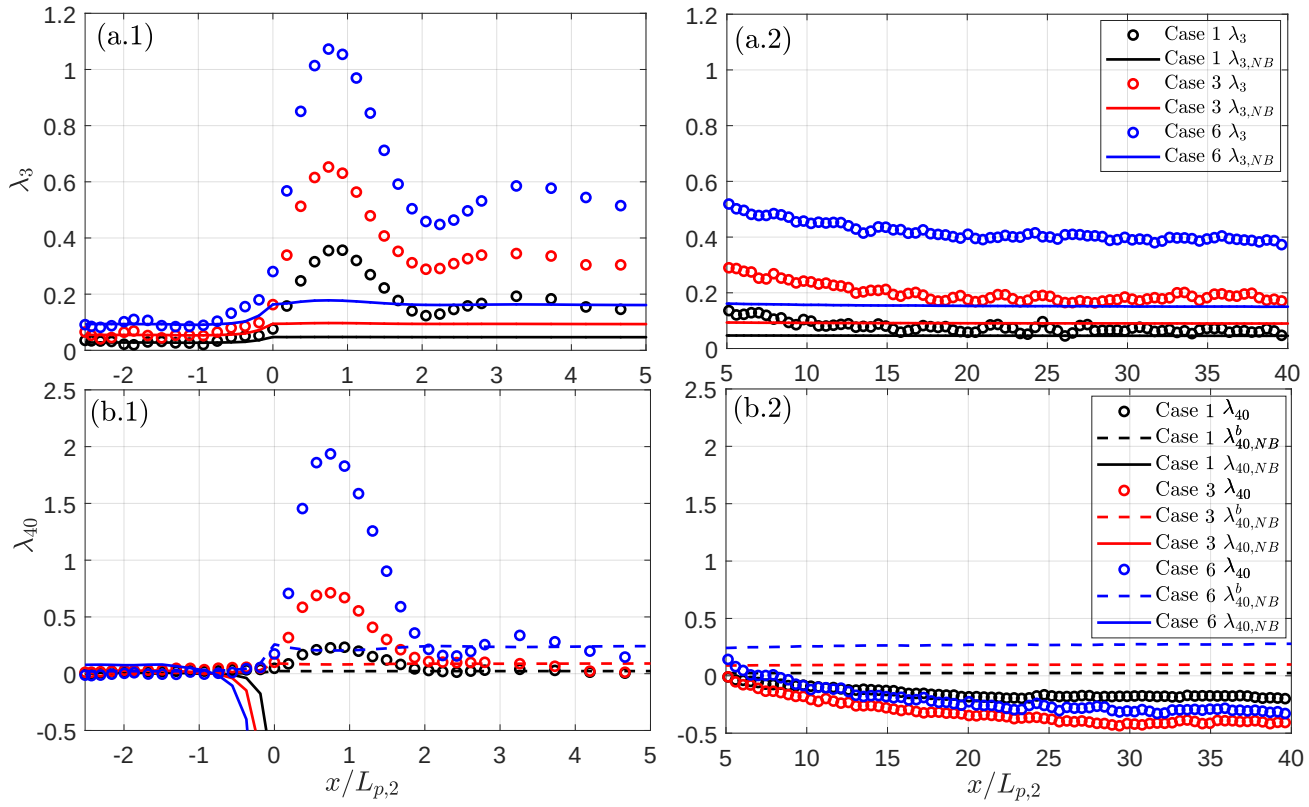


**Figure 7:** Comparison of the spatial evolution of simulated empirical $\lambda_3$ (upper panels) and $\lambda_{40}$ (lower panels) with that of the theoretical $\lambda_{3,NB}$ and $\lambda_{40,NB}$ in cases 1, 3, and 6. The left panels correspond to short-scale evolution and the right panels to long-scale evolution.

dashed lines, with the characteristic locations marked in the figure. Comparing panels (a–c) corresponding to cases 1, 3 and 6 respectively, it is noticed that, although the levels of incident wave nonlinearity and maximum values of $\lambda_3$ and $\lambda_{40}$ vary among the three cases, the evolution of $\lambda_{40}(\lambda_3)$ follows the same trend in all three cases. In Fig. 9(c), the empirical S–K relationship extracted from the experimental results of Trulsen et al. (2020) before and over the flat shallower region (i.e., without the de-shoaling area) is superimposed as an additional verification of the S–K relationships. It is seen that the empirical S–K relationship is close to the predictions of $p_H$ and $p_{LN}$ (thus well captured by the second-order theoretical prediction by Mori and Kobayashi (1998)).

The excess kurtosis $\lambda_{40}(\lambda_3)$ first increases then decreases in the short scale ($x < 5L_{p,2}$), because of the development and attenuation of the NED effect over the shallower region. The increase and decrease trends of $\lambda_{40}(\lambda_3)$ are slightly different, i.e., different values of $\lambda_{40}$ can be achieved for the same $\lambda_3$. This feature is more evident for larger $\lambda_3$. The evolution $\lambda_{40}(\lambda_3)$ is well described by the S–K relations of the exponential Gamma distribution $p_H$ and the lognormal distribution $p_{LN}$. The S–K relations based on the Gamma distribution $p_\Gamma$ and the second-order distribution $p_{TA}$ slightly underestimate $\lambda_{40}$. In general, all theoretical S–K relations provide reasonably good estimation of $\lambda_{40}(\lambda_3)$. The good agreement between the empirical S–K relation and the theoretical predictions of $p_{LN}$ in the short scale ($x < 5L_{p,2}$) indicates that $\lambda_3$ and $\lambda_{40}$ follow a polynomial relationship. The $p_{LN}$ model introduced in this work shows some merits: On the one hand, the S–K relationship of $p_{LN}$ coincides with the prediction of second-order wave theory, and it could be used for waves of higher-order nonlinearity (for example, in our case 6, the waves are of third-order nonlinearity on average, according to the Le Méhauté's diagram). On the other hand, it is capable of predicting the non-equilibrium statistics in both short and long scales (as illustrated in sec. 4), without the limitation of anomalous negative PDF values as in the G–C type models nor the concerns on the unrealistic troughs in second-order Stokes-type models (Tayfun and Alkhalidi, 2020).

We noticed in Figs. 9(a–c) that none of the four models can explain the secondary local peaks of kurtosis corresponding to the point (0.36,0.16) in panel (b) and (0.57,0.26) in panel (c). These secondary peaks appear around $x = 5.3$ m, as the NED decays in the short scale. Moreover, as waves propagate in the longer scale ($x > 5L_{p,2}$, marked with dotted black lines), the empirical skewness $\lambda_3$ continues to decrease, and the corresponding kurtosis $\lambda_{40}$ does not reduce back to 0, but to a negative value. Again, the S–K relationship in the long scale cannot be predicted by any of the theoretical PDF models considered here. The secondary local peak of empirical $\lambda_{40}$ in the short scale and the negative val-

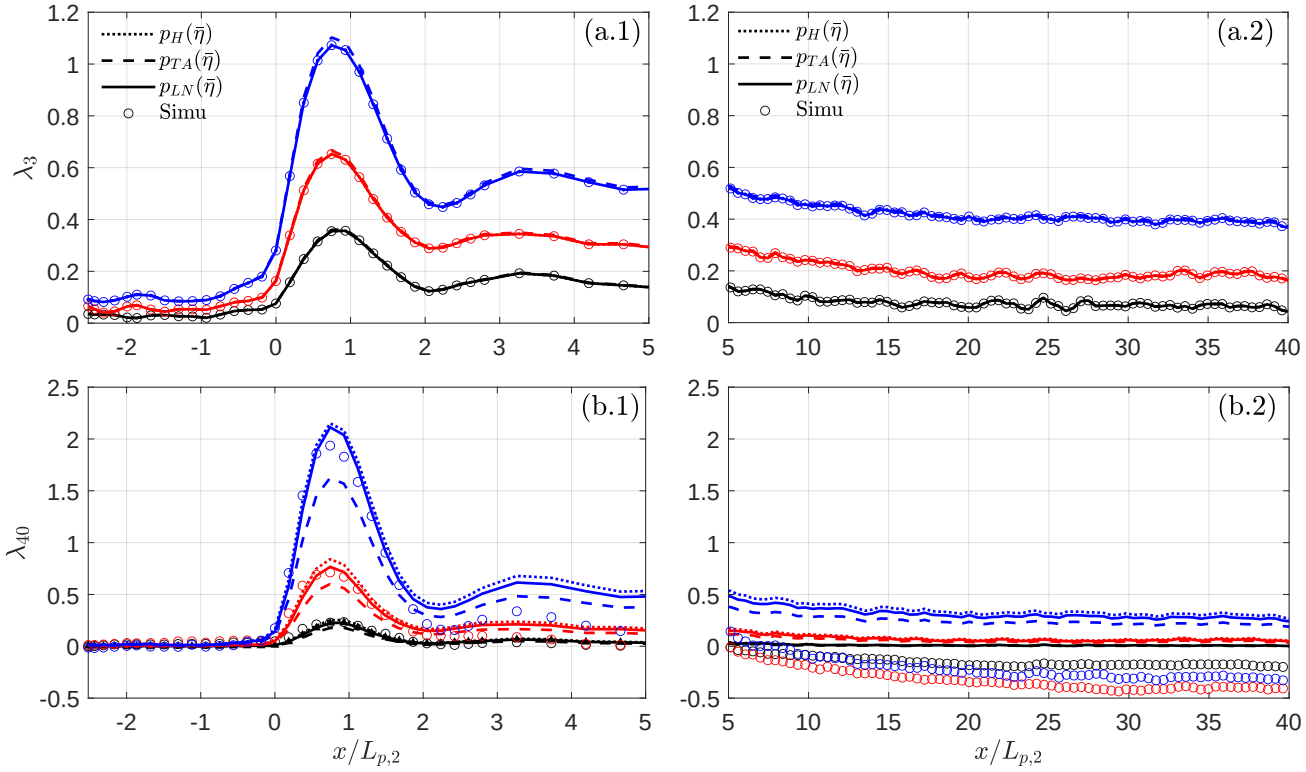


**Figure 8:** Evolution of $\lambda_3$ (upper panels) and $\lambda_{40}$ (lower panels) computed using the integration of three theoretical models of FSE distribution. The left panels correspond to short-scale evolution and the right panels to long-scale evolution. The results of different models are marked with different line styles, and the line colors indicate different test cases (black for case 1, red for case 3 and blue for case 6).

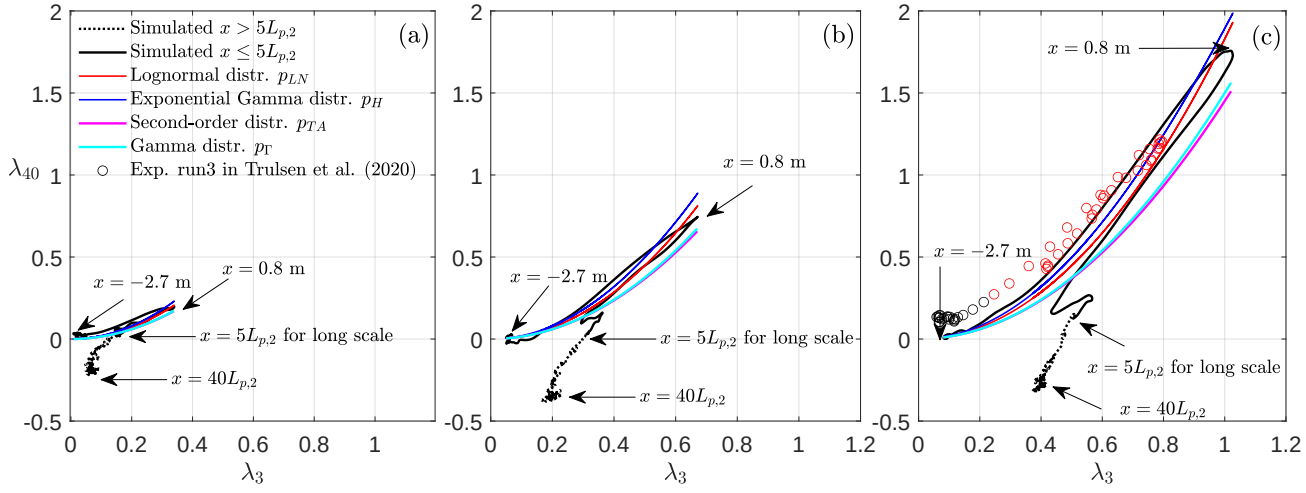


**Figure 9:** Evolution of excess kurtosis $\lambda_{40}$ as a function of skewness $\lambda_3$, predicted by four theoretical models in (a) case 1, (b) case 3 and (c) case 6. The simulated empirical S–K relationships are marked with different line styles for the short scale (black solid line) and the long scale (black dash line) sea-state evolution. Theoretical model predictions are marked with different colors, as indicated in the legend box of panel (a).

ues achieved in the long scale, indicate that re-equilibration process of an out-of-equilibrium sea-state reflects a complex physics. To better describe the different trends of S–K relation during the two-scale sea-state equilibration process, a four- (or even more) parameter model would be needed.

A good S–K relation indicates that a theoretical PDF model describes well the underlying physics, and could be applied for several practical purposes. Being able to express $\lambda_{40}$ as a function of $\lambda_3$ offers the possibility to reduce the dimension of inputs, and to simplify the formulations of other

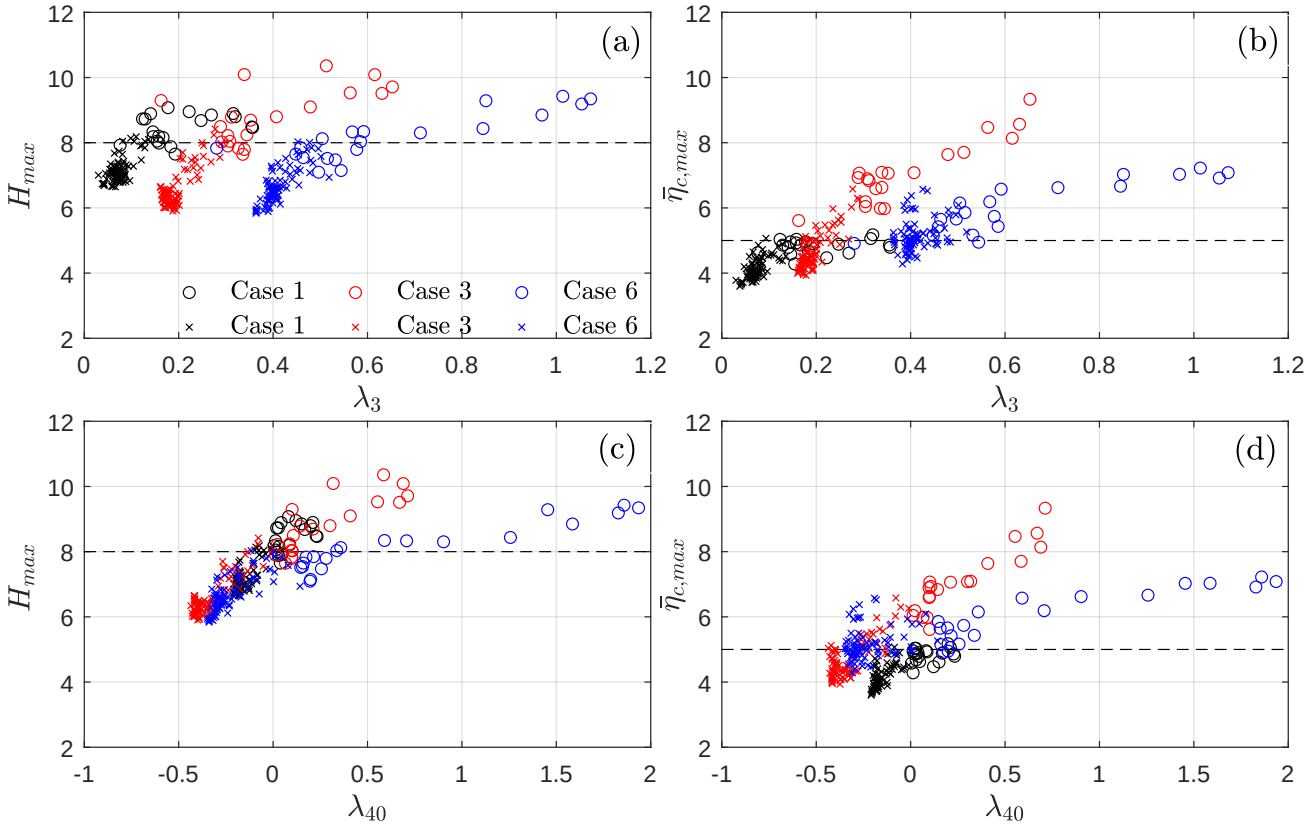


**Figure 10:** Variations of maximum crest-to-trough wave heights $\bar{H}_{max}$ (upper panels a and b) and maximum wave crest $\bar{\eta}_{c,max}$ (lower panels c and d) as functions of $\lambda_3$ (left panels a and c) and as functions of $\lambda_{40}$ (right panels b and d) for the cases 1 (black symbols), 3 (red symbols) and 6 (blue symbols). In all panels, hollow circles denote the results in the short scale ($x \in [0, 5]L_{p,2}$), and crosses in the long scale ($x \in [5, 40]L_{p,2}$).

statistical models. For example, as the crest height or wave height distributions of G–C type require both $\lambda_3$ and $\lambda_{40}$ as inputs, they could be simplified as a function of $\lambda_3$ only. A suitable S–K relation could also be used for the estimation of kurtosis from in-situ measurements. Because the evaluation of kurtosis is more sensitive to the sampling frequency than skewness, the results could be different if the highest waves were not properly captured by the measuring device. Finally, in a freak wave warning system, which often depends on the computation of skewness and kurtosis, a proper S–K relationship of a simple form could reduce the computation effort considerably.

### 5.3. Relationship between the maximum wave statistics and statistical moments

In the literature, there is some indication that both the maximum wave height $\bar{H}_{max}$ and the maximum crest height $\bar{\eta}_{c,max}$ (i.e. the maximum FSE $\bar{\eta}_{max}$ of a wave identified with the zero-up-crossing method) are positively correlated with $\lambda_3$ and $\lambda_{40}$ in the short scale after strong depth variations (see e.g. Kashima and Mori, 2019). Here, we extend the discussion of the relationship between the statistical maximums and high-order moments to the long-scale wave evolution.

In Fig. 10, both $\bar{H}_{max}$ and $\bar{\eta}_{c,max}$ are shown as functions of $\lambda_3$ and $\lambda_{40}$. Again, cases 1, 3, and 6 are analyzed to illustrate the effects related to incident wave nonlinearity. In order to include some spatial information in this figure, the maximum waves recorded in the short scale ($x < 5L_{p,2}$) are marked with hollow circles, and those in the long scale ($x > 5L_{p,2}$) with crosses. The criteria for freak waves are represented by the horizontal dash lines, indicating $\bar{H}_{max} = 2H_s/\sigma = 8$ or $\bar{\eta}_{max} = 1.25H_s/\sigma = 5$ depending on the case.

From Figs. 10(a–b), the maximum wave height shows different features in short and long scales. The maximum wave heights $\bar{H}_{max}$ taking place in the short scale are mostly above the freak wave criterion, whereas the majority of the $\bar{H}_{max}$ values recorded in the long scale are below the criterion. This observation implies that the choice of the length $5L_{p,2}$ is fair in the sense that it separates two regions with different wave dynamics and freak wave risks. The maximum wave height $\bar{H}_{max}$ is correlated with $\lambda_3$ and $\lambda_{40}$ in a nearly linear way, especially for $\bar{H}_{max}$ below the freak wave criterion. The incident wave nonlinearity leads to a shift in the relationship between $\bar{H}_{max}$ and $\lambda_3$.

From Figs. 10(c–d), it is seen that the maximum wave crest height $\bar{\eta}_{c,max}$ exceeds the freak wave criterion in both short and long scales. This is related to the asymmetric wave profile due to finite water depth effects, wave crests being higher in magnitude than wave troughs. Notice that the mag-

nitudes of some maximum crest heights in Figs. 10(c) and (d) are close to their counterparts in Figs. 10(a) and (b). It means that the troughs of these waves are so shallow that the contribution of wave troughs to the wave height becomes small. $\bar{\eta}_{c,max}$ shows a nearly linear correlation with $\lambda_3$, however such a dependence is not clear for $\lambda_{40}$.

## 6. Conclusion

Recent works (e.g. Trulsen et al., 2012; Viotti and Dias, 2014; Trulsen et al., 2020) have shown that the freak wave occurrence probability could be enhanced due to the NED effects when a (quasi-) steady sea-state passes over a steep shoal or seabed step, and then propagates in a new, shallower water depth. According to ZBM2022, in the new depth environment, the NED first develops to reach its maximum level within one wavelength (i.e., $x < L_{p,2}$) and then attenuates rapidly in the near field after the depth change (the so-called "short scale" $x \in [0, 5]L_{p,2}$), the NED continuously affects the sea-state over a long distance (the so-called "long scale" $x \in [5, 40]L_{p,2}$), and finally, the sea-state adapts to the reduced water depth and reaches a corresponding new equilibrium state. In ZBM2022, the characteristics of the two-scale equilibration process have been discussed based on spectral analyses. The present work continues the investigation of this equilibration process based on the same dataset as reported in ZBM2022, focusing on the non-stationary statistics caused by the NED and aiming at providing recommendations for coastal engineering applications. In particular, the validity and accuracy of several theoretical or semi-empirical distributions of FSE (7 models) and crest-to-trough wave height (5 models) have been assessed regarding the representation of non-equilibrium statistics.

The main conclusions from this assessment of theoretical distribution models are summarized below:

- for FSE: in the short scale, the exponential Gamma distribution $p_H(\bar{\eta})$ of Herrman et al. (1997) and Kobayashi et al. (1998) and the lognormal distribution $p_{LN}(\bar{\eta})$ introduced in the present work both predict the evolution of FSE distribution very well, even when non-Gaussianity is prominent. The lognormal distribution appears as a balanced choice for practical use, as it shows good performance and is convenient to build. In the long scale, the G–C type model of Bitner (1980) $p_B(\bar{\eta})$ has an overall better performance.
- for the wave height: the model of Alkhalidi and Tayfun (2013) $P_{AT}(\bar{H})$ provides reasonable prediction despite some underestimation for high waves, in both short and long scales. It has been shown that the LoWiSh model $P_W(\bar{H})$ (Wu et al., 2016) would overestimate the probability of high waves when the sea-state is characterized by weak nonlinearity.

The high-order statistical moments, skewness and kurtosis and their correlations with maximums of wave height and FSE, have been investigated as well. The results show that the theoretical computation of $\lambda_3$ and $\lambda_{40}$ based on wave spectrum (derived theoretically by assuming uniform water depth) is not capable of describing the non-equilibrium statistics in the short scale. This is especially true for the excess kurtosis, usually considered as a proxy for freak wave probability. The theoretical prediction overestimates the stabilizing effects due to nonlinear wave-wave interaction for shallow water depth, this is because the contribution of the dynamic component exponentially decreases toward $-\infty$ when $kh < 1.363$. In the long scale, the theoretical excess kurtosis does not capture the slightly negative values observed in the simulations. It should be noticed that, for a freak wave warning system based on $\lambda_3$ and $\lambda_{40}$ analytically determined from the wave spectrum, the freak wave probability could be considerably underestimated for rapidly changing environments due to NED effects. Particular attention should be paid to the areas with strong depth transitions, like trenches, submerged mounds or bars, coral reefs, etc. The analyses of the FSE and wave height maximums and statistical moments show that the maximum wave height correlates with $\lambda_3$ and $\lambda_{40}$ in a nearly linear manner, yet the correlation between maximum FSE and $\lambda_{40}$ is not so clear.

The S–K relationships (expressing excess kurtosis $\lambda_{40}$ as a function of skewness $\lambda_3$) have been estimated by integrating four theoretical FSE distribution models and compared with the simulation results. In the short scale, the rapid and large variations of $\lambda_{40}(\lambda_3)$ are closely described by the S–K relationships of the exponential Gamma distribution $p_H$ and the lognormal distribution $p_{LN}$ for the three cases considered with increasing wave nonlinearity (see figure 9). In contrast, the S–K relationships of the Gamma-type $p_\Gamma$ and the second-order $p_{TA}$ distributions slightly underestimate $\lambda_{40}$. The lognormal model $p_{LN}$ characterizes the S–K relation approximately as a simple quadratic expression(see eq. (A.28)), which is also in line with the prediction of second-order model of Kobayashi et al. (1998). Such a feature of $p_{LN}$ is of great potential in practical applications, like reducing the model dimension of wave height distributions or accelerating the computation in a freak wave warning system.

For future works, we aim to enlarge the database of non-equilibrium waves by performing more experiments and numerical simulations with different wave and bottom configurations in 2DH setup, and by collecting in-situ measurements under rapid changes of environment. Other models of distribution of wave height, e.g. the recent second-order model put forward by Mendes et al. (2022), will be considered and compared with data. Other statistical properties of the non-equilibrium sea-states deserve to be investigated as well, for example, the distributions of wave crest, wave envelope, wave phase, and maximum wave height.

## CRediT authorship contribution statement

**Jie Zhang:** Simulation performing, Data processing, Writing – Original draft preparation. **Yuxiang Ma:** Discussion, Revising – contribution to the final version of the manuscript. **Michel Benoit:** Conceptualization of this study, Supervis-

ing the findings of this work, Revising – contribution to the final version of the manuscript.

## 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.

## Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grant No. 52101301), and the China Postdoctoral Science Foundation (Grant No. 2023T160078, 2021M690523). The authors would also like to acknowledge Dr. Mehmet Aziz Tayfun, for providing numerous insightful comments.

## Appendix A Distributions of FSE

### A.1 Gaussian distribution (Longuet-Higgins, 1952)

Assuming the sea-state is a linear superposition of an infinite number of sinusoidal waves with random amplitudes and phases, the normalized FSE $\bar{\eta}$ is governed by the normal (Gaussian) distribution

$$p_G(\bar{\eta}) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{\bar{\eta}^2}{2}\right) \tag{A.1}$$

### A.2 Gram-Charlier type distribution (Bitner, 1980)

Assuming the waves are weakly nonlinear, stationary, ergodic, and a quasi-normal random process, the distribution of $\bar{\eta}$ is expressed in Edgeworth's form of G–C series (Longuet-Higgins, 1963). Such a model consists of two parts: the linear part (Gaussian distribution) and the nonlinear correction (with skewness $\lambda_3$ and excess kurtosis $\lambda_{40}$ involved in the expression). It has been applied to describe the distributions of FSE and other variables in shallow water (Bitner, 1980; Ochi and Wang, 1984). The formulation of the FSE distribution given in Bitner (1980) is

$$p_B(\bar{\eta}) = p_G\left[1 + \frac{\lambda_3}{6}H_3(\bar{\eta}) + \frac{\lambda_{40}}{24}H_4(\bar{\eta})\right], \tag{A.2}$$

where $H_3$ and $H_4$ denote 3-rd and 4-th order Hermite polynomials, defined as

$$H_3(\bar{\eta}) = \bar{\eta}^3 - 3\bar{\eta}, \quad H_4(\bar{\eta}) = \bar{\eta}^4 - 6\bar{\eta}^2 + 3. \tag{A.3}$$

In a Gaussian sea-state, $\lambda_3 = \lambda_{40} = 0$, thus eq. (A.2) reduces to the Gaussian distribution (A.1). The major disadvantage of a G–C type model is that, in high sea-states, it predicts an EDF of wave height/crest height with local negative gradients, indicating the corresponding probability density to take negative values, which is evidently non-physical.

### A.3 Herrman distribution (Herrman et al., 1997)

The exponential Gamma distribution introduced by Herrman et al. (1997); Kobayashi et al. (1998) has shown good performance in describing the distribution of FSE over a wide range of constant water depth. Based on the assumption of stationarity and ergodicity, it is formulated as (Herrman et al., 1997)

$$p_H(\bar{\eta}) = \left[\Gamma(a_0)\right]^{-1} \sqrt{\psi_1(a_0)} e^{-a_0 y} \exp(-e^{-y}), \tag{A.4}$$

where the formulation of $y(a_0)$ is

$$y = \sqrt{\psi_1(a_0)}\bar{\eta} - \psi(a_0), \tag{A.5}$$

and $a_0$ is the shape parameter, determined from $\lambda_3$ by solving the nonlinear equation

$$\lambda_3 = -\psi_2(a_0)\left[\psi_1(a_0)\right]^{-1.5}. \tag{A.6}$$

In eqs. (A.4–A.6), $\Gamma$, $\psi$, $\psi_1$, and $\psi_2$ denote the gamma, digamma, trigamma, and tetragamma functions, respectively. Given $a_0$ determined from eq. (A.6), the excess kurtosis can be computed as

$$\lambda_{40} = -\psi_3(a_0)\left[\psi_1(a_0)\right]^{-2}, \tag{A.7}$$

where $\psi_3$ denotes the pentagamma function. The shape parameter $a_0$ decreases from $+\infty$ to 0 for $\lambda_3$ varying from 0 to 2. For $\lambda_3 = 0$ and $a_0 \to +\infty$, eq. (A.4) reduces to the Gaussian distribution, and for $\lambda_3 = 2$ and $a_0 \to 0$, eq. (A.4) becomes the exponential distribution.

### A.4 Socquet-Juglard distribution (Socquet-Juglard et al., 2005)

Following the approach of Longuet-Higgins (1963) and assuming waves are narrow-banded, Tayfun (1980) derived a second-order correction to the Gaussian distribution of FSE. Socquet-Juglard et al. (2005) used an approximation of the second-order model of Tayfun (1980) to investigate the evolution of wave statistics during spectral changes (which is also a non-stationary scenario), showing good agreement with the simulated data for at least $\bar{\eta} < 4$. For this reason, the second-order model of Socquet-Juglard et al. (2005) is included for comparison, although the model assumes a flat bottom. The formulation of Socquet-Juglard et al. (2005) is

$$p_{SJ}(\bar{\eta}) = \frac{1 - 7\epsilon^2/8}{\sqrt{2\pi\left(1 + 3G + 2G^2\right)}} \exp\left(-\frac{G^2}{2\epsilon^2}\right), \tag{A.8}$$

where $G = \sqrt{1 + 2\epsilon\bar{\eta}} - 1$. Note that, theoretically, $\epsilon = k_p\sigma$ is computed from the second-order sea-state. In practice, it is computed from the simulated sea-state which is fully nonlinear. This model is valid for $\bar{\eta} > -3/(8\epsilon)$, otherwise the formulation given in eq. (A.8) diverges. As $\epsilon \to 0$, eq. (A.8) reduces to the Gaussian distribution.

### A.5 Tayfun distribution (Tayfun and Alkhalidi, 2020)

When applied to shallow water waves of moderate to high nonlinearity, second-order Stokes-type models are systematically subject to anomalous troughs with spurious (negative) crests instead of usual troughs. This non-realistic feature in second-order models could result in serious errors while modelling wave statistics. Recently, Tayfun and Alkhalidi (2020) put forward a modified second-order model valid for narrow-banded waves in finite water depth, which is free from anomalies for shallow water waves. We adopt the simplified form of this second-order model given in Tayfun and Alkhalidi (2020), as it is of more practical value

$$p_{TA}(\bar{\eta}) = \begin{cases} c_0 p_0(z)/(1+\epsilon' z); & z > 0, \\ c_0 p_0(z)/\exp(\epsilon' z); & -2/\epsilon' < z \le 0, \\ 0; z \le -2/\epsilon', \end{cases} \quad \text{(A.9)}$$

with the following relation between $z$ and $\bar{\eta}$

$$\bar{\eta} = \begin{cases} \left(z + \epsilon' z^2/2 - \eta'_m\right)/\sigma'; & z > 0, \\ \left[z \exp\left(\epsilon' z^2/2\right) - \eta'_m\right]/\sigma'; & -2/\epsilon' < z \le 0. \end{cases} \quad \text{(A.10)}$$

This model is restricted to $\bar{\eta} > -[2\exp(-1)/\epsilon' + \eta'_m]/\sigma'$. The coefficient $c_0 = \sigma'/P_0(2/\epsilon')$. In this formulation, $p_0$ denotes the standard normal PDF with zero mean and $\sigma = 1$, and $P_0$ is the corresponding cumulative distribution.

The input parameters $\eta'_m$, $\sigma'$, $\epsilon' = k_p\sigma'$, denoting the mean, standard deviation and the steepness respectively, are computed from the second-order FSE. To skip the numerical calculation of the second-order model, Tayfun and Alkhalidi (2020) gave the corresponding regression formulations of $\eta'_m$, $\sigma'$, $\epsilon'$ as functions of $\lambda_3$ (valid for $\lambda_3 \in [0, 1.5]$) reported below, including in addition the expression of $\lambda_{40}$

$$\epsilon' = 0.0259\lambda_3^3 + 0.0174\lambda_3^2 + 0.3377\lambda_3, \quad \text{(A.11)}$$

$$\eta'_m = 0.0101\lambda_3^3 - 0.0012\lambda_3^2 + 0.1687\lambda_3, \quad \text{(A.12)}$$

$$\sigma' = 0.0104\lambda_3^3 + 0.0396\lambda_3^2 + 0.0025\lambda_3 + 1, \quad \text{(A.13)}$$

$$\lambda'_{40} = -0.0259\lambda_3^3 + 1.4696\lambda_3^2 + 0.0075\lambda_3. \quad \text{(A.14)}$$

### A.6 Gamma distribution (Bolles et al., 2019)

The Gamma distribution was recently proposed by Bolles et al. (2019) to model the distribution of FSE in the area following an abrupt depth change (with a vertical step), and compared favourably with irregular wave flume experiments. For the branch of positive values of FSE, the Gamma distribution exhibits an exponential decay which is much slower than Gaussian, thus representing an increased occurrence of extreme events in the area following the seabed step. The general 3-parameter form of the Gamma distribution is

$$p_\Gamma(\eta) = \begin{cases} 0 \quad ; \quad \eta \le \epsilon, \\ \dfrac{1}{\lambda^\alpha \Gamma(\alpha)} (\eta - \epsilon)^{\alpha-1} \mathrm{e}^{-\frac{\eta-\epsilon}{\lambda}}; \eta > \epsilon, \end{cases} \quad \text{(A.15)}$$

where $-\infty < \epsilon < +\infty$ is the location parameter, $\lambda > 0$ is the scale parameter and $\alpha > 0$ is the shape parameter. This distribution has the following first four moments

$$\langle\eta\rangle = \epsilon + \lambda\alpha, \quad \text{(A.16)}$$

$$\sigma(\eta) = \lambda\sqrt{\alpha}, \quad \text{(A.17)}$$

$$\lambda_3(\eta) = 2/\sqrt{\alpha}, \quad \text{(A.18)}$$

$$\lambda_4(\eta) = 3 + 6/\alpha. \quad \text{(A.19)}$$

Using eqs. (A.16–A.17), the Gamma PDF for the normalized FSE with zero mean and unit variance can be written as

$$p_\Gamma(\bar{\eta}) = \begin{cases} 0 \quad ; \quad \bar{\eta} \le -\sqrt{\alpha}, \\ \dfrac{\alpha^{\alpha/2}\mathrm{e}^{-\alpha}}{\Gamma(\alpha)} \left(\bar{\eta} + \sqrt{\alpha}\right)^{\alpha-1} \mathrm{e}^{-\sqrt{\alpha}\bar{\eta}}; \bar{\eta} > -\sqrt{\alpha}, \end{cases} \quad \text{(A.20)}$$

which is a function of the shape parameter only, directly related to the skewness after eq. (A.18): $\alpha = 4/\lambda_3^2(\bar{\eta})$. Finally, eq. (A.19) provides a simple and explicit S–K relationship

$$\lambda_{40}(\bar{\eta}) = \frac{6}{\alpha} = \frac{3}{2}\lambda_3^2(\bar{\eta}). \quad \text{(A.21)}$$

### A.7 Lognormal distribution

In the present study, we propose another option for modelling the FSE distribution, using a fitted lognormal function. It is formulated as

$$p_{LN}(\bar{\eta}) = \begin{cases} 0 \quad ; \quad \bar{\eta} < a_p, \\ \dfrac{1}{\left(\bar{\eta} - a_p\right)\tau\sqrt{2\pi}} e^{-\frac{\left[\ln\left(\bar{\eta}-a_p\right)-a_s\right]^2}{2\tau^2}}; \bar{\eta} \ge a_p, \end{cases} \quad \text{(A.22)}$$

where the location $a_p$, scale $a_s$, and shape $\tau$ parameters are required as input. They are determined from the following equations

$$\langle\bar{\eta}\rangle = 0 = a_p + q^{\frac{1}{2}} e^{a_s}, \quad \text{(A.23)}$$

$$\sigma(\bar{\eta}) = 1 = e^{a_s}\sqrt{q^2 - q}, \quad \text{(A.24)}$$

$$\lambda_3(\bar{\eta}) = (q+2)\sqrt{q-1}, \quad \text{(A.25)}$$

where $q \equiv e^{(\tau^2)}$. For the scaled FSE $\bar{\eta}$, the lognormal model requires $\lambda_3(\bar{\eta})$ as the only input. The excess kurtosis $\lambda_{40}(\bar{\eta})$ is given by

$$\lambda_{40}(\bar{\eta}) = (q-1)\left(q^3 + 3q^2 + 6q + 6\right), \quad \text{(A.26)}$$

where $q$ can be obtained from $\lambda_3(\bar{\eta})$ by inverting eq. (A.25):

$$q(\lambda_3) = \left[1 + \frac{\lambda_3}{2}\left(\lambda_3 + \sqrt{\lambda_3^2 + 4}\right)\right]^{1/3} + \left[1 + \frac{\lambda_3}{2}\left(\lambda_3 - \sqrt{\lambda_3^2 + 4}\right)\right]^{1/3} - 1 \quad \text{(A.27)}$$

The S-K relation $\lambda_{40}(\lambda_3)$ can not be obtained in a simple closed form from (A.26) and (A.27), but over the range $\lambda_3(\bar{\eta}) \in [0, 1.5]$, it can be shown that a close approximation is

$$\lambda_{40}(\bar{\eta}) \approx 1.85\,\lambda_3^2(\bar{\eta}). \quad \text{(A.28)}$$

This S-K relation is also quadratic in $\lambda_3(\bar{\eta})$ as for the Gamma distribution eq. (A.21), but with a higher multiplying factor.

To the limit of our knowledge, the lognormal distribution has never been applied to describe the FSE distribution.

## Appendix B Distributions of wave height

### B.1 Boccotti distribution (Boccotti, 2000)

The applicability of Rayleigh distribution for wave heights critically relies on the narrow-band assumption. For finite-banded linear waves, the Q-D model of Boccotti (2000) is adopted as the Gaussian expectation of wave height statistics

$$P_B(z > \bar{H}) = c_0 \exp\left(-c_1 \bar{H}^2\right), \tag{B.29}$$

where the coefficients $c_1$ and $c_2$ are

$$c_0 = \frac{1+b}{\sqrt{2b\,(1+a)}}, \quad c_1 = \frac{\bar{H}^2}{4\,(1+a)} \tag{B.30}$$

with

$$a = \left| \int_0^\infty S(f) \cos\left(2\pi f \tau^*\right) \mathrm{d}f \right| / m_0, \tag{B.31}$$

$$b = \left| \int_0^\infty f^2 S(f) \cos\left(2\pi f \tau^*\right) \mathrm{d}f \right| / m_2, \tag{B.32}$$

with $\tau^*$ denoting the time-lag of the global minimum of the auto-correlation function $\rho(\tau) = \langle \eta(t)\eta(t+\tau) \rangle$. This model describes the distribution of large wave heights ($\bar{H} \gg 1$).

### B.2 Forristall distribution (Forristall, 1978)

Forristall (1978) proposed an empirical model to fit the field measurements of hurricane sea-states in the Gulf of Mexico. This model is often adopted for deep or transitional water depth. It is formulated as

$$P_F(z > \bar{H}) = \exp\left(-\frac{\bar{H}^{2.126}}{8.42}\right), \tag{B.33}$$

### B.3 Alkhalidi distribution (Alkhalidi and Tayfun, 2013)

The G–C type wave height distribution (also known as modified Edgeworth–Rayleigh (MER) distribution) involves kurtosis to describe the wave statistics due to third-order nonlinear interactions (Mori and Janssen, 2006). A more general expression of G–C type model (assuming the wave height is twice the envelope height) is given in Tayfun and Fedele (2007) and reads

$$P_{TF}(z > \bar{H}) = \exp\left(-\frac{\bar{H}^2}{8}\right)\left[1 + \frac{\Lambda}{1024}\bar{H}^2\left(\bar{H}^2 - 16\right)\right]. \tag{B.34}$$

When replacing $\Lambda$ by $\Lambda_{app}$, eq. (B.34) returns to the MER distribution given in Mori and Janssen (2006), which assumes the waves to be narrow-banded.

Combining the G–C type model in eq. (B.34) and the linear Q-D model in eq. (B.29), Alkhalidi and Tayfun (2013) put forward a generalized Boccotti distribution to take the effects of third-order nonlinearity and finite bandwidth into account. It is formulated as

$$P_{AT}(z > \bar{H}) = c_0 \exp\left(-c_1 \bar{H}^2\right)\left[1 + \frac{\Lambda}{16} c_1 \bar{H}^2\left(c_1 \bar{H}^2 - 2\right)\right]. \tag{B.35}$$

This model describes the distribution of large wave heights ($\bar{H} \gg 1$) in relatively deep water, as in the linear Q-D model (Boccotti, 2000). It should be mentioned that we adopt $\Lambda_{app}$ in practice as it does not require the information of $\eta$ a priori and thus is a model for forecast.

### B.4 Glukhovskiy-type distributions

Glukhovskiy (1966) put forward a shallow-water wave height distribution with a Weibull shape. This empirical model is capable of describing wave breaking effects. This model has two modified versions provided by Van Vledder (1991) and Klopman (1996). The formulation of modified Glukhovskiy distribution is

$$P_G(z > \bar{H}) = \exp\left[-A\left(\sigma \bar{H}/H_m\right)^\kappa\right], \tag{B.36}$$

where $\sigma$ denotes the standard deviation of $\eta$.

In the version of Van Vledder (1991), the coefficients $\kappa$ and the mean wave height $H_m$ are evaluated iteratively based on $H_{rms}$, until the following relationships are fulfilled

$$\kappa = \frac{2}{1 - H_m/h}, \tag{B.37}$$

$$H_m = H_{rms}\Gamma\left(\frac{1}{\kappa}+1\right)\left[\Gamma\left(\frac{2}{\kappa}+1\right)\right]^{-1/2}. \tag{B.38}$$

Then, with $\kappa$ determined, the coefficient $A$ is computed as follows

$$A = \left[\Gamma\left(\frac{2}{\kappa}+1\right)\right]^{\kappa/2}. \tag{B.39}$$

Klopman (1996) proposed another version of Glukhovskiy distribution, circumventing the iterative procedure by directly taking $H_m = H_{rms}$, and computing $\kappa$ as

$$\kappa = \frac{2}{1 - 0.7H_{rms}/h}. \tag{B.40}$$

In both versions, $H_{rms} \approx H_s/\sqrt{2}$ is assumed, indicating that a narrow-banded Gaussian process is taken for granted.

### B.5 LoWiSh distribution (Wu et al., 2016)

In the project "Limits on Waves in Shallow Water (LoWiSh)", (Katsardi et al., 2013; Wu et al., 2016), an empirical two-part Weibull-generalized Pareto (WGP) distribution for shallow water wave heights is introduced. In analogy with the composite Weibull distribution (CWD) by Battjes and Groenendijk (2000), WGP involves a Weibull distribution for describing the lower wave heights and a Pareto distribution for the higher wave heights. Compared to CWD, WGP

shows two improvements: (1) the corresponding PDF is continuous; (2) it contains an upper bound for wave height.

The WGP model requires the local parameters including $H_s$, $k_p$ and $h$ as input, and reads

$$P_W(z > \bar{H}) = \begin{cases} \exp\left[-\mu_0 \left(\dfrac{\bar{H}}{4\rho}\right)^K\right], & 0 \le \bar{H} \le 4\rho, \\ \dfrac{\exp\left(-\mu_0\right)}{\left[1+\xi\left(\bar{H}-4\rho\right)\right]^{\frac{1}{\xi}}}, & 4\rho < \bar{H} < \bar{H}_{max}, \end{cases} \tag{B.41}$$

where $\mu_0$, $K$ are the scale and shape parameters of the Weibull distribution respectively, $4\rho$ is the transition wave height of two formulations, $\xi$ the shape parameter of the Pareto distribution, and $\bar{H}_{max}$ the normalized maximum wave height according to the Miche limiting criterion. They are defined as

$$\bar{H}_{max} = 2\beta\pi \frac{\tanh\left(k_p h\right)}{k_p \sigma}, \quad \mu_0 = \frac{1}{\alpha K}, \tag{B.42}$$

$$K = \frac{2}{1-\lambda(H_s/h)^{1.7}}, \quad \xi = \frac{4\rho\alpha}{4\rho - \bar{H}_{max}}. \tag{B.43}$$

In Wu et al. (2016), the recommended values are $\alpha = 0.22$, $\beta = 0.15$, $\lambda = 1$ and $\rho = 1$, and the transition wave height for two models is $\rho H_s \equiv 4\rho\sigma$. We notice a misprint in eq. (21) of Karmpadakis et al. (2020) for the expression of the Pareto distribution.

## References


Alkhalidi, M.A., Tayfun, M.A., 2013. Generalized Boccotti distribution for nonlinear wave heights. Ocean Eng. 74, 101–106. doi:10.1016/j.oceaneng.2013.09.014.

Annenkov, S., Shrira, V., 2015. Modelling the impact of squall on wind waves with the generalized kinetic equation. J. Phys. Oceanogr. 45, 807–812. doi:10.1175/jpo-d-14-0182.1.

Battjes, J.A., Groenendijk, H.W., 2000. Wave height distributions on shallow foreshores. Coastal Eng. 40, 161–182. doi:10.1016/s0378-3839(00)00007-7.

Benjamin, T.B., Feir, J.E., 1967. The disintegration of wave trains on deep water Part 1. theory. J. Fluid Mech. 27, 417–430. doi:10.1017/s002211206700045x.

Benoit, M., Raoult, C., Yates, M.L., 2017. Analysis of the linear version of a highly dispersive potential water wave model using a spectral approach in the vertical. Wave Motion 74, 159–181. doi:10.1016/j.wavemoti.2017.07.002.

Bitner, E.M., 1980. Non-linear effects of the statistical model of shallow-water wind waves. Appl. Ocean Res. 2, 63–73. doi:10.1016/0141-1187(80)90031-0.

Boccotti, P., 2000. Wave Mechanics for Ocean Engineering. Elsevier Science, Oxford.

Bolles, C.T., Speer, K., Moore, M.N.J., 2019. Anomalous wave statistics induced by abrupt depth change. Phys. Rev. Fluids 4, 011801. doi:10.1103/physrevfluids.4.011801.

Chabchoub, A., 2016. Tracking breather dynamics in irregular sea state conditions. Phys. Rev. Lett. 117, 144103. doi:10.1103/physrevlett.117.144103.

Chakrabarti, S.K., 2005. Handbook of offshore engineering. Elsevier B.V.

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, 599–615. doi:10.1016/j.coastaleng.2005.02.006.

Cherneva, Z., Tayfun, M.A., Guedes-Soares, C., 2009. Statistics of nonlinear waves generated in an offshore wave basin. J. Geophys. Res. 114, C08005. doi:10.1029/2009jc005332.

Dematteis, G., Grafke, T., Onorato, M., Vanden-Eijnden, E., 2019. Experimental evidence of hydrodynamic instantons: The universal route to rogue waves. Phys. Rev. X 9, 041057. doi:10.1103/physrevx.9.041057.

Didenkulova, E., Didenkulova, I., Medvedev, I., 2023. Freak wave events in 2005–2021: statistics and analysis of favourable wave and wind conditions. Nat. Hazards Earth Syst. Sci. 23, 1653–1663. doi:10.5194/nhess-23-1653-2023.

Didenkulova, E.G., Pelinovsky, E.N., 2020. Freak waves in 2011–2018. Dokl. Earth Sci. 491, 187–190. doi:10.1134/s1028334x20030046.

Ducrozet, G., Gouin, M., 2017. Influence of varying bathymetry in rogue wave occurrence within unidirectional and directional sea-states. J. Ocean Eng. Mar. Energy 3, 309–324. doi:10.1007/s40722-017-0086-6.

Dysthe, K., Krogstad, H.E., Müller, P., 2008. Oceanic rogue waves. Annu. Rev. Fluid Mech. 40, 287–310. doi:10.1146/annurev.fluid.40.111406.102203.

Fedele, F., 2015. On the kurtosis of deep-water gravity waves. J. Fluid Mech. 782, 25–36. doi:10.1017/jfm.2015.538.

Fedele, F., Brennan, J., Ponce de León, S., Dudley, J., Dias, F., 2016. Real world ocean rogue waves explained without the modulational instability. Sci. Rep. 6, 27715. doi:10.1038/srep27715.

Forristall, G.Z., 1978. On the statistical distribution of wave heights in a storm. J. Geophys. Res. 83, 2353–2358. doi:10.1029/jc083ic05p02353.

Glukhovskiy, B.K., 1966. Investigation of sea wind waves (in Russian), in: Proc. of Sea Climatology Conf., pp. 51–71.

Goda, Y., 2010. Random seas and design of maritime structures, Third Edition. World Scientific Publishing Company.

Godoi, V.A., Bryan, K.R., Stephens, S.A., Gorman, R.M., 2017. Extreme waves in New Zealand waters. Ocean Modell. 117, 97–110. doi:10.1016/j.ocemod.2017.08.004.

Herrman, M.N., Kobayashi, N., Johnson, B.D., Orzech, M.D., 1997. Experiments on surface elevation probability distribution and statistics in surf and swash zones. Technical Report. Research Report No.CACR-97-01. Center for Applied Coastal Research, Ocean Engineering Laboratory, University of Delaware, Newark, Delaware, p. 127, 19716.

Janssen, P.A.E.M., 2003. Nonlinear four-wave interactions and freak waves. J. Phys. Oceanogr. 33, 863–884. doi:10.1175/1520-0485(2003)33<863:nfiafw>2.0.co;2.

Janssen, P.A.E.M., 2009. On some consequences of the canonical transformation in the Hamiltonian theory of water waves. J. Fluid Mech. 637, 1–44. doi:10.1017/s0022112009008131.

Janssen, P.A.E.M., Bidlot, J.R., 2009. On the extension of the freak wave warning system and its verification. ECMWF , Tech. memo 558.

Janssen, P.A.E.M., Onorato, M., 2007. The intermediate water depth limit of the Zakharov equation and consequences for wave prediction. J. Phys. Oceanogr. 37, 2389–2400. doi:10.1175/jpo3128.1.

Karmpadakis, I., Swan, C., Christou, M., 2020. Assessment of wave height distributions using an extensive field database. Coastal Eng. 157, 103630. doi:10.1016/j.coastaleng.2019.103630.

Kashima, H., Mori, N., 2019. Aftereffect of high-order nonlinearity on extreme wave occurrence from deep to intermediate water. Coastal Eng. 153, 103559. doi:10.1016/j.coastaleng.2019.103559.

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. doi:10.1016/j.coastaleng.2012.09.007.

Kharif, C., Pelinovsky, E., Slunyaev, A., 2009. Rogue Waves in the Ocean. Springer Berlin Heidelberg.

Klopman, G., 1996. Extreme wave heights in shallow water, Report H2486. Technical Report. WL/Delft Hydraulics, The Netherlands.

Kobayashi, N., Herrman, M.N., Johnson, B.D., Orzech, M.D., 1998. Probability distribution of surface elevation in surf and swash zones. J. Waterw. Port, Coast. Ocean Eng. 124, 99–107. doi:10.1061/(asce)0733-950x(1998)124:3(99).

Lawrence, C., Trulsen, K., Gramstad, O., 2021. Statistical properties of wave kinematics in long-crested irregular waves propagating over non-

uniform bathymetry. Phys. Fluids 33, 046601. doi:10.1063/5.0047643.

Li, Y., Draycott, S., Zheng, Y., Lin, Z., Adcock, T.A.A., van den Bremer, T.S., 2021. Why rogue waves occur atop abrupt depth transitions. J. Fluid Mech. 919, R5. doi:10.1017/jfm.2021.409.

Longuet-Higgins, M.S., 1952. On the statistical distributions of sea waves. J. Mar. Res. 11, 245–265. doi:citeulike:5677786.

Longuet-Higgins, M.S., 1957. The statistical analysis of a random, moving surface. Philos. Trans. R. Soc. A 249, 321–387. doi:10.1098/rsta.1957.0002.

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. doi:10.1017/s0022112063001452.

Ma, Y., Zhang, J., Chen, Q., Tai, B., Dong, G., Xie, B., Niu, X., 2022. Progresses in the research of oceanic freak waves: mechanism, modeling, and forecasting. Int. J. Ocean Coast. Eng. 4, 2250002. doi:10.1142/S2529807022500026.

Machado, U.B., 2003. Probability density functions for non-linear random waves and responses. Ocean Eng. 30, 1027–1050. doi:10.1016/s0029-8018(02)00083-5.

Mendes, S., Kasparian, J., 2023. Non-homogeneous approximation for the kurtosis evolution of shoaling rogue waves. J. Fluid Mech. 966, A42. doi:10.1017/jfm.2023.453.

Mendes, S., Scotti, A., Brunetti, M., Kasparian, J., 2022. Non-homogeneous analysis of rogue wave probability evolution over a shoal. J. Fluid Mech. 939, A25. doi:10.1017/jfm.2022.206.

Mori, N., Janssen, P.A.E.M., 2006. On kurtosis and occurrence probability of freak waves. J. Phys. Oceanogr. 36, 1471–1483. doi:10.1175/jpo2922.1.

Mori, N., Kobayashi, N., 1998. Nonlinear distribution of neashore free surface and velocity, in: Coast. Eng. 1998, American Society of Civil Engineers. pp. 189–202. doi:10.1061/9780784404119.013.

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, 1484–1497. doi:10.1175/2011jpo4542.1.

Mori, N., Yasuda, T., 2002. Effects of high-order nonlinear interactions on unidirectional wave trains. Ocean Eng. 29, 1233–1245. doi:10.1016/s0029-8018(01)00074-9.

Nieto-Reyes, A., 2021. On the non-Gaussianity of the height of sea waves. J. Mar. Sci. Eng. 9, 1446. doi:10.3390/jmse9121446.

O'Brien, L., Renzi, E., Dudley, J.M., Clancy, C., Dias, F., 2017. Extreme wave events in Ireland: 2012-2016. Nat. Hazards Earth Syst. Sci. Discuss. 1, 1–46. doi:10.5194/nhess-2017-206.

Ochi, M.K., 1998. Ocean Waves: The Stochastic Approach. Cambridge University Press. doi:10.1017/cbo9780511529559.

Ochi, M.K., Ahn, K., 1995. Non-Gaussian probability distribution of coastal waves, in: Coast. Eng. 1994, pp. 482–496. doi:10.1061/9780784400890.037.

Ochi, M.K., Wang, W.C., 1984. Non-Gaussian characteristics of coastal waves, in: Coast. Eng. 1984, pp. 516–531. doi:10.1061/9780872624382.036.

Onorato, M., Osborne, A.R., Serio, M., Cavaleri, L., Brandini, C., Stansberg, C.T., 2006. Extreme waves, modulational instability and second order theory: Wave flume experiments on irregular waves. Eur. J. Mech. B Fluids 25, 586–601. doi:10.1016/j.euromechflu.2006.01.002.

Onorato, M., Residori, S., Bortolozzo, U., Montina, A., Arecchi, F., 2013. Rogue waves and their generating mechanisms in different physical contexts. Phys. Rep. 528, 47–89. doi:10.1016/j.physrep.2013.03.001.

Raoult, C., Benoit, M., Yates, M.L., 2016. Validation of a fully nonlinear and dispersive wave model with laboratory non-breaking experiments. Coastal Eng. 114, 194–207. doi:10.1016/j.coastaleng.2016.04.003.

Shemer, L., Sergeeva, A., Liberzon, D., 2010. Effect of the initial spectrum on the spatial evolution of statistics of unidirectional nonlinear random waves. J. Geophys. Res. 115, C12039. doi:10.1029/2010jc006326.

Simon, B., Papoutsellis, C.E., Benoit, M., Yates, M.L., 2019. Comparing methods of modeling depth-induced breaking of irregular waves with a fully nonlinear potential flow approach. J. Ocean Eng. Mar. Energy 5, 365–383. doi:10.1007/s40722-019-00154-7.

Socquet-Juglard, H., Dysthe, K., Trulsen, K., Krogstad, H.E., Liu, J., 2005. Probability distributions of surface gravity waves during spectral changes. J. Fluid Mech. 542, 195. doi:10.1017/s0022112005006312.

Tayfun, M.A., 1980. Narrow-band nonlinear sea waves. J. Geophys. Res. 85, 1548. doi:10.1029/jc085ic03p01548.

Tayfun, M.A., Alkhalidi, M.A., 2016. Distribution of surface elevations in nonlinear seas, in: Proceedings of Offshore Technology Conference, Kuala Lumpur, Malaysia. doi:10.4043/26436-ms.

Tayfun, M.A., Alkhalidi, M.A., 2020. Distribution of sea-surface elevations in intermediate and shallow water depths. Coastal Eng. 157, 103651. doi:10.1016/j.coastaleng.2020.103651.

Tayfun, M.A., Fedele, F., 2007. Wave-height distributions and nonlinear effects. Ocean Eng. 34, 1631–1649. doi:10.1016/j.oceaneng.2006.11.006.

Tayfun, M.A., Lo, J.M., 1990. Nonlinear effects on wave envelope and phase. J. Waterw. Port Coast. Ocean Eng. 116, 79–100. doi:10.1061/(asce)0733-950x(1990)116:1(79).

Trulsen, K., 2018. Rogue waves in the ocean, the role of modulational instability, and abrupt changes of environmental conditions that can provoke non equilibrium wave dynamics, in: The Ocean in Motion. Springer International Publishing, pp. 239–247. doi:10.1007/978-3-319-71934-4_17.

Trulsen, K., Raustøl, A., Jorde, S., Rye, L.B., 2020. Extreme wave statistics of long-crested irregular waves over a shoal. J. Fluid Mech. 882, R2. doi:10.1017/jfm.2019.861.

Trulsen, K., Zeng, H., Gramstad, O., 2012. Laboratory evidence of freak waves provoked by non-uniform bathymetry. Phys. Fluids 24, 097101. doi:10.1063/1.4748346.

Van Vledder, G.P., 1991. Modification of the Glukhovskiy distribution, Report H1203. Technical Report. WL/Delft Hydraulics, The Netherlands.

Viotti, C., Dias, F., 2014. Extreme waves induced by strong depth transitions: Fully nonlinear results. Phys. Fluids 26, 051705. doi:10.1063/1.4880659.

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. doi:10.1016/j.coastaleng.2016.01.015.

Zeng, H., Trulsen, K., 2012. Evolution of skewness and kurtosis of weakly nonlinear unidirectional waves over a sloping bottom. Nat. Hazards Earth Syst. Sci. 12, 631–638. doi:10.5194/nhess-12-631-2012.

Zhang, J., Benoit, M., 2021. Wave–bottom interaction and extreme wave statistics due to shoaling and de-shoaling of irregular long-crested wave trains over steep seabed changes. J. Fluid Mech. 912, A28. doi:10.1017/jfm.2020.1125.

Zhang, J., Benoit, M., Kimmoun, O., Chabchoub, A., Hsu, H., 2019. Statistics of extreme waves in coastal waters: Large scale experiments and advanced numerical simulations. Fluids 4, 99. doi:10.3390/fluids4020099.

Zhang, J., Benoit, M., Ma, Y., 2022. Equilibration process of out-of-equilibrium sea-states induced by strong depth variation: Evolution of coastal wave spectrum and representative parameters. Coastal Eng. 174, 104099. doi:10.1016/j.coastaleng.2022.104099.

Zhang, J., Ma, Y., Tan, T., Dong, G., Benoit, M., 2023. Enhanced extreme wave statistics of irregular waves due to accelerating following current over a submerged bar. J. Fluid Mech. 954, A50. doi:10.1017/jfm.2022.1022.

Zheng, Y., Lin, Z., Li, Y., Adcock, T.A.A., Li, Y., van den Bremer, T.S., 2020. Fully nonlinear simulations of unidirectional extreme waves provoked by strong depth transitions: The effect of slope. Phys. Rev. Fluids 5, 064804. doi:10.1103/physrevfluids.5.064804.