



ELSEVIER

Applied Ocean Research 25 (2003) 137–155

Applied Ocean  
Research

[www.elsevier.com/locate/apor](http://www.elsevier.com/locate/apor)

# Horizontal asymmetry and steepness distributions for wind-driven ocean waves from severe storms

Paul Stansell<sup>a,\*</sup>, Julian Wolfram<sup>b</sup>, Stan Zachary<sup>c</sup>

<sup>a</sup>*School of Physics, The University of Edinburgh, James Clerk Maxwell Building, The King's Buildings, Mayfield Road, Edinburgh, UK*

<sup>b</sup>*School of the Built Environment, Heriot-Watt University, Edinburgh EH14 4AS, UK*

<sup>c</sup>*Department of Actuarial Mathematics and Statistics, School of Mathematical and Computer Sciences, Heriot-Watt University, Edinburgh EH14 4AS, UK*

Received 28 October 2002; accepted 7 October 2003

## Abstract

Results on horizontal asymmetry and steepness distributions from analyses of ocean wave data collected during 10 severe storms in the northern North Sea are presented. The data have been collected at a sampling rate of 5 Hz using laser altimeters mounted on a fixed platform permitting the shapes of individual waves to be quite closely defined. This has allowed the steepness of the fronts and backs of wave crests and troughs to be examined. It is found that, on an average, as the non-dimensionalised wave height increases, the horizontal asymmetry becomes more pronounced. That is, the fronts of large wave crests tend to be steeper than their backs. Regression has been used to establish relationships between individual non-dimensionalised wave heights and steepness measures. The generalised Pareto distribution has then been used to establish a simple model for predicting the probability of extreme wave steepness conditional on the non-dimensionalised wave height.

© 2003 Elsevier Ltd. All rights reserved.

*Keywords:* Ocean wave steepness; Horizontal asymmetry; Steepness distributions conditional on wave height

## 1. Introduction

The shape of sea waves is of interest to those involved in designing and operating ships and offshore structures. Of particular interest is the steepness of large waves that can cause significant impact damage and give rise to so-called *greenwater* events, where several tonnes of water from the wave crest can sweep across the deck causing structural damage and possible flooding. The wave steepness of the incident waves causes important non-linear effects [3], and a number of recent incidents of damage in the Atlantic west of Shetland and in the North Sea have served to focus the attention of the offshore industry on the question of wave front steepness. In the deterministic models used for design purposes, waves are assumed to be symmetrical; and in random linear theory, wave front steepness is, on an average, equal to wave back steepness. The same is true for the random second-order theories that are now becoming adopted for the design and analysis of deep water structures.

Of course, when waves move into shallow water, their fronts become much steeper than their backs and they break. It might thus seem reasonable to assume that during deep water storms, where waves frequently break that, on an average, wave fronts would be steeper than wave backs. Yet Vinje and Haver [12] concluded from an analysis of wave data collected using laser altimeters mounted on the Gullfaks platform in the North sea that, on an average, the horizontal asymmetry is more or less zero. The same is concluded by Jonathon et al. [4] in their analysis of the North Sea Tern platform data in support of the New Wave theory developed by Tromans et al. [11].

In this paper, we describe the statistical analyses that have been undertaken using wave data collected using wave altimeters mounted on the North Alwyn platform in the northern North Sea. We examine 10 severe storms comprising a total of 217,083 individual waves of heights ranging up to a maximum of 24 m. We start by introducing several definitions in the context of time series measurements: six different measures of steepness for the various parts of the wave; and also two horizontal asymmetry coefficients, pertaining to crest and trough horizontal

\* Corresponding author. Fax: +44-131-451-5078.

E-mail address: [p.stansell@ed.ac.uk](mailto:p.stansell@ed.ac.uk) (P. Stansell).

asymmetry. After a detailed description of the data used for our analyses, and a discussion on how best to non-dimensionalise the data, we go on to look at horizontal asymmetry in waves from 10 storms and examine the relationship between asymmetry, non-dimensionalised wave height and wave steepness. We find that greater average horizontal asymmetry occurs in higher waves. We use a regression analysis to establish relationships between the steepness of various parts of the wave and non-dimensionalised wave height, and then we use a *generalised Pareto distribution* (GPD) as the basis for a model to predict the extremes of wave steepness given wave height.

**2. Some definitions**

The definitions introduced in this section are appropriate to the study of horizontal wave asymmetry in deep water ocean waves. An individual wave is defined as that part of the surface elevation profile which lies between any two consecutive zero down-crossings.

Let  $H_c$  be the vertical distance from the mean sea level to the crest maximum, and let  $H_t$  be the vertical distance from the mean sea level to the trough minimum. Both  $H_c$  and  $H_t$  are always positive and the wave height is given by their sum:

$$H = H_t + H_c. \tag{1}$$

Let  $\lambda$  and  $T$  be the wavelength and wave period, respectively. The four quarter wavelengths associated with each wave are defined as:  $\lambda_1$ , the horizontal distance between the position of the first zero down-crossing to the position of the trough minimum;  $\lambda_2$ , the horizontal distance between the position of the trough minimum to the position of the zero up-crossing;  $\lambda_3$ , the horizontal distance between the position of the zero up-crossing to the position of the crest maximum;  $\lambda_4$ , the horizontal distance between the position of the crest maximum to the position of the second zero down-crossing. The four quarter wave periods,  $T_i$ ,  $i = 1, 2, 3, 4$ , are defined analogously but in the time domain. For each wave  $\sum_{i=1}^4 \lambda_i = \lambda$  and  $\sum_{i=1}^4 T_i = T$ . These definitions are shown schematically in Fig. 1.

Traditionally wave steepness is defined by

$$s \equiv \frac{H}{\lambda}, \tag{2}$$

but, because the change in height  $H$  occurs over the distance  $\lambda/2$  and not  $\lambda$ ,  $s$  is in fact a measure representative of half of the spatial gradient likely to be encountered in a wave. To have a consistent set of wave steepness measures, each appropriate to different parts of a wave, it is convenient to define the following:

wave steepness,  $s_0 \equiv 2s$

trough front steepness,  $s_1 \equiv \frac{H_t}{\lambda_1}$ ,

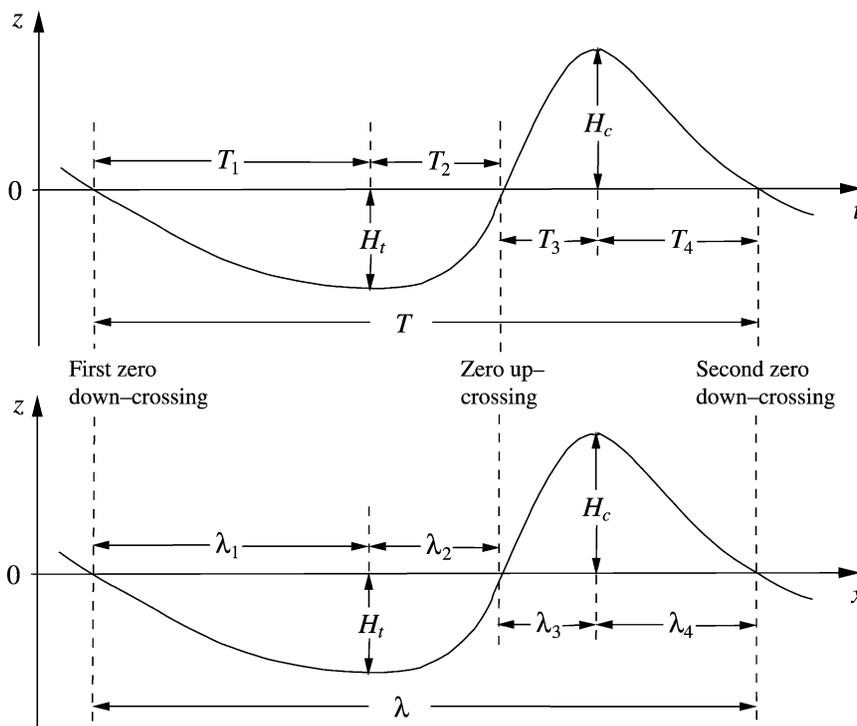


Fig. 1. Schematic illustration of definitions of crest height, trough depth and quarter wave lengths and periods. Note the wave is travelling in the negative  $x$ -direction.

trough back steepness,  $s_2 \equiv \frac{H_t}{\lambda_2}$ ,

crest front steepness,  $s_3 \equiv \frac{H_c}{\lambda_3}$ ,

crest back steepness,  $s_4 \equiv \frac{H_c}{\lambda_4}$ .

wave front steepness,  $s_{23} \equiv \frac{H}{\lambda_2 + \lambda_3}$ .

This set of six wave steepness measures is representative of the spatial gradients in various parts of a wave. The use of  $s_0 = 2s$  serves to relate the results presented here with other results concerning wave steepness which are traditionally given in terms of  $s$ .

Transforming to the time domain, we assume the linear dispersion relation,  $\omega^2 = gk$ , holds and that the shape of the wave does not vary significantly over a time scale equal to the period of the wave. The phase speed of a wave is, therefore, considered to be constant (at least over one wave period) and is equal to

$$c = \frac{\omega}{k} = \frac{gT}{2\pi}.$$

From this, and the relation  $\lambda = cT$ , the previously defined wave steepness measures are rewritten in terms of the time variable as

$$s = \frac{2\pi H}{gT^2}, \quad (3)$$

$$s_0 = \frac{4\pi H}{gT^2}. \quad (4)$$

Also, from the relations  $\lambda_i = cT_i$ , the crest and trough steepnesses are rewritten in terms of the time variable as

$$s_i = \frac{2\pi H_i}{gT_i^2}, \quad i = 1, 2, 3, 4, \quad (5)$$

where  $H_i = H_t$  for  $i = 1, 2$ , and  $H_i = H_c$  for  $i = 3, 4$ . Finally, the wave front steepness is rewritten as

$$s_{23} = \frac{2\pi H}{gT(T_2 + T_3)}. \quad (6)$$

Other definitions used in this study are those of the dimensionless quarter periods

$$\tau_i = \frac{T_i}{T}, \quad i = 1, 2, 3, 4,$$

the sum of which equals unity, and the dimensionless partial period

$$\tau_{23} = \frac{T_2 + T_3}{2T}.$$

Also defined are the dimensionless trough and crest horizontal asymmetry coefficients

$$A_t = (T_1 - T_2)/(T_1 + T_2), \quad (7)$$

$$A_c = (T_3 - T_4)/(T_3 + T_4). \quad (8)$$

This definition of  $A_c$  has an advantage over the horizontal asymmetry coefficient defined by Myrhaug and Kjeldsen [7, page 551] (that is to say,  $T_4/T_3$ ) in that for a statistically stationary Gaussian sea surface profile composed of a linear random sum of symmetric Airy waves, the distributions of  $A_t$  and  $A_c$  will be symmetric about its mean of zero.

### 3. The data

The raw data used in this study were collected from three Thorn EMI infra-red laser altimeters sampling at 5 Hz and mounted on three of the corners of the North Alwyn fixed steel-jacket oil and gas platform. The Alwyn North field, operated by TotalFinaElf, is situated in the northern North Sea about 100 miles east of the Shetland Islands (60°48.5' North and 1°44.17' East) in a water depth of approximately 130 m. There are two jacket platforms in close proximity connected by a walkway. The field processing platform, NAA, is the site of all the sensor and data logging equipment. The logging system is configured such that each sensor takes five measurement of the sea surface elevation every second. These are recorded for a duration of 20 min and the significant wave height,  $H_{m_0}$ , for this period is calculated as four times the square root of the variance of the measurements. If  $H_{m_0}$  is greater than 3 m, all three 20-min sea surface records are saved to optical disc for detailed analysis. We define a storm as the period between the start of the first, and the end of the last, of a continues sequence of 20 min records each satisfying  $H_{m_0} > 3$  m.

In this study, we analyse data collected over the full durations of 10 separate severe storm periods. The 10 storms are of varying bandwidth, but all are essentially uni-modal wind-driven seas without significant swell. To ensure the cleanest data, for each storm, we only use data from the altimeter which is upwind of the platform.

The raw data were stored as 1,879 20-min records of surface elevation measurements. Note that all wave records are wholly unfiltered: not being smoothed by any means other than that arising from the finite sampling rate, 5 Hz, of the measurement instruments. This rate of sampling is sufficiently high to yield an accurate representation of the sea surface (see Ref. [9] for a discussion of the effect of sampling rate on the measured distribution of wave heights). In a preliminary analysis of these individual 20-min records, the mean surface elevation was subtracted from each elevation measurement to give a wave record, denoted by  $\eta(t)$ , having a mean elevation of zero. In each wave record  $\eta(t)$ , the times of all zero-crossings were estimated by a linear interpolation from its positive and negative bracketing

Table 1  
Summary statistics for the data used in this study

Storm ID	No. of 20-min records	Total No. of waves	No. of waves, $H_c^* > 0.1$ and $H_t^* > 0.1$	Mean( $H$ ) (m)	Maxm( $H$ ) (m)	Mean( $T_z$ ) (s)
23	177	21,303	16,927	4.03	21.94	8.4
25	111	13,640	10,658	3.45	15.88	7.73
26	161	20,007	15,105	2.25	9.111	6.56
28	150	18,796	15,701	3.47	19.51	7.89
29	93	11,241	9251	3.51	20.27	7.87
90	310	39,145	33,069	3.65	23.85	8.0
124	173	17,910	15,928	4.9	21.14	9.64
146	91	10,859	7753	2.29	9.151	6.8
149	412	40,529	33,854	4.19	24.19	8.8
195	201	23,653	18,557	3.5	18.72	8.01
Combined	1879	217,083	176,803	–	24.19	–

points by

$$t(\eta = 0) = t_i - \frac{\eta_i \delta t}{\eta_{i+1} - \eta_i},$$

where  $t_i$  is the time of the  $i$ th measurement, and  $\delta t = t_{i+1} - t_i$  is the sampling period (equal to 0.2 s). From the set of zero-crossing times, all zero down-crossing waves were identified in these records, and the values of  $T_1$ ,  $T_2$ ,  $T_3$  and  $T_4$  for each were calculated. Also, the values of  $H_c$  and  $-H_t$  for each wave were taken, respectively, as the maximum and minimum values of  $\eta(t_i)$  lying between the waves' zero down-crossings. From these basic parameters, we calculated the six steepness measures and the two horizontal wave asymmetry coefficients as defined in Section 2.

A more detailed summary of the data from these storms is given in Table 1.

#### 4. Scaling and stationarity

The raw data is (in the statistical sense) highly non-stationary, the values for mean wave height and period differing widely over the full set of 20-min records (although the difference is only small between consecutive 20-min records). This makes it impossible to combine data from different records into a single data set. In an attempt to overcome this problem, we examined use of various non-dimensionalised variables. The result of non-dimensionalising the equations governing the behaviour of surface gravity waves (see, for example, Ref. [13, page 434]) shows that the single dimensionless parameter governing the behaviour of dynamically similar systems is  $\mathcal{L}/g\mathcal{T}^2$ , where  $\mathcal{T}$  and  $\mathcal{L}$  are representative measures of time and distance, respectively (see Appendix A). Some possible choices for  $\mathcal{L}$  are  $\sqrt{m_0}$  or  $H_s$ , and for  $\mathcal{T}$ ,  $2\pi m_0/m_1$  or  $2\pi\sqrt{m_0/m_2}$ , where  $m_n$  is the  $n$ th spectral moment of the wave record,

defined by

$$m_n = \int_0^\infty \omega^n S(\omega) d\omega, \quad n = 0, 1, 2, \dots,$$

and where  $S(\omega)$  is the one-sided frequency spectrum and  $\omega = 2\pi f$  is the angular frequency. One may have thought that choosing spectral measures for  $\mathcal{L}$  and  $\mathcal{T}$  would have the advantage that they would be directly available from the frequency spectrum. But in what follows, we analyse measurements from the time series directly and not from the energy spectrum. Also, one must be aware that the relations which relate the spectral measures to the time series measure<sup>1</sup> apply strictly only to Gaussian processes, and many, only to narrow-banded Gaussian processes. Our wave data is non-Gaussian and broad-banded. Thus, in this study, we choose to use the fundamental measures

$$\mathcal{L} = \bar{H},$$

$$\mathcal{T} = T_z,$$

where  $\bar{H}$  is the mean wave height and  $T_z$  is the mean zero-crossing period.

Generally, in the following sections, all variables which are non-dimensionalised by  $\bar{H}$  or  $T_z$ , or a combination of these, are denoted with the same symbol as used for the original variable, but differentiated by the addition of a superscripted asterisk (\*). Thus, for each wave record, all individual wave height and period measures are non-dimensionalised by dividing, respectively, by  $\mathcal{L} = \bar{H}$  and  $\mathcal{T} = T_z$ , where  $\bar{H}$  and  $T_z$  are the means measured from the

<sup>1</sup> Examples of relations of time series parameters to spectral parameters are as follows. For a Gaussian process,  $\eta(t) : \text{std}(\eta) \equiv \sqrt{m_0}$ , where  $\text{std}(\eta)$  is the standard deviation of the  $\eta$ ;  $T_z \equiv 2\pi\sqrt{m_0/m_2}$ , where  $T_z$  is the mean zero-crossing wave period; and  $\bar{T} \equiv 2\pi m_0/m_1$ . If  $\eta(t)$  is a narrow band Gaussian process then:  $\bar{H} = \sqrt{2\pi m_0}$ , and  $H_{m_0} = H_{\text{sig}} = 4.004\sqrt{m_0}$ . Note that from our data of the measured values of  $2\pi m_0/m_1$  and  $2\pi\sqrt{m_0/m_2}$ , the former seems far closer to  $T_z$  than the latter, although the former are slightly above and the latter slightly below  $T_z$ .

record containing the individual wave. For example,  $H_i^* = H_i/\bar{H}$  and  $T_i^* = T_i/T_z$ .

As each measure of steepness is proportional to a wave height measure divided by the product of two wave period measures, all wave steepness measures,  $s_i$ , from the same record will also be scaled by an amount  $T_z^2/\bar{H}$  and denoted

$s_i^*$ . As we expect the distributions of the various wave steepness measures to be constant across records with the same value of  $\mathcal{L}/g\mathcal{T}^2$ , we group all wave records based on their values of  $\bar{H}/T_z^2$ .

We expect some statistical variation in  $\bar{H}/T_z^2$ , even between 20-min segments of a continuous stationary wave

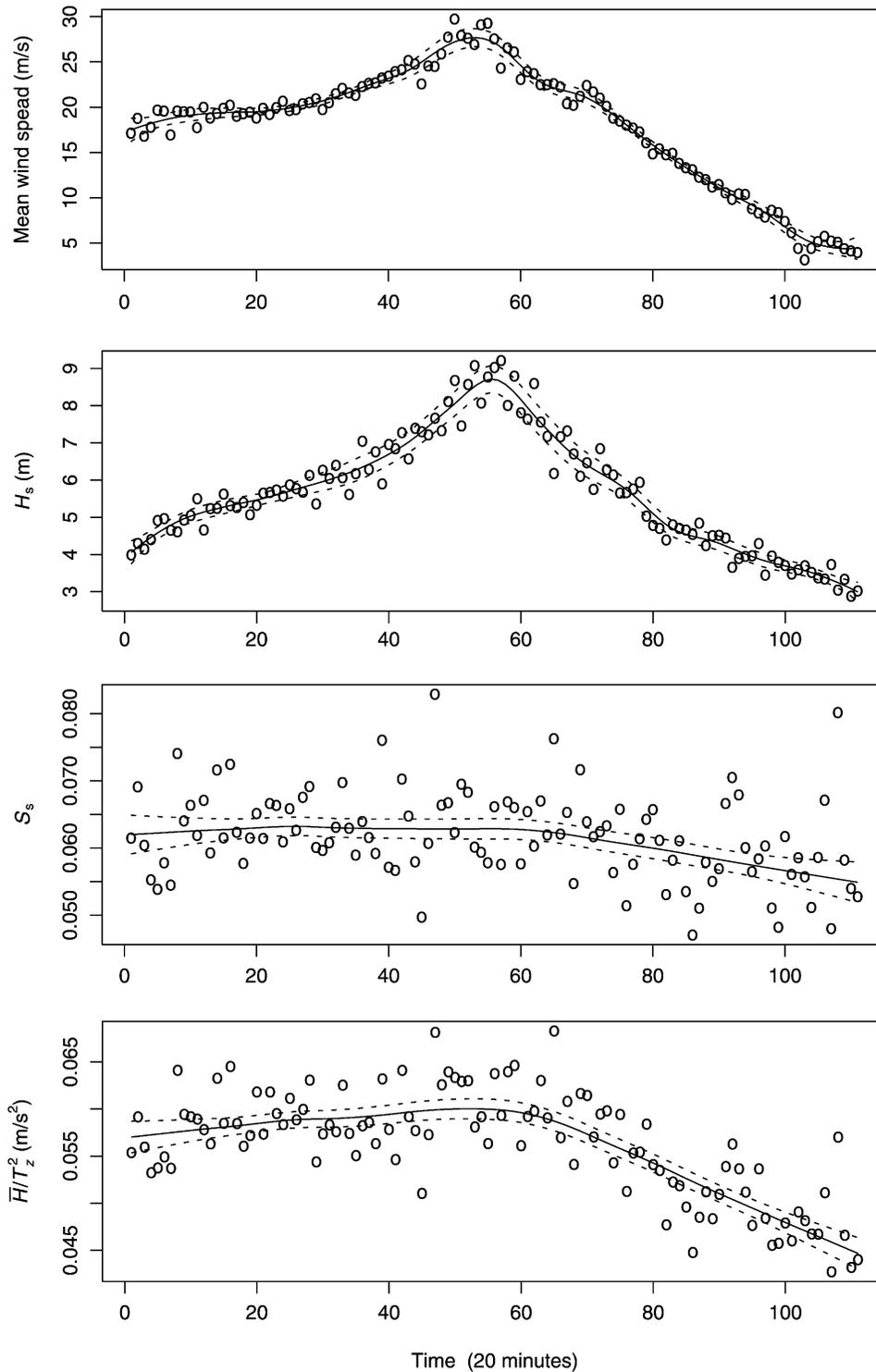


Fig. 2. Detail of Storm 25: points are averages over each 20-min record, the solid line is a local regression, and the dashed lines represent its 95% confidence band. Plots are shown for: the mean wind speed; the significant wave height,  $H_s$ ; the significant wave steepness,  $S_s$ ; and the scaling factor  $\bar{H}/T_z^2$ .

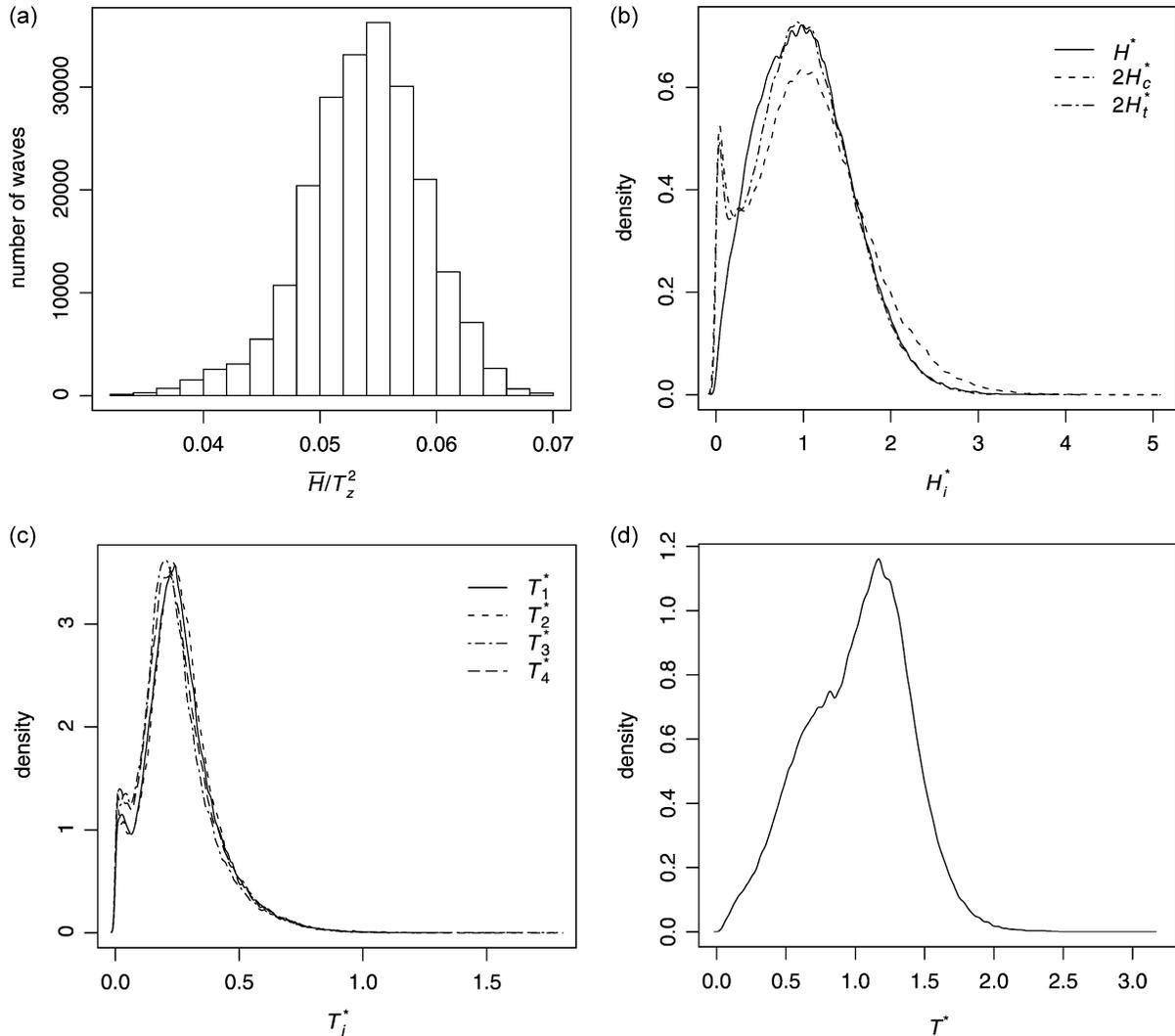


Fig. 3. Histogram of the number of individual waves from records with scaling values of  $\bar{H}/T_z^2$ . (a) Histogram of number of waves against scaling factors. (b) Distributions of dimensionless wave height measures from Bin 2. (c) Distribution of dimensionless quarter periods from Bin 2. (d) Distributions of dimensionless period from Bin 2.

record. The level of this variation is shown in Fig. 2 which shows, for Storm 25, a local regression<sup>2</sup> along with the average over each 20-min record of the following variables: the mean wind speed; the significant wave height,  $H_s$  (defined as the average height of the highest third of the waves); the significant wave steepness  $S_s$  (defined by  $S_s = 2\pi H_s/gT_z^2$ ); and the scaling factor  $\bar{H}/T_z^2$ . Assuming our data to be approximately stationary over a 20-min period, we use the variation in  $\bar{H}/T_z^2$  between consecutive 20-min records as a measure of the size of the natural statistical variation in  $\bar{H}/T_z^2$ . Analysis of the data has shown that about 95% of the differences lie between  $-0.01$  and  $0.01$ . Thus, a bin size for  $\bar{H}/T_z^2$  of  $0.01$  or less would be adequate.

Another way to calculate the expected variance in the measurements of  $\bar{H}/T_z^2$  from each record is to estimate its

standard deviation based on the standard deviations of  $\bar{H}$  and  $T_z$ .<sup>3</sup> For Storm 25, the mean value of  $\text{std}(\bar{H}/T_z^2) = 0.007$ , which gives a bin width for  $\bar{H}/T_z^2$  of approximately  $0.014$ . This value is in broad agreement with that previously estimated from the variance of the differences of  $\bar{H}/T_z^2$  between consecutive records.

Shown in Fig. 3(a) is a histogram of the number of individual waves plotted against the value  $\bar{H}/T_z^2$  of the record from which the waves were measured (the minimum and maximum values are  $0.033$  and  $0.068$ ). By consideration of this histogram, and in order to obtain bins with large populations of reasonably stationary wave

<sup>3</sup> That is to say

$$\text{std}\left(\frac{\bar{H}}{T_z^2}\right) \approx \left|\frac{1}{T_z^2}\right| \text{std}(\bar{H}) + \left|\frac{2\bar{H}}{T_z^3}\right| \text{std}(T_z),$$

where  $\text{std}(\bar{H}) = \text{std}(H)/\sqrt{N}$ ,  $\text{std}(T_z) = \text{std}(T)/\sqrt{N}$  and  $N$  is the number of waves in the record.

<sup>2</sup> See Section 5 for a description of the local regression algorithm used.

Table 2  
Summary of binned data used in this study

Bin	Number of records	Number of waves	Min( $\bar{H}/T_z^2$ )	Mean( $\bar{H}/T_z^2$ )	Max( $\bar{H}/T_z^2$ )
1	386	44,547	0.04402	0.04838	0.05066
2	966	111,099	0.05067	0.05403	0.05732
3	421	49,635	0.05734	0.05975	0.06396

data, we chose to ignore data from records with  $\bar{H}/T_z^2 < 0.044$  and  $0.064 < \bar{H}/T_z^2$ . We grouped the remaining data into three bins of width 2/300. These bin widths are within the 0.01 width suggested by the analysis of differences above. Table 2 summarises aspects of the data after partitioning. Pairwise quantile–quantile plots (not shown) between various sections of the data from each bin confirm that the distributions of the dimensionless parameters  $H^*$ ,  $H_i^*$ ,  $T^*$ ,  $T_i^*$  and  $S_i^*$  are, to an adequate level of approximation, unchanging between records within each bin.

Other plots in Fig. 3 show the densities of the dimensionless wave height and wave period measures for data from Bin 2 only (which are representative of the other two bins too). They show a high density of very small waves. This feature has been documented before (see, for example, Refs. [8,10]) and results from measurements of data; that is broad-banded. Such data is not well fitted by a Rayleigh distribution. In the analysis that follows, we ignore the effects of these small, waves by considering only those waves for which either  $H_c^* < 0.1$  or  $H_t^* < 0.1$  (see Table 1).

Since we non-dimensionalise our data using  $\bar{H}$  and  $T_z$ , and then examine the horizontal asymmetry of

the non-dimensionalised data, it is important that there is no dependency of the horizontal asymmetry coefficients  $A_t$  and  $A_c$  on either  $\bar{H}$  or  $T_z$ . This is demonstrated in Fig. 4 which shows, via local regressions of  $A_t$  and  $A_c$  on  $\bar{H}$  and  $T_z$  that there is no significant dependency of either  $A_t$  and  $A_c$  on either  $\bar{H}$  or  $T_z$ . Although there does appear to be some slight increase in both  $A_t$  and  $A_c$  with increasing  $\bar{H}$  greater than about 6 and increasing  $T_z$  greater than about 10, this is not significant at the 95% level since it is possible to draw a horizontal line within each of the 95% confidence intervals.

### 5. The distribution of wave quarter periods conditional on wave height

In this section, we investigate horizontal asymmetry by an analysis of the distributions of the dimensionless quarter periods  $\tau_i$ ,  $i = 1, 2, 3, 4$ , and the asymmetry coefficients  $A_c$  and  $A_t$ , in each case conditional on wave height.

We first verify that the dimensionless partial periods  $\tau_i$  may reasonably be treated as independent observations. Any failure of independence is almost certain to show in the corresponding autocorrelation functions: in particular in the lag-one autocorrelations. For each of the variables  $\tau_i$ ,  $i = 1, 2, 3, 4$ ,  $A_c$  and  $A_t$ , and for each storm, the above autocorrelations were all close to zero. For example, for Storm 23, the lag-one autocorrelations are:  $R_1(\tau_1) = -0.102$ ,  $R_1(\tau_2) = -0.134$ ,  $R_1(\tau_3) = -0.134$ ,  $R_1(\tau_4) = -0.0884$ ,  $R_1(A_c) = -0.0091$  and  $R_1(A_t) = -0.0108$ . These compare to lag-one autocorrelations for wave height and period of  $R_1(H^*) = 0.317$  and  $R_1(T^*) = 0.245$ , which indicates that observations of these variables are not independent—see Ref. [5] for a Markov chain model of these.

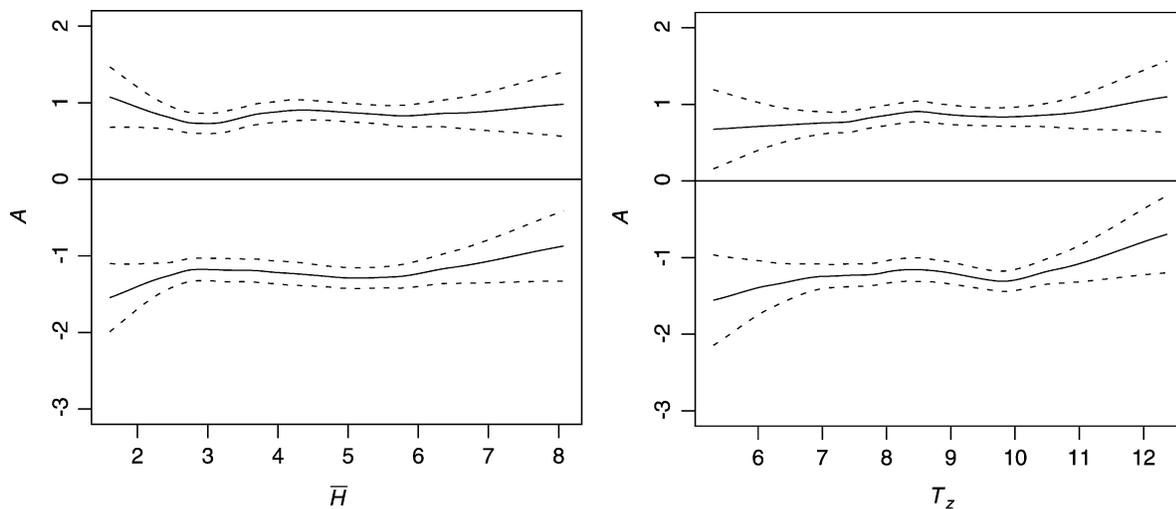


Fig. 4. Local regressions (with 95% confidence intervals) of dimensionless trough (upper line) and crest (lower line) horizontal asymmetry coefficients against unscaled significant mean wave height and mean zero-crossing wave period (the fitting parameters used were  $\beta = 0.5$  and  $\Delta = 1$ ).

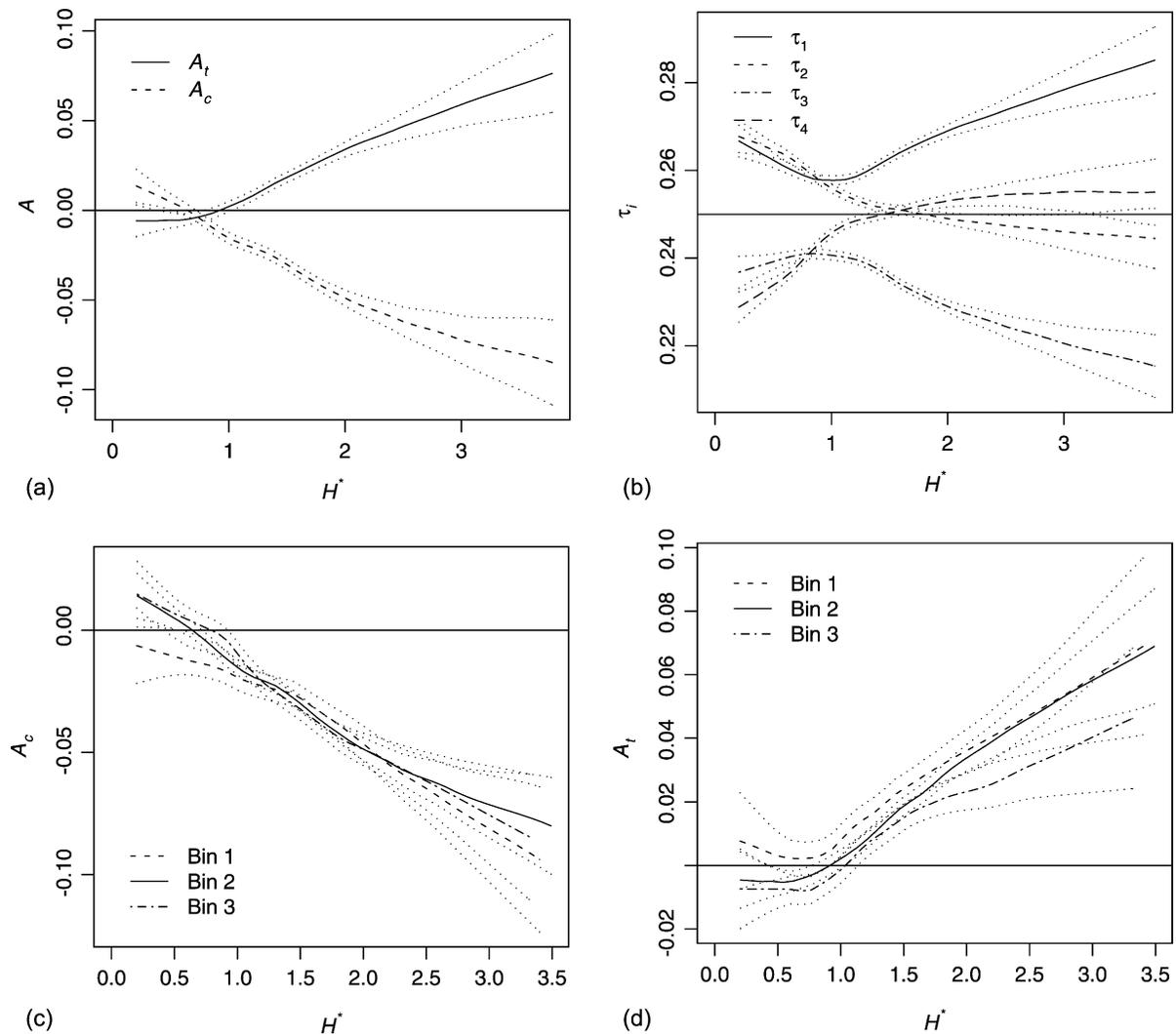


Fig. 5. Local regressions (with 95% confidence intervals) of dimensionless wave quarter period measures and the horizontal asymmetry coefficients against dimensionless wave height for the three steepness bins (the fitting parameters used were  $\beta = 0.5$  and  $\Delta = 1$ ). (a) Asymmetry parameters for data from Bin 2. (b) Dimensionless quarter periods for data from Bin 2. (c) Crest asymmetry parameter for data from all bins. (d) Trough asymmetry parameter for data from all bins.

Throughout this study, we make use of a non-parametric local regression procedure called *locfit*<sup>4</sup> [6]. For each value of a predictor variable,  $x$ , *locfit* estimates the response variable,  $y$ , as  $y = f(x) + \epsilon$ , where  $f(x)$  is a non-parametric function obtained by a local regression for those observations in the neighbourhood of  $x$  and  $\epsilon$  is a residual random variable. The *locfit* procedure has two tuneable parameters called the *bandwidth* the *degree* fitting parameter, denoted  $\beta$  and  $\Delta$ , respectively. One advantage of using a non-parametric regression is that it is not required to specify, a priori, the functional form of  $f(x)$ . In particular, the *locfit* algorithm is very flexible, making it ideal for modelling complex processes for which no theoretical models exist. The disadvantage is

that the estimated functions  $f(x)$  have no simple mathematical description.

Fig. 5 shows local regressions of the dimensionless quarter period measures and the horizontal asymmetry coefficients against dimensionless wave height. It is evident from Fig. 5(a) that, for the data from Bin 2, there is a strong dependency of the horizontal asymmetry coefficients on dimensionless wave height. Both  $A_c$  and  $A_t$  show a definite and pronounced increase with  $H^*$ . Regression of  $A_c$  and  $A_t$  for all bins are shown together in Fig. 5(c) and (d). It is clear that the data from each bin shows the same dependency, only the size of the confidence intervals vary, owing to the smaller size of the data sample in Bins 1 and 3. Also, it is suggested by Fig. 5(c) and (d) that there is no significant difference in  $A_c$  or  $A_t$  across bins. This is confirmed by quantile–quantile plots (not shown).

<sup>4</sup> *Locfit* is implemented in the *R* and *S-Plus* statistical analysis packages.

Fig. 5(b) shows local regressions of each dimensionless quarter periods,  $\tau_i$ , against  $H^*$  for the data from Bin 3. There is very clear evidence of horizontal asymmetry, and this asymmetry is seen to increase with  $H^*$ . Generally, and on an average, for small waves (with  $H^* < 1$ ), it appears that  $\tau_1 \approx \tau_2$  and  $\tau_3 \approx \tau_4$ , whereas for large waves (with  $H^* > 2$ ), it appears that  $\tau_3 < \tau_2 < \tau_4 < \tau_1$ . These results are consistent with small waves being symmetric in the horizontal direction, but asymmetric in the vertical direction (like a Stokes wave). Large waves, however, are clearly asymmetric in both the vertical and horizontal directions (with similar ratios of quarter periods as pictured schematically in Fig. 1.

### 6. The distribution of wave steepness conditional on wave height

We now investigate the distribution of wave steepness conditional on wave height. The traditional definition of wave steepness given by Eq. (2) represents an average steepness over a whole zero-crossing wave. For a Gaussian probability model of ocean waves, the statistics of the various wave steepness measures, defined by Eqs. (4)–(6), will be identical. However, in Section 5, and particularly Fig. 5(b), it was clearly demonstrated that there is a significant horizontal asymmetry present in waves for which  $H^* > 1.5$ . We, therefore, treat separately each of the steepness measures  $s_i^*$ ,  $i = 1, 2, 3, 4, 0, 23$ , and we consider models of the form

$$s_i^* = f_i(H^*) + \epsilon_i, \tag{9}$$

describing the distribution of  $s_i^*$  conditional on wave height  $H^*$ . Here,  $f_i(H^*)$  is a location measure for this conditional distribution and  $\epsilon_i$  is a residual random variable. In Section 6.1, we consider estimation of the functions  $f_i$  using both non-parametric and parametric methods. In Section 6.2, we show that the distributions of the residual random variables  $\epsilon_i$  have no significant further dependence on  $H^*$ . We use extreme value theory to model the tails of these distributions. Thus, from knowledge of the functions  $f_i$  and of the tails of the distributions of  $\epsilon_i$ , we may predict the probabilities of extreme levels of steepness conditional on  $H^*$ .

#### 6.1. Local regression of steepnesses on wave height

In this section, we consider estimation of the function  $f_i$  defined in Eq. (9). We restrict attention to waves for which both  $H_c^* \geq 0.1$  and  $H_t^* \geq 0.1$ .

We consider first the locfit non-parametric estimation. In order to estimate confidence limits, it is important to verify that the residuals can be treated as observations of independent random variables. Analysis of the autocorrelation function of each  $\epsilon_i$  strongly suggests that this is the case. For example, the lag-one autocorrelations are:  $R_1(\epsilon_1) = 0.0379$ ,  $R_1(\epsilon_2) = 0.0603$ ,  $R_1(\epsilon_3) = 0.0598$ ,

$R_1(\epsilon_4) = 0.0588$ ,  $R_1(\epsilon_0) = 0.122$  and  $R_1(\epsilon_{23}) = 0.0935$ . These values are not zero, but are considered sufficiently small to make the assumption of independence reasonable. Further, autocorrelation plots (not shown) give no evidence of higher-order autocorrelation.

Fig. 6 shows the estimated functions  $f_i$  together with their 95% confidence envelopes for the data from each bin fitted separately. There is an interesting structure to these fits of  $s_i$  and all show the same trends: the differences in  $f_i$  between the bins vanish at around  $H^* = 1$ ; for  $H^* > 2$  the magnitudes of  $f_i$  are based on bin order, with the values from Bin 3 being the largest and those from Bin 1 being the smallest; for  $H^* < 1$ , the converse is true, that is to say, the magnitudes of  $f_i$  from Bin 1 are largest and those from Bin 3 are the smallest.

The fact that  $f_i$  for large waves is higher for bins of higher  $\bar{H}/T_z$  is not surprising since  $\bar{H}/T_z$  is a measure of average steepness of the waves in the bin. But, the observation that the difference vanishes and  $H^* = 1$ , and then reverses for  $H^* < 1$  is surprising (this observation has also been confirmed by using more flexible and responsive values for the locfit bandwidth and degree parameters, specifically  $\beta = 0.25$  and  $\Delta = 2$ , and the same trends were observed).

Fig. 7 shows plots similar to those in Fig. 6 except in this case only data from Bin 2 is shown and the regressions of each  $s_i^*$  against  $H^*$  and  $H_i^*$  are shown together on single plots. From Fig. 7(a), it is clear that for large waves, the order of the steepness measures within Bin 2 is  $s_3^* > s_2^* > s_{23}^* > s_0^* > s_2^* > s_1^*$ . This result is consistent with the asymmetry observed in the quarter periods. All regressions displayed to this point have used  $H^*$ , and not  $H_i^*$ , as the predictor variable. This means that we have eliminated any effects of vertical asymmetry in the wave. Fig. 7(b) shows the local regression of  $s_i^*$  on  $H_i^*$ , and it can be seen that the vertical asymmetry has a greater effect than the horizontal asymmetry on the regressions of  $s_i^*$ .

Since parametric estimates are often more useful for prediction, we also estimate each  $f_i$  using robust quadratic regression obtained via iteratively re-weighted least squares. That is, for each  $i$ , we also obtain a robust fit of the model

$$s_i^* = \alpha_i + \beta_i H^* + \gamma_i H^{*2} + \epsilon_i, \tag{10}$$

where  $\alpha_i$ ,  $\beta_i$  and  $\gamma_i$  are estimated coefficients, and  $\epsilon_i$  is again a residual random variable.

For each bin, a model of the form of Eq. (10) was fitted separately and for all waves for which  $H^* > 1.25$ . The results are displayed in Fig. 8 which shows the values of  $\alpha_i$ ,  $\beta_i$  and  $\gamma_i$ , and bars representing their standard errors, for each bin and steepness measure.

Fig. 9 shows the quadratic fits obtained from the data in Bin 2 superimposed over the corresponding locfit for each of the steepness measures. In each case, the similarity between the locfit and the quadratic fits is very high: the latter lying

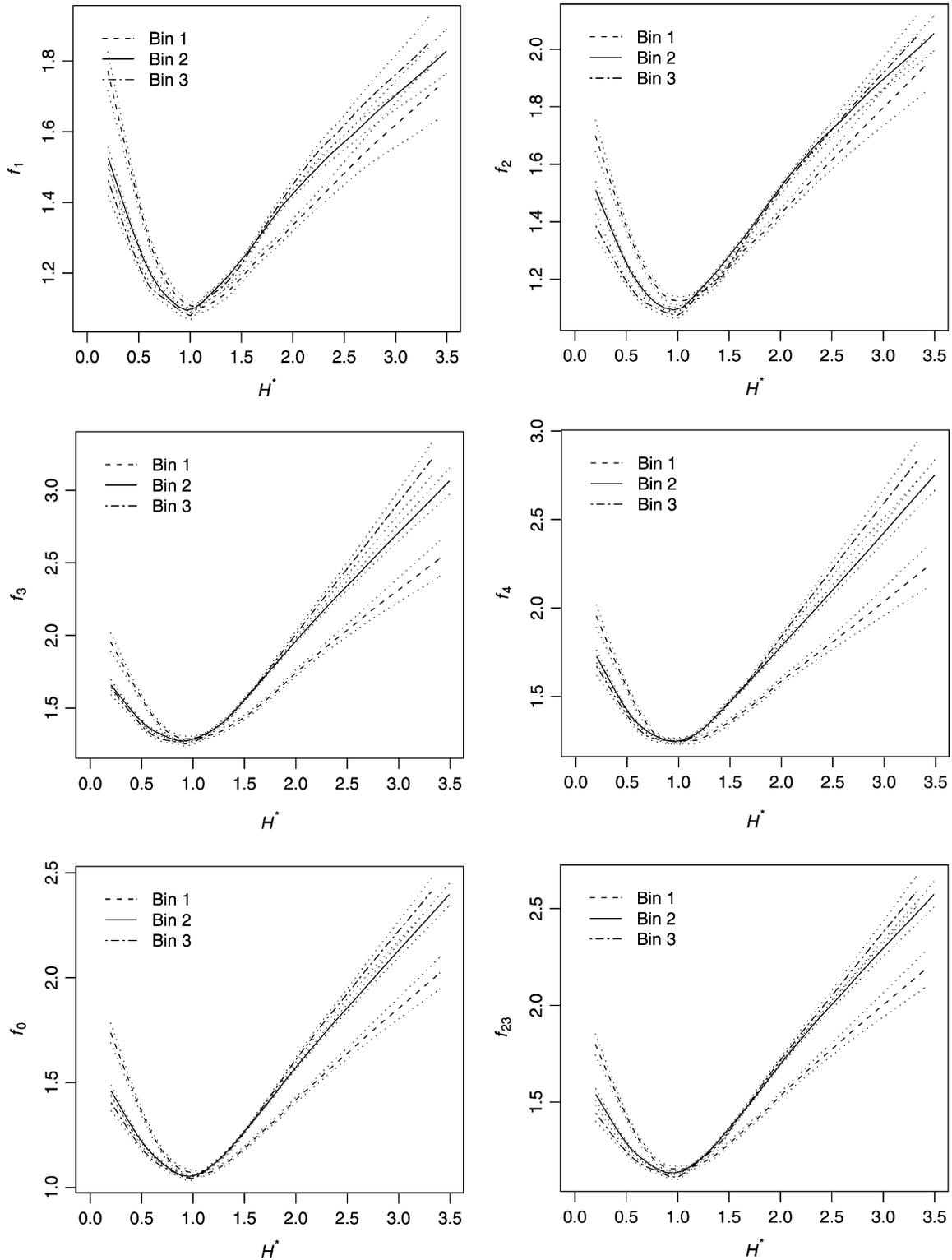


Fig. 6. Local non-parametric regressions (with 95% confidence intervals) of the various steepness measures,  $s_i^*$ , against  $H^*$  (the fitting parameters used were  $\beta = 0.25$  and  $\Delta = 1$ ).

within the 95% confidence envelope of the former over the range  $H^* > 1.25$ . This suggests that the simple relation given by Eq. (10) does in fact appear adequate to model the process over this range of  $H^*$ . The quadratic regressions

provide formulae with which to develop mathematical models for the variation of steepness with wave height. Table 3 gives the fitted parameters for the robust quadratic regression for each of the steepness measures in each bin.

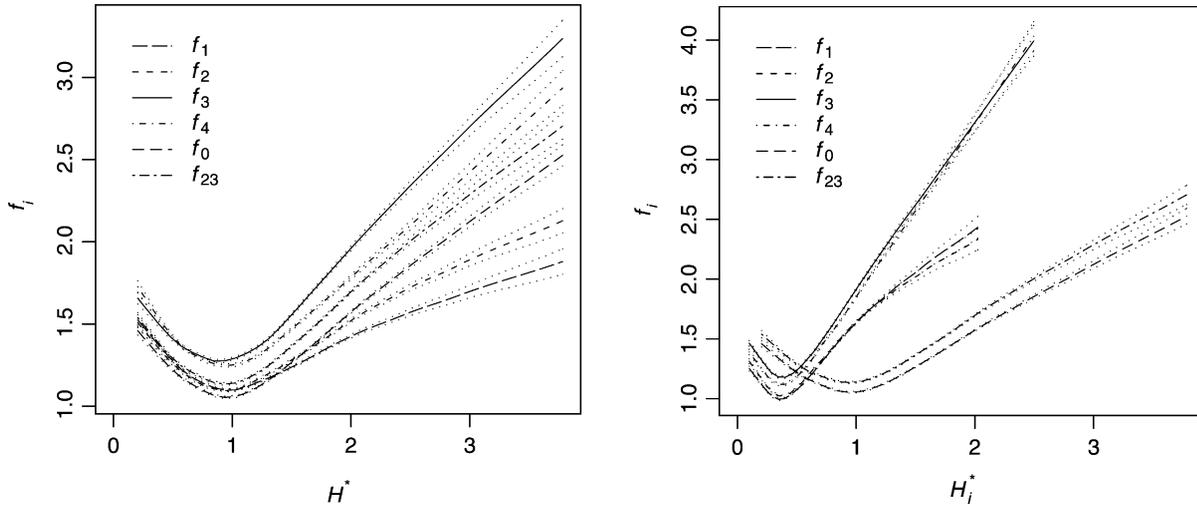


Fig. 7. Local non-parametric regressions (with 95% confidence intervals) of the various steepness measures,  $s_i^*$ , regressed against  $H^*$  and  $H_i^*$  for the data from Bin 2 (the fitting parameters used were  $\beta = 0.25$  and  $\Delta = 1$ ). (a)  $s_i^*$  regressed against  $H^*$ . (b)  $s_i^*$  regressed against  $H_i^*$ .

6.2. The distributions of extremes of wave steepness

We now estimate the distribution of the extremes of wave steepness conditional on wave height. For each

steepness measure,  $s_i^*$ , the distribution of the residuals  $\epsilon_i$  from the quadratic regression fitted by Eq. (10) appears, to a good approximation, to be independent of  $H^*$  in the range  $H^* > 1.25$ . This is verified graphically

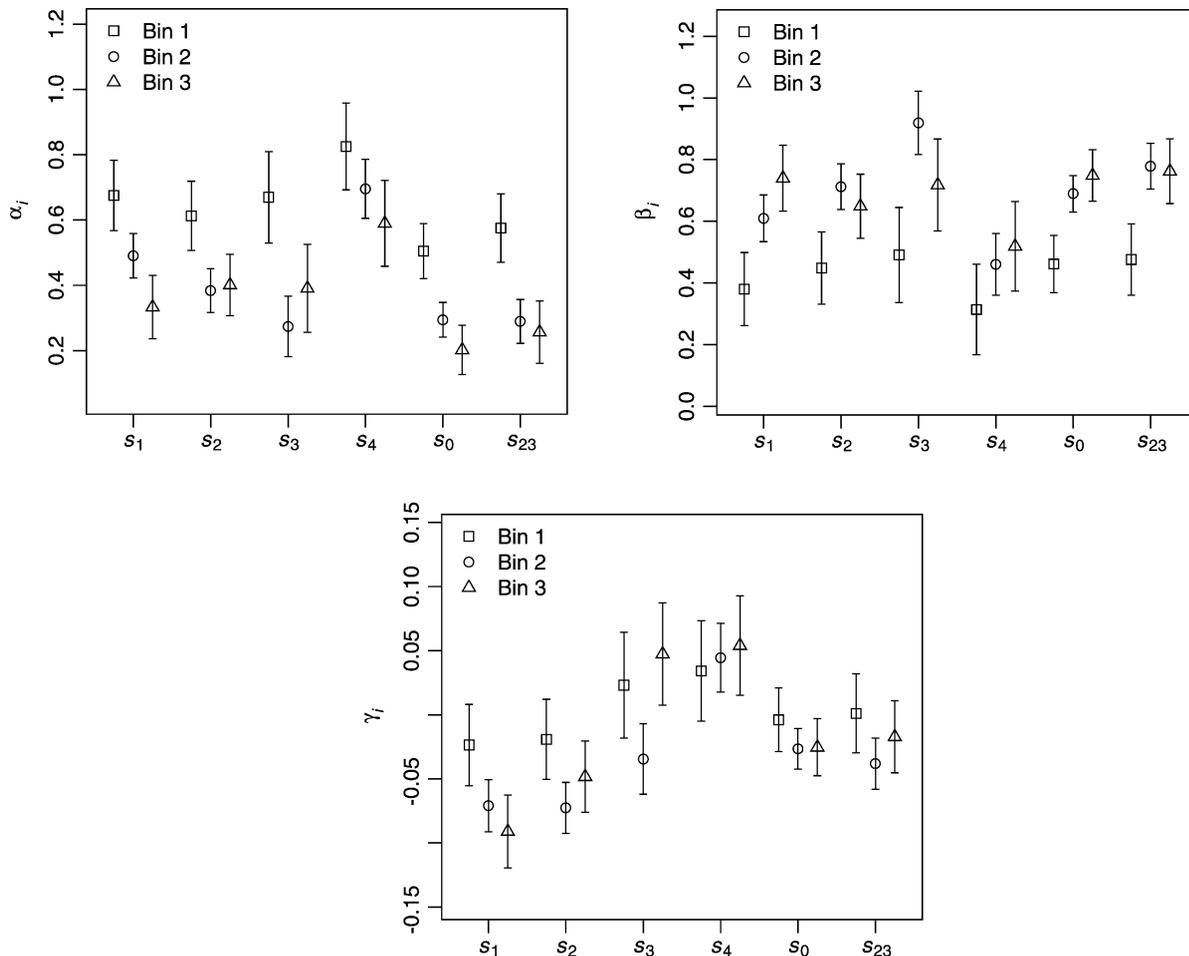


Fig. 8. Robust regression estimates of the fitted parameters  $\alpha_i$ ,  $\beta_i$  and  $\gamma_i$  for  $i = 0, 1, 2, 3, 4, 23$  for each bin. Error bars of  $\pm 1$  standard error are also shown.

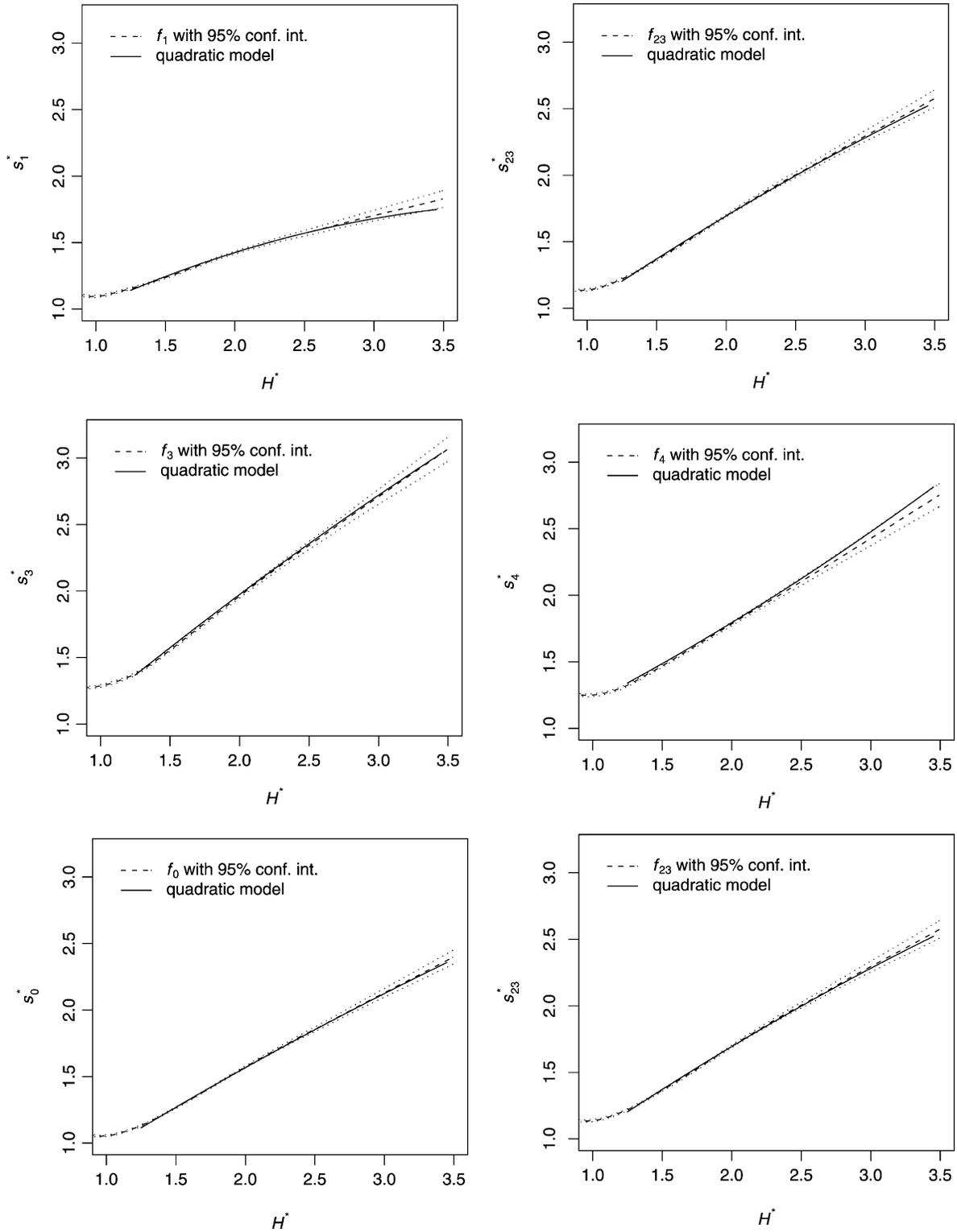


Fig. 9. Local non-parametric (with 95% confidence intervals) and quadratic regressions of the various steepness measures,  $s_i^*$ , against  $H^*$  (the fitting parameters used were  $\beta = 0.25$  and  $\Delta = 1$ ).

as follows. The data are sorted into 20 equally sized groups according to their values of  $H^*$ . For each  $i$  and for each of these groups, Fig. 10 shows a set of quantiles for the above residuals. These correspond to the eight probability levels,  $p = \{0.1, 0.3, 0.5, 0.7, 0.8,$

$0.9, 0.95, 0.99\}$ . For each  $i$  and for each probability level, the quantiles are connected across groups to obtain a series of probability contours; for  $H^* > 1.25$  and  $\epsilon_i > 0$ , these quantiles indeed appear to be independent of  $H^*$ .

Table 3  
Values of the fitted parameters for the robust quadratic regression of steepness on wave height.

Bin index	Steepness index $i$	$\alpha_i$	$\beta_i$	$\gamma_i$
1	1	0.6752	0.3804	-0.02361
	2	0.6128	0.4484	-0.01919
	3	0.6699	0.4904	0.02308
	4	0.8253	0.314	0.03417
	0	0.5047	0.4615	-0.003933
2	23	0.5755	0.476	0.001079
	1	0.4906	0.6095	-0.07095
	2	0.3842	0.7122	-0.07264
	3	0.2737	0.9193	-0.03449
	4	0.6954	0.4605	0.0445
3	0	0.2943	0.6895	-0.02652
	23	0.2894	0.7788	-0.03813
	1	0.3333	0.7396	-0.09106
	2	0.4011	0.6489	-0.04831
	3	0.3907	0.3907	0.04739
	4	0.5897	0.5189	0.054
	0	0.2023	0.7486	-0.02533
	23	0.2567	0.7622	-0.0172

In order to estimate extremes, for each bin and each  $i$ , we seek a model for the tail of the distribution of the residuals  $\epsilon_i$ , valid for all  $H^* > 1.25$ . Asymptotic theory [1,2] suggests that it is appropriate to consider modelling this tail by a generalised Pareto Distribution (GPD). This theory is applicable to all distributions, subject only to mild regularity conditions, even when these distributions arise as mixtures of those corresponding to different physical conditions. For the purposes of extrapolation to very extreme values, the modelling of the tail of a distribution by a GPD may only work well when there are a sufficiently large number of observations that may themselves be considered reasonably extreme. However, our purpose here is essentially that of interpolation, and the family of GPD distributions is also a sufficiently large and flexible class for this purpose.

The GPD has the distribution function  $F_{\xi\mu\sigma}(x) = 1 - \bar{F}_{\xi\mu\sigma}(x)$  where the complementary distribution function is given by

$$\bar{F}_{\xi\mu\sigma}(x) = \begin{cases} \left[ 1 + \frac{\xi(x - \mu)}{\sigma} \right]^{-1/\xi}, & \text{if } \xi \neq 0, \\ \exp\left(-\frac{x - \mu}{\sigma}\right), & \text{if } \xi = 0, \end{cases} \quad (11)$$

and where

$$\begin{aligned} x &\geq \mu, & \text{if } \xi \geq 0, \\ \mu \leq x \leq \mu - \sigma/\xi, & \text{if } \xi < 0. \end{aligned}$$

Here  $\xi$ ,  $\mu$  and  $\sigma > 0$  are, respectively, shape, location and scale parameters for the GPD (note in particular that  $\bar{F}_{\xi\mu\sigma}(x)$  is continuous in  $\xi$  at  $\xi = 0$ ).

Hence, for each wave steepness measure  $i$ , we consider the model

$$\Pr(\epsilon_i > x) = \bar{F}_{\xi_i\mu_i\sigma_i}(x), \quad (12)$$

valid for all  $x$  greater than or equal to some appropriately chosen threshold  $u_i$ . While various graphical techniques, for example, *mean excess* or *mean residual* plots [1,2], can be used to assist in the choice of appropriate thresholds, the resulting plots can in practise be very difficult to interpret (see Ref. [2]), and we have decided instead simply to choose the lowest thresholds above which the fits are clearly satisfactory (see below).

Under the model (12) we have

$$p_i = \Pr(\epsilon_i > u_i) = \bar{F}_{\xi_i\mu_i\sigma_i}(u_i), \quad (13)$$

while elementary calculation shows that

$$\Pr(\epsilon_i > u_i + z | \epsilon_i > u_i) = \frac{\bar{F}_{\xi_i\mu_i\sigma_i}(u_i + z)}{\bar{F}_{\xi_i\mu_i\sigma_i}(u_i)}, \quad (14)$$

$$= \bar{F}_{\xi_i,0,\tilde{\sigma}_i}(z), \quad (15)$$

where

$$\tilde{\sigma}_i = \sigma_i + \xi_i(u_i - \mu_i). \quad (16)$$

The parameters  $p_i$ ,  $\xi_i$  and  $\tilde{\sigma}_i$  may be determined by maximum likelihood estimation where, for each observation of  $S_i^*$ , we record whether or not it exceeds the threshold  $u_i$  and, if so, the value of its excess (the number of observations exceeding  $u_i$  is sufficient for the estimation of  $p_i$  from Eq. (13); the values of the excesses are sufficient for the estimation of  $\xi_i$  and  $\tilde{\sigma}_i$ ). The original (threshold independent) parameters  $\mu_i$  and  $\sigma_i$  may then be recovered via the relations (11), (13) and (16). Thus, for all  $\xi_i$ ,

$$\begin{aligned} \sigma_i &= \tilde{\sigma}_i p_i^{\xi_i}, \\ \mu_i &= u_i + \frac{\tilde{\sigma}_i}{\xi_i} (p_i^{\xi_i} - 1). \end{aligned}$$

The values of the thresholds and the fitted parameters for each steepness measure are given in Table 4.<sup>5</sup>

For each steepness measure  $S_i^*$ , Fig. 11 shows a quantile–quantile plot comparing the empirical and fitted residual distributions. For this purpose, we have projected the fitted GPD below the threshold  $u_i$ . The linearity of the quantile–quantile plots above the chosen thresholds shows the choice of these to be satisfactory. Similarly, Fig. 12 shows kernel density estimates for each residual distribution together with the corresponding GPD densities.

These results now enable estimation of the probability of any given extreme of steepness  $S_i^*$  conditional on  $H^*$ , for any  $H^* > 1.25$ . Thus, for any  $S^*$

$$\Pr(S_i^* > S^* | H^*) = \bar{F}_{\xi_i\mu_i\sigma_i}(\epsilon_i), \quad (17)$$

provided  $\epsilon_i = S^* - \alpha_i - \beta_i H^* - \gamma_i H^{*2}$  is greater than the threshold  $u_i$ . Here,  $\alpha_i$ ,  $\beta_i$  and  $\gamma_i$  are the estimated parameters

<sup>5</sup> The *S-Plus* code of Ref. [1] was used to obtain these fits.

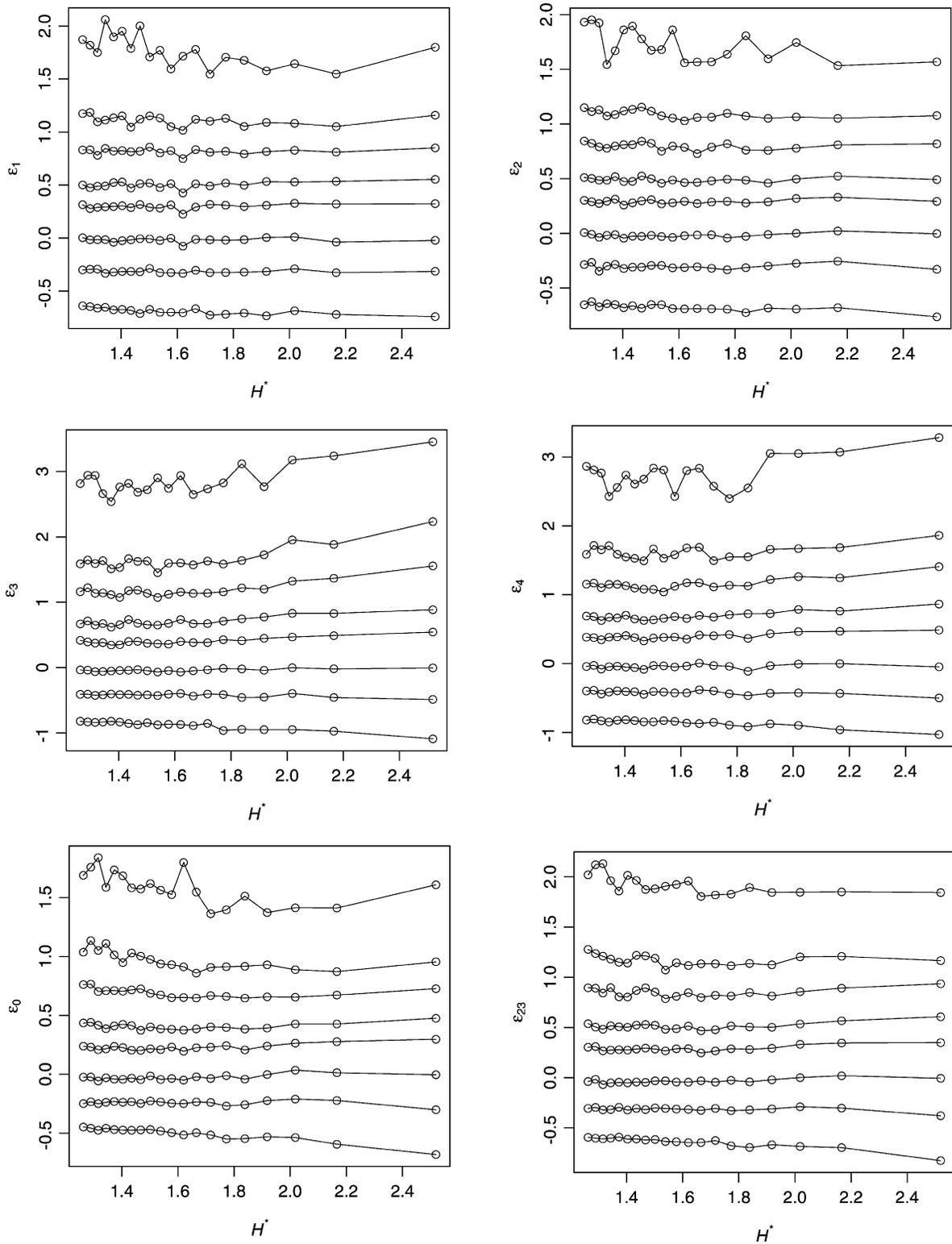


Fig. 10. Probability contours from data in Bin 2 showing, for each steepness measure  $i$ , the dependence on dimensionless wave height of the quantiles of the distribution of  $\epsilon_i$ . The quantiles of each probability level are joined to form the contours.

given in Table 3; the functional form of  $\bar{F}_{\xi_i, \mu_i, \sigma_i}(x)$  is given by Eq. (11); and estimates of the parameters  $\xi_i$ ,  $\mu_i$  and  $\sigma_i$  are given in Table 4. For  $S^*$  such that  $\epsilon_i < u_i$ —corresponding to less extreme values of steepness—the above probability

may be estimated directly by interpolation from Fig. 10. We are in general, however, interested in extremes such that  $\epsilon_i > u_i$ . The region of applicability of the model is shown schematically in Fig. 13.

Table 4  
Thresholds and fitted parameters for the GPD of the tails of wave steepness measures

Bin index	Steepness index, $i$	$u_i$	$\bar{\sigma}_i$	$\xi_i$	$p_i$	$\sigma_i$	$\mu_i$
1	1	1	0.3957	0.1087	0.07335	0.2979	0.1000
	2	1	0.4534	0.02106	0.07034	0.4287	-0.1704
	3	1	0.6550	0.1010	0.1325	0.534	-0.1976
	4	1	0.6712	0.02356	0.1213	0.6387	-0.3815
	0	1	0.4296	0.01408	0.05009	0.4119	-0.2595
	23	1	0.4921	0.007496	0.07727	0.4827	-0.2478
2	1	1	0.3827	0.1147	0.06522	0.2798	0.1030
	2	1	0.3862	0.03583	0.06215	0.3496	-0.02128
	3	1	0.6548	0.1047	0.1320	0.5297	-0.1947
	4	1	0.6382	0.07412	0.1277	0.5479	-0.2184
	0	1	0.3934	0.01489	0.0468	0.3759	-0.1775
	23	1	0.4523	0.01672	0.07333	0.4329	-0.1562
3	1	1	0.3293	0.1720	0.05587	0.2005	0.2512
	2	1	0.3524	0.08917	0.05465	0.2719	0.09773
	3	1	0.6906	0.08508	0.1352	0.5825	-0.2709
	4	1	0.6413	0.1209	0.1258	0.4991	-0.1761
	0	1	0.3725	0.03942	0.04128	0.3285	-0.1158
	23	1	0.4484	0.01569	0.07053	0.4301	-0.1646

Unconditional probabilities for extremes of steepness are given by

$$\Pr(s_i^* > S^*) = \int p(H^*) \Pr(s_i^* > H^* | H^*) dH^*,$$

where  $p(H^*)$  is some probability density function for wave height such as the Rayleigh density—see Fig. (13).

### 7. Discussion

The data used here were obtained in wind-driven sea conditions without significant swell. The significant wave-height increases from around 3 m at the start of each storm and rises to a maximum value of up to 13 m before falling again to around 3 m. The conditions are typical of those encountered in North Sea storms and the wave spectra have been found to be modelled well by the JONSWAP spectra. They are therefore representative of extra-tropical storms where the fetch is limited. Since the data includes that obtained during some of the largest storms seen in the northern North Sea in the last eight years, they are as close to extreme design conditions as is likely to be recorded consistent with having a data set of reasonable size and representation.

Strictly speaking the predictive model (17) is only applicable to similar storm conditions. Considering Fig. 2 we see that the steepness tends to be higher during the growth of the storm up to its peak than during its subsequent decay. By non-dimensionalising the data and splitting it into bins according to  $\bar{H}/T_z^2$ , we distinguish effectively between

these phases, and, from the ten storms, there are sufficient numbers of individual waves at all points along the dimensionless waveheight range for relationships to be established between waveheight and each measure of steepness. Thus, the models developed here can be applied to fetch-limited wind-driven sea conditions where  $\bar{H}/T_z^2$  varies from around 0.044 to around 0.064 (as specified in Table 2).

The model (17) is one for the conditional distribution of steepness measure  $s_i^*$  given dimensionless waveheight  $H^*$ . Equivalently, through Eq. (4), it is a model for the conditional distribution of dimensionless period  $T_i^*$  given  $H^*$ . Together with a knowledge of the unconditional distribution of  $H^*$  itself, it may be used to predict the unconditional distribution of  $s_i^*$  or  $T_i^*$ , or just the unconditional means of either of these variables.

It is interesting to note from Figs. 6 and 7 that, for each of the measures of steepness considered, there is a very clear trend, with tight confidence intervals, between mean steepness and wave height for waves with  $H^* > 1$ . Obviously there are few extremely large waves and the confidence intervals here become wider due to lack of data. For small waves, with  $H^* < 1$ , this trend is reversed. It is worth noting that from the analysis of wave buoy records collected under moderate conditions, it has, in the past, been assumed in ocean engineering practise that wave steepness is largely independent of wave height. Further, while vertical wave asymmetry in deep water is well recognised and reflected in non-linear wave theories, horizontal asymmetry is not; and the second-order wave theories now beginning to be adopted by designers do not allow for

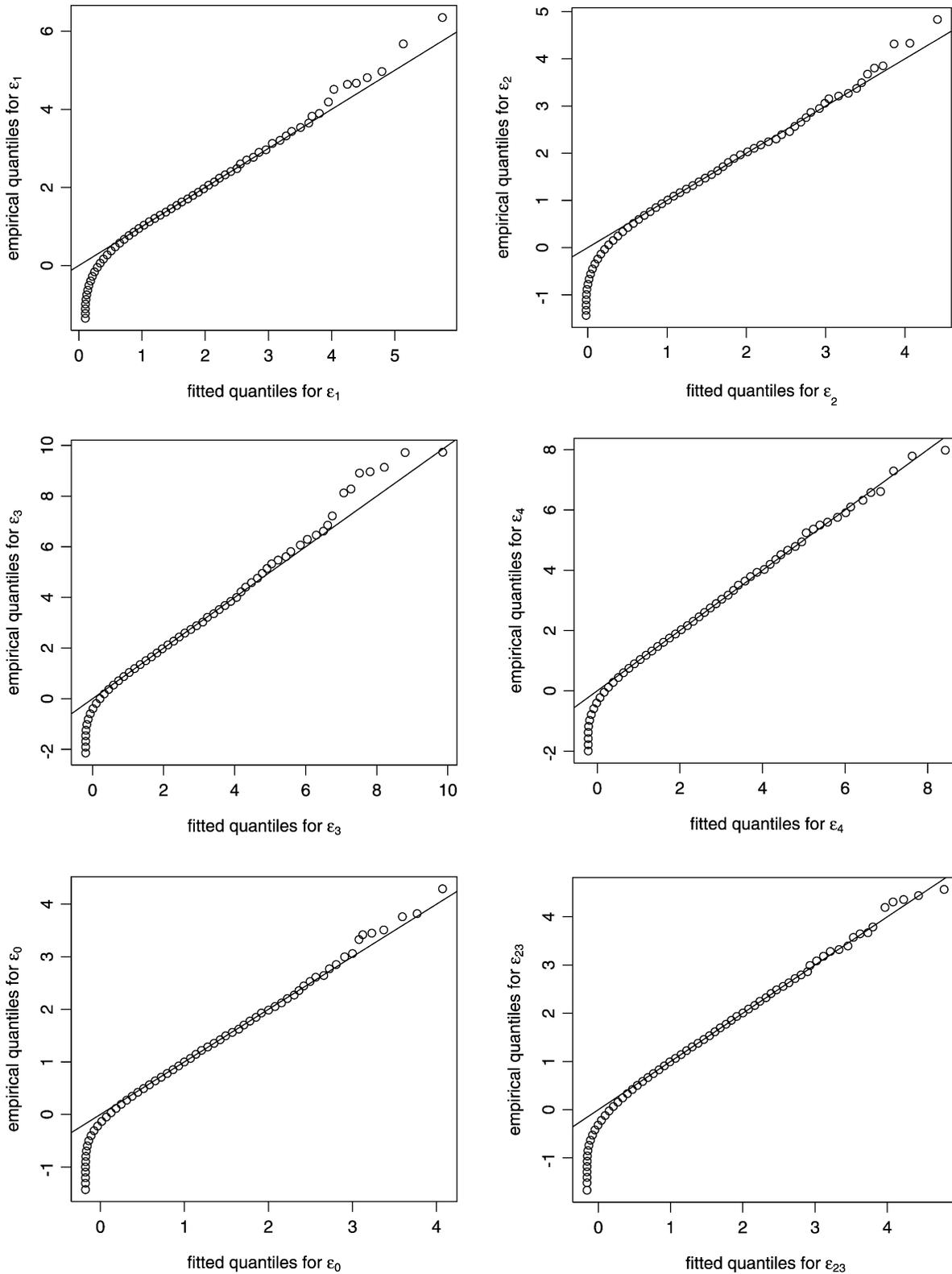


Fig. 11. Quantile–quantile plots of empirical distributions and GPD fits of steepness residuals  $\epsilon_i$  from the data from Bin 2.

it. Yet, it has considerable significance for wave impact forces and greenwater on the decks of floating vessels.

It would be possible to calculate confidence intervals for the predictions of extreme wave steepness

using bootstrap techniques, but the authors have not done so as it is not clear how such confidence intervals could be usefully be employed by designers in practise.

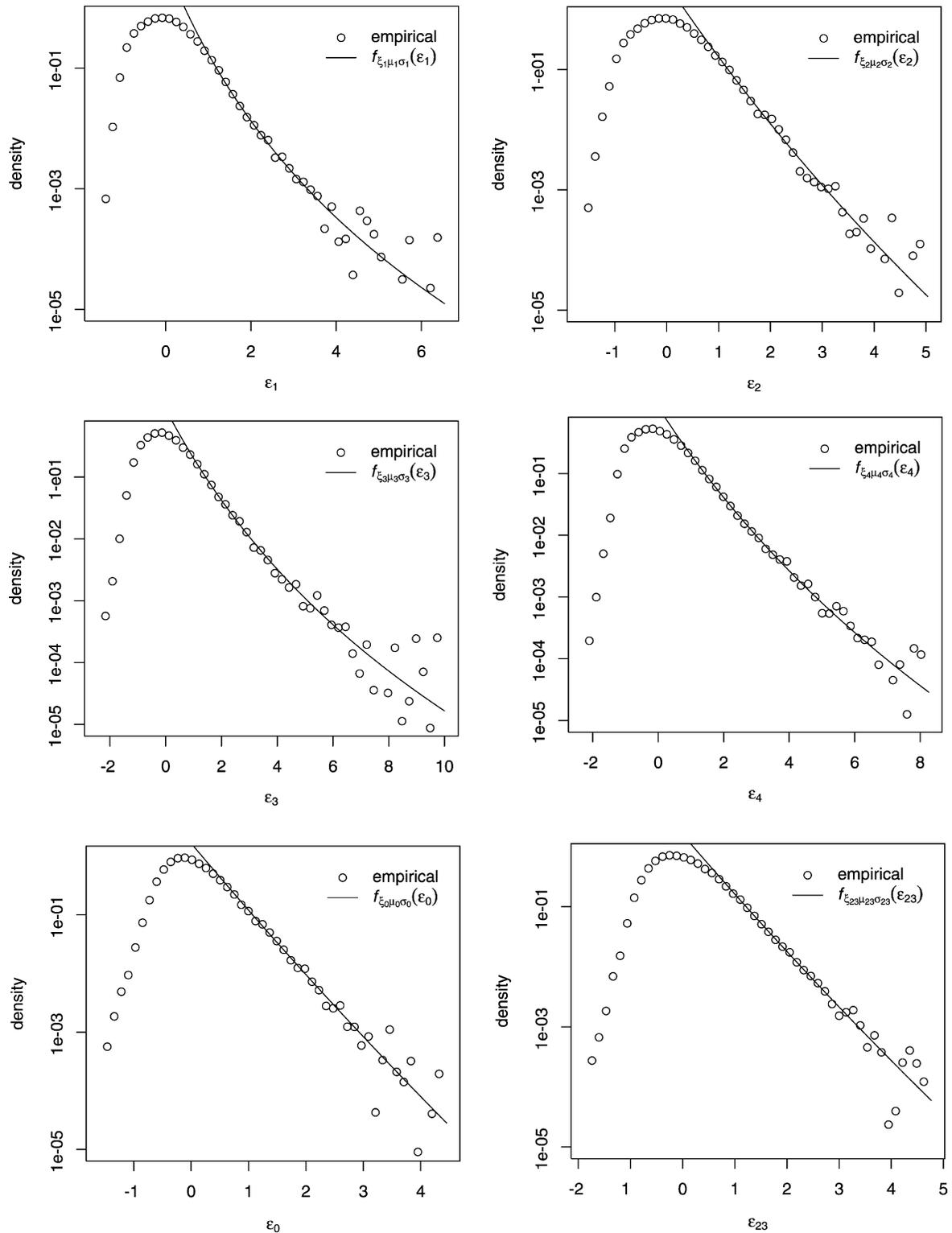


Fig. 12. Empirical and fitted GPD densities of steepness residuals  $\epsilon_i$  from the data from Bin 2.

### 8. Conclusions

There is horizontal asymmetry in the deep water storm waves of a wind-driven sea. The average horizontal asymmetry is not discernible among small waves, but

among waves of mean wave height and greater, it seems to increase with wave height. There is a clear relationship between steepness and wave height for waves with dimensionless height  $H/\bar{H} = H^* > 1$ , with steepness increasing as wave height increases. A simple statistical model has been

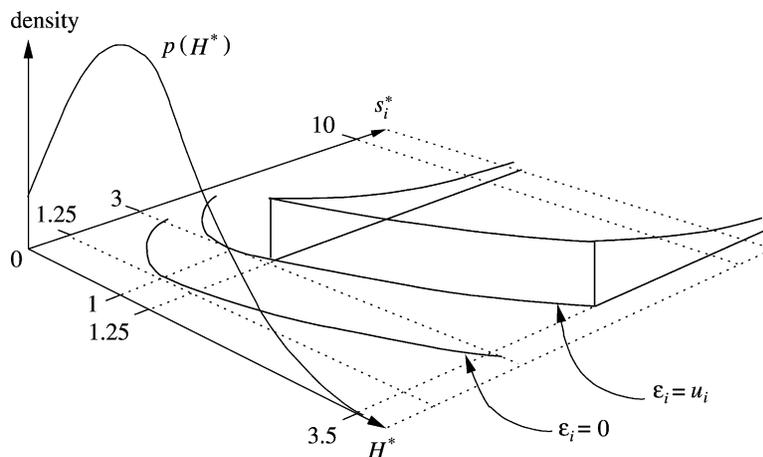


Fig. 13. Schematic illustration in  $(H^*, s_i^*)$ -space of the wedge-shaped region of conditional density, given  $H^*$ , fitted by the GPD. Also shown is a possible (for example, Rayleigh) density of  $H^*$ .

developed for predicting the probability of extreme wave steepness given dimensionless wave height. This model is applicable to extreme storm conditions of the type encountered in the northern North Sea, similar to those used in the analysis.

#### Acknowledgements

The authors are grateful to UK Engineering & Physical Sciences Research Council (EPSRC Grant Number GR/M15439) for funding Paul Stansell and to TotalFinaElf who provide support for Julian Wolfram [14]. The views expressed, however, are those of the authors' and do not necessarily coincide with those of either of the above mentioned organisations. The authors would also like to acknowledge the useful comments from the referees.

#### Appendix A. Non-dimensionalisation of surface gravity wave equations

The equations governing the behaviour of water waves are [13, page 434]

$$\left( \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) \phi = 0, \quad -h_0 < z < \eta, \quad (\text{A1})$$

$$\frac{\partial \phi}{\partial x} \frac{\partial h}{\partial x} + \frac{\partial \phi}{\partial y} \frac{\partial h}{\partial y} + \frac{\partial \phi}{\partial z} = 0, \quad z = -h_0 \quad (\text{A2})$$

$$\frac{\partial \eta}{\partial t} + \frac{\partial \phi}{\partial x} \frac{\partial \eta}{\partial x} + \frac{\partial \phi}{\partial y} \frac{\partial \eta}{\partial y} - \frac{\partial \phi}{\partial z} = 0, \quad z = \eta, \quad (\text{A3})$$

$$\frac{\partial \phi}{\partial t} + \frac{1}{2} \left[ \left( \frac{\partial \phi}{\partial x} \right)^2 + \left( \frac{\partial \phi}{\partial y} \right)^2 + \left( \frac{\partial \phi}{\partial z} \right)^2 \right] + gz = 0, \quad (\text{A4})$$

$$z = \eta,$$

where  $\eta = \eta(x, y, t)$  is the surface elevation,  $\phi = \phi(x, y, z, t)$  is the velocity potential,  $g$  is the acceleration due to gravity, and  $z$  is the distance in the vertical. The dimensions are the following:

$$[\eta] = L,$$

$$[\phi] = \frac{L^2}{T},$$

$$[g] = \frac{L}{T^2},$$

One can non-dimensionalise with characteristic length and time measures to obtain:

$$x = Lx^*,$$

$$y = Ly^*,$$

$$z = Lz^*,$$

$$h_0 = Lh_0^*,$$

$$t = Tt^*,$$

$$\eta(x, y, z, t) = \eta(Lx^*, Ly^*, Tt^*), = L\eta^*(x^*, y^*, t^*),$$

$$\phi(x, y, z, t) = \phi(Lx^*, Ly^*, Lz^*, Tt^*), = \frac{L^2}{T} \phi^*(x^*, y^*, z^*, t^*),$$

where all the variables superscripted by asterisk are dimensionless. Thus, the water wave governing equations can be written in the dimensionless variables as

$$\left( \frac{\partial^2}{\partial x^{*2}} + \frac{\partial^2}{\partial y^{*2}} + \frac{\partial^2}{\partial z^{*2}} \right) \phi^* = 0, \quad -h_0^* < z^* < \eta^*,$$

$$\frac{\partial \phi^*}{\partial x^*} \frac{\partial h^*}{\partial x^*} + \frac{\partial \phi^*}{\partial y^*} \frac{\partial h^*}{\partial y^*} + \frac{\partial \phi^*}{\partial z^*} = 0, \quad z^* = -h_0^*,$$

$$\frac{\partial \eta^*}{\partial t^*} + \frac{\partial \phi^*}{\partial x^*} \frac{\partial \eta^*}{\partial x^*} + \frac{\partial \phi^*}{\partial y^*} \frac{\partial \eta^*}{\partial y^*} - \frac{\partial \phi^*}{\partial z^*} = 0, \quad z^* = \eta^*,$$

$$\frac{\partial \phi^*}{\partial t^*} + \frac{1}{2} \left[ \left( \frac{\partial \phi^*}{\partial x^*} \right)^2 + \left( \frac{\partial \phi^*}{\partial y^*} \right)^2 + \left( \frac{\partial \phi^*}{\partial z^*} \right)^2 \right] + \frac{gT^2}{L} z^* = 0,$$

$$z^* = \eta^*.$$

We see, from the last term in the last of these equations, that the condition for dynamical similarity is

$$g^* = \frac{gT^2}{L} = \text{constant},$$

which implies that the scaling relationship  $T^2 = L$  must be imposed if the term including the acceleration due to gravity be equal to  $g$  alone.

This is equivalent to the following. If  $\eta = \eta(x, y, t)$  and  $\phi = \phi(x, y, z, t)$  are solutions to the governing equations (A1)–(A4) in a reference frame,  $\Sigma$  say, then in a scaled reference frame,  $\Sigma'$ , defined by the transformations

$$x = a^2 x',$$

$$y = a^2 y',$$

$$z = a^2 z',$$

$$t = at',$$

which imply

$$\frac{\partial}{\partial x} = \frac{1}{a^2} \frac{\partial}{\partial x'},$$

$$\frac{\partial}{\partial y} = \frac{1}{a^2} \frac{\partial}{\partial y'},$$

$$\frac{\partial}{\partial z} = \frac{1}{a^2} \frac{\partial}{\partial z'},$$

$$\frac{\partial}{\partial t} = \frac{1}{a} \frac{\partial}{\partial t'},$$

then the scaled functions

$$\eta' = L\eta = a^2 \eta,$$

$$\phi' = \frac{L^2}{T} \phi = a^3 \phi,$$

are solutions of the governing equations (A1)–(A4) in  $\Sigma'$ .

## References

- [1] Coles S. An introduction to statistical modeling of extreme values. Series in Statistics, Berlin: Springer; 2001. ISBN 1-85233-459-22.
- [2] Embrechts P, Klüppelberg C, Mikosch T. Modelling extremal events. Berlin: Springer; 1997.
- [3] Faltinse OM, Greco M, Landrini M. Green water loading on a FPSO. J Offshore Mech Arctic Engng 2002;124:97–103.
- [4] Jonathon P, Taylor PH, Tromans PS. Storm waves in the northern North Sea. In Proceedings of the 17th International Conference on the Behaviour of Offshore Structures, Cambridge, MA, USA, vol. 2.; 1994. p. 481–94.
- [5] Kimura A. Statistical properties of random wave groups. In Proceedings of the 17th International Conference on Coastal Engineering, Sydney, Australia; 1980. p. 2955–73.
- [6] Loader C. Local regression and likelihood. Berlin: Springer; 1999. ISBN 0-387-98775-4.
- [7] Myrhaug D, Kjeldsen SP. Steepness and asymmetry of extreme waves and the highest waves in deep water. Ocean Engng 1986;13(6): 549–68.
- [8] Podgórski K, Rychlik I, Machado UEB. Exact distributions for apparent wave in irregular seas. Ocean Engng 2000;27:979–1016.
- [9] Stansell P, Wolfram J, Linfoot B. Effect of sampling rate on wave height statistics. Ocean Engng 2002;29:1023–47.
- [10] Stansell P, Wolfram J, Linfoot B. Improved joint probability distribution for ocean wave heights and periods. Journal of fluid Mechanics, in press.
- [11] Tromans PS, Anatürk AR, Hagemeyer PM. A new model for the kinematics of large ocean waves—application as a design wave. In Proceedings of the First International Offshore and Polar Engineering Conference, Edinburgh, UK, vol. 3.; 1991. p. 54–71.
- [12] Vinje T, Haver S. On the non-Gaussian structure of ocean waves. In Proceedings of the 17th International Conference on the Behaviour of Offshore Structures, Cambridge, MA, USA, vol. 2.; 1994. p. 453–80.
- [13] Whitham GB. Linear and nonlinear waves. Pure and applied mathematics: a Wiley-interscience series of texts, monographs and tracts, New York: Wiley; 1974. ISBN 0-471-94090-9.
- [14] Wolfram J, Linfoot B, Venugopal V. Some results from the analysis of Metocean data collected during storms in the northern North Sea. J Soc Underwater Technol 2001;24(4):153–63.