Short-term statistics of waves observed in deep water

Mercè Casas-Prat[1] and Leo H. Holthuijsen[2]

Received 26 August 2009; revised 21 January 2010; accepted 8 March 2010; published 29 September 2010.

[1] The short-term statistics of 10 million individual waves observed with buoys in deep water have been investigated, corrected for a sample-rate bias, and normalized with the standard deviation of the surface elevation (the range of normalized wave heights is $0 < \tilde{H} < 10$). The observed normalized trough depths are found to be Rayleigh distributed with near-perfect scaling. The normalized crest heights are also Rayleigh distributed but 3% higher than given by the conventional Rayleigh distribution. The observed normalized wave heights are not well predicted by the conventional Rayleigh distribution (overprediction by 9.5% on average), but they are very well predicted by Rayleigh-like distributions obtained from linear theories and by an empirical Weibull distribution (errors <1.5%). These linear theories also properly predict the observed monotonic variation of the normalized wave heights with the (de-)correlation between crest height and trough depth. The theoretical Rayleigh-like distributions may therefore be preferred over the empirical Weibull distribution and certainly over the conventional Rayleigh distribution. The values of the observed expected maximum wave height (normalized) as a function of duration are consistent with these findings. To inspect nonlinear effects, the buoy observations were supplemented with 10,000 waves observed with laser altimeters mounted on a fixed platform ($0 < \tilde{H} < 7$). The (normalized) crest heights thus observed are typically 5% higher than those observed with the buoys, whereas the (normalized) trough depths are typically 12% shallower. The distribution of the normalized wave heights thus observed is practically identical to the distribution observed with the buoys. These findings suggest that crest heights and trough depths are affected by nonlinear effects, but wave heights are not. One wave in our buoy observations may qualify as a freak wave.

Citation: Casas-Prat, M., and L. H. Holthuijsen (2010), Short-term statistics of waves observed in deep water, *J. Geophys. Res.*, *115*, C09024, doi:10.1029/2009JC005742.

1. Introduction

[2] The short-term statistics of wind-generated waves in deep water are often based on the assumption that the sea surface elevation is a stationary, Gaussian process, resulting in a Rayleigh distribution for crest heights, through depths and wave heights (assuming that the wave heights are twice the crest heights [*Longuet-Higgins* 1952]). Many studies have been carried out, either to find a better distribution or simply to demonstrate agreement with observations in the field [e.g., *Cartwright and Longuet-Higgins*, 1956; *Forristall*, 1978, 1984; *Nolte and Hsu*, 1979; *Longuet-Higgins*, 1980; *Larsen*, 1981; *Vinje*, 1989; *Boccotti*, 1989, 2000; *Tayfun*, 1990, 1994, 2004, 2006; *Mori and Yasuda*, 2002; *Stansell*, 2004, 2005; *Mori and Janssen*, 2006]. The number of observed or numerically simulated waves in these studies is typically 10,000–50,000 waves, with notable exceptions with fewer waves [e.g., *Mori and Janssen*, 2006] or more waves [e.g., *Stansell*, 2004; *Vinje*, 1989]. An excellent review of the literature on the subject has been given by *Tayfun and Fedele* [2007]. With our records of some 10 million waves, we can inspect more extreme statistics than those in the previous studies.

[3] Our data were measured with Waverider buoys of the Xarxa d'Instruments Oceanogràfics i Meteorològics de la Generalitat de Catalunya network off the Catalan coast of Spain in the Mediterranean Sea [*Bolaños et al.*, 2009]. We are well aware of the fact that, owing to the hydrodynamic characteristics of a buoy, the peaks of steep waves tend to be flatter in the time records than they actually are. In fact, a buoy seems to linearize the waves, i.e., the waves look more sinusoidal than they would be when observed with a fixed instrument [*James*, 1986; *Magnusson et al.*, 1999]. Nonlinear effects such as the sharp peaks of steep waves therefore cannot be properly investigated with a buoy. Another potential problem is that a buoy may be dragged through or swerve around the 3-D peaks of waves [*Allender et al.*, 1989]. However, most wave measurements at sea are made with buoys, and understanding the statistics thus obtained is important, leaving the interpretation in terms of a fixed instrument as a separate issue. But to see the effect of

[1]Escola Tècnica Superior d'Enginyers de Camins, Canals i Ports de Barcelona, Technical University of Catalonia, Barcelona, Spain.
[2]Faculty of Civil Engineering and Geosciences, Delft University of Technology, Delft, Netherlands.

0148-0227/10/2009JC005742

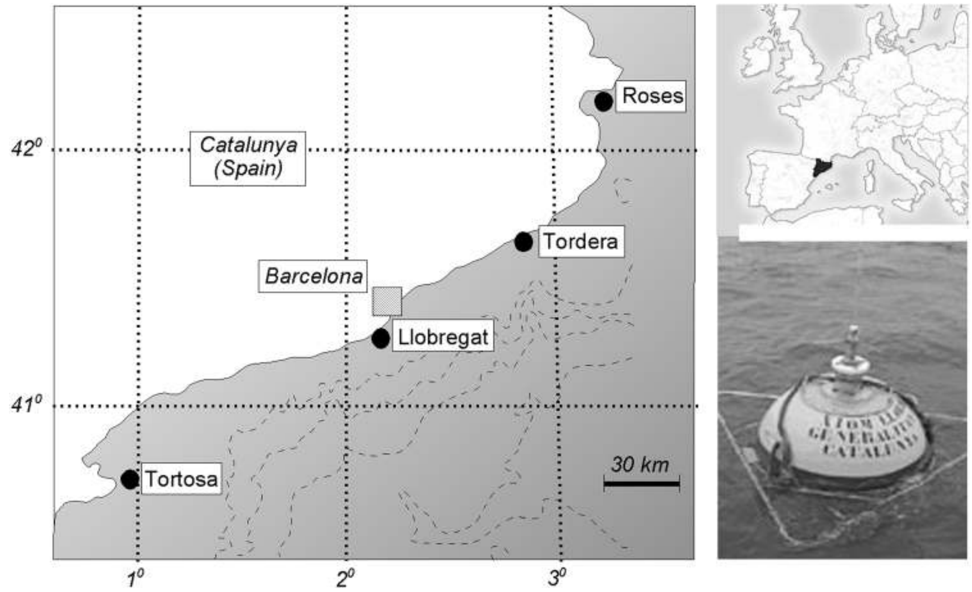

Figure 1. The location of the four WAVERIDER buoys off the Catalan coast used in this study.

using a fixed instrument rather than a floating buoy, we also considered some 10,000 waves observed with laser altimeters from an offshore platform in the North Sea during the Wave Direction Measurement Calibration Project (WADIC) experiment [*Allender et al.*, 1989].

[4] The questions that seem most relevant to scientists and engineers are as follows: (1) To what extent are wave heights, crest heights, and trough depths Rayleigh distributed (or should modifications be applied)? (2) If they are Rayleigh distributed, can the scale of this distribution be predicted from the spectrum of the waves? and (3) Can the maximum crest height and wave height in a given duration correspondingly be predicted from the spectrum? These questions have been addressed extensively in the review of *Tayfun and Fedele* [2007]. They concluded, on the basis of comparing theoretical expectations with real ocean waves, that nonlinear effects do not affect wave heights and that the linear model of *Boccotti* [1989, 2000] describes observed wave heights extremely well.

[5] We therefore focus initially on the conventional Rayleigh distribution (i.e., with its scale parameter equal to the standard deviation of the sea surface elevation). But because of its shortcomings, we also consider Rayleigh-like distributions with other values of the scale parameter that account for the (de-)correlation between the crest height and the trough depth of a wave. We do not consider nonlinear theories [e.g., *Longuet-Higgins*, 1963; *Tayfun*, 1994; *Mori and Yasuda*, 2002; *Janssen*, 2003; *Kharif and Pelinovsky*, 2003; *Mori and Janssen*, 2006; *Onorato et al.*, 2009] because (1) the surface elevations in our data are nearly Gaussian distributed (so that nonlinear theories reduce to the linear theory), which is closely related to the fact that (2) the buoys cannot properly observe nonlinear effects and to the extent that they do, they add their own nonlinear effects which are not included in the nonlinear wave theories, and (3) the nonlinear effects seem to be small for the wave heights as shown by *Tayfun and Fedele* [2007], although they may be larger for the wave crests.

[6] We describe our primary data and initial analysis in section 2 and the theoretical and empirical expectations in section 3. In sections 4 and 5, we show the results of our statistical analysis, and in section 6 we discuss our findings and formulate our conclusions.

2. Observations and Initial Analysis

2.1. Observations

[7] Most of the observations that we analyze are time series of the sea surface elevation measured by four Waverider buoys in the Mediterranean Sea off the Catalan coast of Spain at locations Roses, Tordera, Llobregat, and Tortosa (see Figure 1) during a 15 year period (1991–2006). The exact location, depth, and other specifics of the buoys are given in Table 1. The Tortosa buoy is a directional buoy (but we used only the vertical buoy motion), and the others are scalar buoys (measuring only the vertical motion).

[8] The discretization of the sea surface elevation in time poses a potential problem inasmuch as wave heights are systematically underestimated by equidistant sampling. *Tayfun* [1993] has shown that the corresponding bias of the average normalized wave heights $\Delta = (\tilde{H}_{1/p} - \tilde{H}'_{1/p})/\tilde{H}_{1/p}$ would be $\Delta = -(\pi\Delta t/T_{m01})^2/6$, in which $\tilde{H}_{1/p}$ is the mean of the $1/p$ highest true normalized wave heights and $\tilde{H}'_{1/p}$ is the same of the observed normalized heights, Δt is the sample rate and T_{m01} is the mean wave period (see equation (4) for definition). Because the mean value in the buoy observations of $(\Delta t/T_{m01})^2$ is 0.0175 and 0.0168 for the laser altimeter observations, the bias value would be 2.9% and 2.8%, respectively. Because *Tayfun* [1993] shows that this applies to the distributions of normalized wave heights, we have upscaled all of our corresponding distributions accordingly.

[9] The buoy data that we received were raw data; i.e., they were obtained by integrating the vertical acceleration of the buoy without any quality control to detect errors, anomalies, or other unwanted phenomena. We therefore

Table 1. Specifics of the Buoy Observations off the Catalan Coast and of the Altimeter Observations in the North Sea

	Roses	Tordera	Llobregat	Tortosa	Edda
Coordinates	03 11.99°E 42 10.79°N	02 48.93°E 41 38.81°N	02 08.48°E 41 16.69°N	00 58.89°E 40 43.29°N	03 28°E 56 28°N
Depth (m)	46	74	45	60	70
Diameter (m)	0.7	0.7	0.7	0.9	–
Sample interval (s)	1/2.56	1/2.56	1/2.56	1/1.28	1
Vertical resolution (m)	0.01	0.01	0.01	0.01	0.001
Instrument	Scalar buoy	Scalar buoy	Scalar buoy	Directional buoy	Laser altimeters
Record length	20 min	20 min	20 min	20 min	17 min 4 s
Period	2001–2006	2002–2006	2001–2004	1991–1997, 2001–2006	5–6 Nov 1985, 21–23 Dec 1985

subjected the data to some rigorous tests. We did not consider repairing the observations, because we had sufficient data to work with (initially some 130,000 records with a total of 40 million waves). Of the original records, we accepted only those for which the following criteria were met (in order of testing):

[10] 1. The record has the nominal length (see Table 1).

[11] 2. All absolute values of the vertical accelerations are $<1/2g$ (the maximum acceleration in a Stokes wave, where g is gravitational acceleration).

[12] 3. After removing the linear trend in the record, no linear sections in the surface elevation occur (i.e., sequences of more than three consecutive data points with zero acceleration).

[13] 4. The significant wave height $H_s = 4\sigma$ (σ is the standard deviation of the surface elevation in the record) is >0.5 m (the size, shape, and weight of the buoy may affect the measurement of smaller waves). With a characteristic wave steepness of 4%, say, this implies a characteristic wave length >12.5 m, so that the buoy diameter (0.7 m and 0.9 m) is only a small fraction of that wave length.

[14] 5. No two consecutive data points occur $>2.83H_s$. A longer sequence is accepted as it may indicate the occurrence of a freak wave (the value $2.83H_s$ is exceeded on average once per 100 storms, each with 2,000 waves and a Benjamin-Feir index $BFI = 0.8$; see Appendix A).

[15] 6. Mean frequency f_{m01} (see equation (3) for definition) is less than the Nyquist frequency $f_{\mathrm{Nyq}} = 1/(2\Delta t)$ divided by 2.2 (i.e., $f_{\mathrm{Nyq}} > 2.2\ f_{m01}$; to avoid excessive aliasing errors). (See *Holthuijsen* [2007] for details about the Nyquist frequency.)

[16] 7. Low-frequency variance density is <0.004 m^2/Hz (i.e., <0.0065 Hz for the buoy data and <0.008 Hz for the altimeter data) to avoid records in which the buoy showed slow oscillations such as those induced by a boat or by flotsam hitting the buoy.

[17] 8. The mean wave length (determined with the linear wave theory from the above mean frequency) is smaller than one half the local water depth.

[18] In the records thus selected, we found a few unusual looking records. These almost always had skewness or kurtosis values deviating considerably from the values for a Gaussian variable (0 and 3). We found 35 records with skewness >0.3 or kurtosis >4.0, and after a visual inspection we rejected 16 of these (none of these contained obvious freak waves). This is admittedly a subjective criterion which we can only illustrate with an example shown in Figure 2. Such a small number does not seem to compromise the statistics of the remaining records. With these nine criteria (including the visual inspection), we accepted 42,377 buoy records with a total of approximately 10 million waves. The highest individual wave height in this data set is 8.53 m, and the highest significant wave height is 5.38 m.

[19] To find some indication of the differences between observations with a moving buoy and observations with a fixed instrument, we supplemented our buoy data with measurements obtained with two laser altimeters mounted

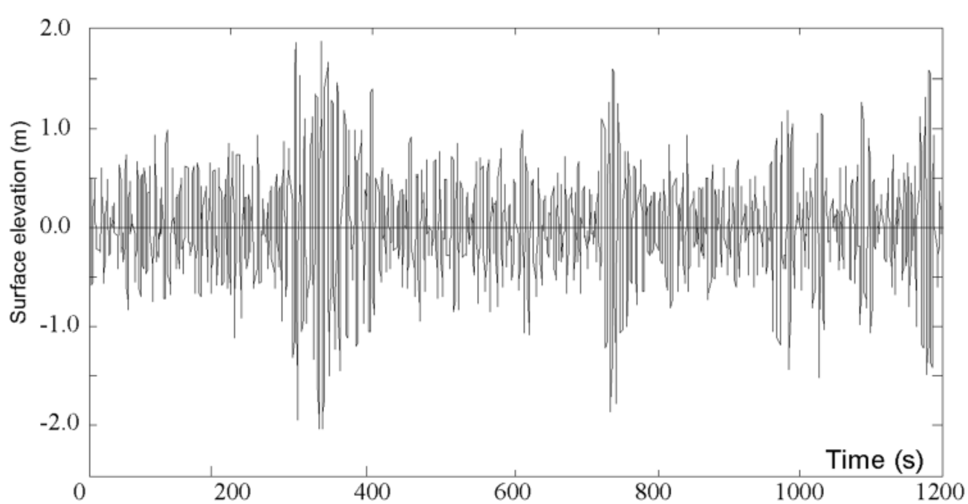

Figure 2. Example of an unusual looking record that was rejected (two unusually long and high wave groups; kurtosis is 4.1).

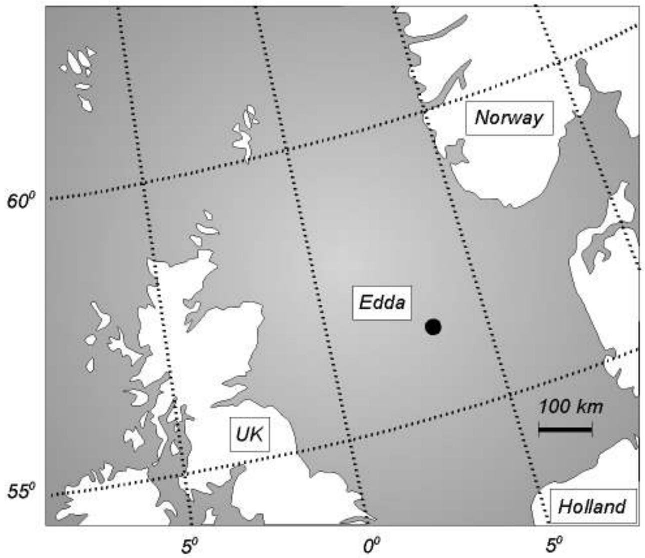

Figure 3. The location of the WADIC experiment, near the EDDA platform.

on the Phillips Edda platform in the North Sea during the WADIC project [*Allender et al.*, 1989]. Most of the data from this project were collected during the period of October 1985 to January 1986 with 20 instruments. In the present study, 92 time series that were registered during November–December 1985 by two laser altimeters from a pentagon array mounted on the Phillips Edda platform (see Figure 3) were at our disposal. Details such as water depth and sample interval time are given in Table 1. These laser observations represent two independent storms, one of which had significant wave heights >10 m. We did not censor these records as they were used in the WADIC experiment after recovery of 87% of the initial raw data. We only considered the requirement of shallow water (see above criteria), which caused 23 records to be rejected (including the ones with the highest significant wave height in November 1985). The remaining 69 records contained about 10,000 waves. The highest individual wave height in these records is 13.55 m, and the highest significant wave height is 8.82 m.

2.2. Initial Analysis

[20] To define wave parameters such as wave height and wave crest we defined a wave as the surface profile between two consecutive zero-down crossings. It is important to note that we used this down-crossing definition because the statistics of the alternative up-crossing definition may well be different [*Haring et al.*, 1976]. The zero-crossing points were obtained by linear interpolation of the surface elevation in time between the nearest two data points (after removing the linear trend in the record). The individual crest height η_{crest}, trough depth η_{trough}, wave height $H = \eta_{\text{crest}} - \eta_{\text{trough}}$, and period T were thus defined from the maximum and minimum elevation in each individual wave and from the zero-crossing points. We subsequently computed for each record the significant wave height as follows:

$$H_{1/3} = \frac{1}{N/3} \sum_{i=1}^{N/3} H_i, \tag{1}$$

in which N is the number of waves in the record and i is the rank number of the wave height (i = 1 being the highest wave). To avoid excessive effects of measurement noise near the zero level, in particular on the estimation of the number of waves, we ignored all wave heights <0.05 m and all crest heights and trough depths <0.025 m and periods smaller than twice the sampling interval time.

[21] We also carried out a spectral analysis for each record. We estimated the spectrum as $E(f) \approx \overline{1/2a_j^2}/\Delta f$, in which a_j is the amplitude of the harmonic with frequency f_j and the overbar indicates averaging over the frequency band Δf. We used a standard fast Fourier transform technique (tapering the first 5% and last 5% of each record with a cosine function). The Nyquist frequency was 0.64 Hz for the Tortosa buoy, 1.28 Hz for the other buoys, and 0.5 Hz for the altimeters. We thus obtained a spectrum for each record with a frequency resolution of about Δf = 0.013 Hz for the buoy data and Δf = 0.016 Hz for the altimeter data and with an error in the spectral density of ~25% (standard deviation). In addition, we computed for each record the following wave parameters on the basis of the moments of the spectrum $m_n = \int_0^\infty f^n\, E(f)\, df$:

$$\text{Significant wave height} \quad H_{1/3,\,\text{Rayleigh}} = 4\sqrt{m_0} \tag{2}$$

$$\text{Mean frequency} \quad f_{m01} = m_1/m_0 \tag{3}$$

$$\text{Mean wave period} \quad T_{m01} = 1/f_{m01} \tag{4}$$

From the definition of the spectrum, it follows that $\sqrt{m_0}$ is the standard deviation of the surface elevation.

3. Theoretical and Empirical Expectations

3.1. The Individual Wave Height and Crest Height

[22] The conventional approach to estimating the short-term statistics of crest height, trough depth, and wave height is a linear approach in the sense that the waves are assumed to be the sum of a large number of independent harmonic waves with random, uniformly distributed phases (leading to a stationary Gaussian process). Under these conditions and for waves with a narrow-band spectrum, *Longuet-Higgins* [1952] showed that the crest height (and by implication trough depth) is Rayleigh distributed:

$$p(\tilde{\eta}_{\text{crest}}) = \tilde{\eta}_{\text{crest}} \exp\left(-1/2\tilde{\eta}_{\text{crest}}^2\right) = p_{\tilde{\eta}_{\text{crest}},\text{Rayleigh}}, \tag{5}$$

in which $p(\tilde{\eta}_{\text{crest}})$ is the probability density function of $\tilde{\eta}_{\text{crest}} = \eta_{\text{crest}}/\sqrt{m_0}$, and $\sqrt{m_0}$ is the scaling factor (and likewise for the trough depth). We refer to this Rayleigh distribution as the conventional Rayleigh distribution for crest heights. *Longuet-Higgins* [1980] pointed out that in his original formulation of 1952, $\tilde{\eta}_{\text{crest}} = \eta_{\text{crest}}/(\sqrt{2}\,\overline{\eta}_{\text{crest}})$, in which $\overline{\eta}_{\text{crest}}$ is the mean crest height rather than the above $\tilde{\eta}_{\text{crest}} = \eta_{\text{crest}}/\sqrt{m_0}$. But for Gaussian sea states, the two normalizations are identical. *Cartwright and Longuet-Higgins* [1956] gave a more complicated expression for local maxima rather than crest heights as defined here, for a spectrum with an arbitrary width. This expression involves a spectral width parameter that contains the fourth-order moment of the spectrum $\varepsilon = \sqrt{1 - m_2^2/(m_0 m_4)}$. But because the spectrum of wind-generated waves often has a high-frequency tail proportional

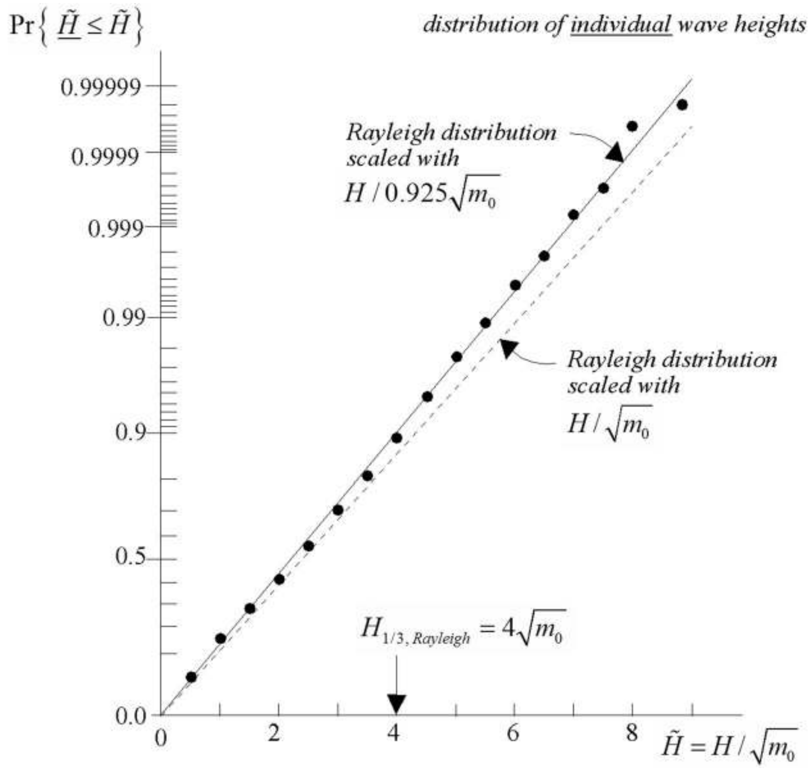

Figure 4. The distribution of normalized observed individual wave heights plotted with Rayleigh scales (plot $\{-\ln[1-\Pr(\underline{x}\leq x)]\}^{1/2}$ against x). The observations arrange themselves along a straight line, illustrating that they are nearly Rayleigh distributed (data from five hurricanes in the Gulf of Mexico from *Forristall* [1978; after *Holthuijsen*, 2007]). The scaling with $0.925\sqrt{m_0}$ was suggested by *Longuet-Higgins* [1980].

to frequency to a power between –4 and –5, the value of this width parameter depends very much on the upper limit of the integration interval. It has therefore been largely abandoned [*Tucker and Pitt*, 2001], and we do not consider this approach.

[23] The short-term statistics of the wave height is often obtained simply by assuming that the wave height is twice the crest height $H = 2\eta_{\text{crest}}$; in other words, that the crest height is equal to the preceding trough depth. Transforming the above Rayleigh distribution, using the proper Jacobian then gives the Rayleigh distribution for the normalized wave height $\tilde{H} = H/\sqrt{m_0}$ as follows:

$$p(\tilde{H}) = \frac{1}{4}\tilde{H}\exp\left(-\frac{1}{8}\tilde{H}^2\right) = p_{\tilde{H},\text{Rayleigh}}. \tag{6}$$

We refer to this Rayleigh distribution as the conventional Rayleigh distribution for wave heights. This assumption of the wave height being equal to twice the crest height seems reasonable as a first approximation, and in fact, the observed shape of the distribution of the wave height is very nearly that of a Rayleigh distribution as shown in Figure 4. However, Figure 4 also shows that often the actual scale of the distribution is smaller than the nominal scale $\sqrt{m_0}$ (*Longuet-Higgins* [1980] suggests for this example scaling with $0.925\sqrt{m_0}$, although *Forristall* [1978] gives $H_{1/3} = 3.77\sqrt{m_0}$, which suggests a scaling of $0.942\sqrt{m_0}$). In other words, the conventional Rayleigh distribution overestimates the wave heights in these observations.

[24] *Longuet-Higgins* [1980] showed that increasing the width of the spectrum by adding a perturbation (i.e., noise) to a narrow-banded process gives a rescaling of the conventional Rayleigh distribution, depending on the width of the spectrum. The result for the wave height, expressed in terms of the spectral width parameter $\nu = \sqrt{m_0 m_2/m_1^2 - 1}$ [*Longuet-Higgins*, 1975] is as follows:

$$p(\tilde{H}) = \frac{\tilde{H}}{4\alpha^2}\exp\left(-\frac{\tilde{H}^2}{8\alpha^2}\right), \tag{7}$$

with the scaling factor $\alpha = \sqrt{1 - (1/8\pi^2 - 1/2)\nu^2}$. Since $H = 2\eta_{\text{crest}}$ is assumed, this correction also applies to the crest heights. In another approach, in which both the crest height and the trough depth are considered to be random variables with a certain correlation ρ (rather than a priori assuming $H = 2\eta_{\text{crest}}$), *Naess* [1985; equation (14)] finds a scaling factor $\alpha = \sqrt{1/2(1-\rho)}$. For large wave heights (*Tayfun* [1990] suggests $\tilde{H} > \sqrt{2\pi} \approx 2.5$), *Boccotti* [1989, 2000; equation (9.69)], and *Vinje* [1989; equation (30)] multiply this distribution (for the same value for α) with a coefficient β, whereas *Tayfun* [1990; equations (29) and (32)] multiplies Vinje's distribution (i.e., for the same value for α and β, referred to in *Tayfun and Fedele* [2007] as the T1 model but with a different correlation coefficient), again

Table 2. The Parameters of the Various Rayleigh(-like) Probability Density Functions for the Normalized Wave Height[a]

	β	α	$f(\tilde{H})$
	All Wave Heights		
Rayleigh [*Longuet-Higgins*, 1952]	1	1	1
Longuet-Higgins [1980]	1	$\sqrt{1-(1/8\pi^2-1/2)\nu^2}$	1
Naess [1985]	1	$\sqrt{1/2(1-\rho_{NB})}$	1
	Large Wave Heights		
*Vinje** [1989][b]	$\sqrt{1/2(1-\rho_{VT}^{-1})}$	$\sqrt{1/2(1-\rho_{VT})}$	1
*Boccotti** [1989][b]	$(1+\rho'')/\sqrt{2\rho''(1-\rho_{NB})}$	$\sqrt{1/2(1-\rho_{NB})}$	1
Tayfun [1990]	$\sqrt{1/2(1-\rho_{VT}^{-1})}$	$\sqrt{1/2(1-\rho_{VT})}$	$[1+(\rho_{VT}^2-1)/(4\rho_{VT}\tilde{H}^2)]$

[a]Variable ρ'' is the (global) minimum correlation between the vertical velocities of the surface elevation (as a function of time lag) in this study taken at $\tau = 1/2T_{m01}$.
[b]With modified correlation coefficients.

with a function $f(\tilde{H})$ that modifies the shape of the distribution. The Rayleigh distribution can thus be generalized to the following:

$$p(\tilde{H}) = \frac{\beta\tilde{H}}{4\alpha^2}\exp\left(-\frac{\tilde{H}^2}{8\alpha^2}\right)f(\tilde{H}). \tag{8}$$

The coefficients and the function $f(\tilde{H})$ are summarized in Table 2.

[25] The correlation ρ may be estimated as $\rho \approx C(\tau_0) = E\{\tilde{\eta}\ (t)\ .\tilde{\eta}\ (t+\tau_0)\}$, where $\tilde{\eta}$ is the normalized surface elevation $\tilde{\eta} = \eta/\sqrt{m_0}$ and τ_0 is the average time interval between the crest and trough of a wave. *Naess* [1989] stated that the precise choice of this interval is not crucial. *Tayfun* [1990] takes $\tau_0 = 1/2T_{m01}$, whereas *Vinje* [1989] and *Boccotti* [1989, 2000] take $\tau_0 = \tau_1$, where the autocorrelation function $C(\tau)$ attains its (global) minimum. To avoid computing $C(\tau)$ and searching for its minimum to find τ_1, we follow *Tayfun* [1990] and take $\tau_0 = 1/2T_{m01}$ in all our estimates of ρ. Following the expressions of *Naess* [1985] and *Boccotti* [1989, 2000], we estimate ρ directly from the spectrum as $\rho_{NB} = m_0^{-1}\int_0^\infty E(f)\cos(2\pi f\tau_0)df$. *Tayfun* [1990] uses a slightly more complicated expression, which we also use in the approach of *Vinje* [1989] (to make it coincide with the T1 model of *Tayfun and Fedele*, 2007]): $\rho_{VT} = -(c^2+s^2)^{1/2}$ with $c = m_0^{-1}\int_0^\infty E(f)\cos(2\pi f\tau_0)df$ and $s = m_0^{-1}\int_0^\infty E(f)\sin(2\pi f\tau_0)df$. We estimate the correlation ρ'' between the vertical velocities (see Table 2) following *Boccotti* [1989, 2000] as $\rho'' = m_2^{-1}\int_0^\infty f^2E(f)\cos(2\pi f\tau_0)df$. Because of replacing τ_1 with $1/2T_{m01}$, we refer to the results thus obtained as the results of the theories of *Vinje* [1989] and *Boccotti* [1989, 2000] with modified correlation coefficients, both denoted with asterisks, i.e., *Vinje** [1989] and *Boccotti** [1989, 2000]. All these Rayleigh-like distributions reduce to the conventional Rayleigh distribution of equation (6) for very narrow spectra ($v = 0$) or for unity correlations ($\rho = 1$ and $\rho'' = 1$).

[26] *Forristall* [1978] suggests, on the basis of observations in five hurricanes in the Gulf of Mexico with $0 < \tilde{H} < 9$, using a two-parameter Weibull distribution:

$$\Pr\{\underline{H} \leq H\} = P(\tilde{H}) = 1-\exp(-\tilde{H}^a/b), \tag{9}$$

with $a = 2.126$ and $b = 8.42$ (note that for $a = 2$ and $b = 8$, this distribution reduces to the conventional Raleigh distribution). *Nolte and Hsu* [1979], using a subset of these data with $0 < \tilde{H} < 6.7$, suggested $a = 2.138$ and $b = 9.08$.

3.2. The Significant Wave Height and Maximum Wave and Crest Height

[27] Given the probability density function for the individual wave height, the value of the significant wave height $H_{1/3}$ (the mean of the one-third highest waves) is readily estimated as follows:

$$H_{1/3} = \int_{H^*}^{\infty} H\,p(H)dH \Big/ \int_{H^*}^{\infty} p(H)dH, \tag{10}$$

where H^* is defined such that $\int_{H^*}^\infty p(H)dH = 1/3$. For the conventional Rayleigh distribution of equation (6), the result in terms of the spectrum is as follows:

$$H_{1/3} \approx H_{1/3,\,\mathrm{Rayleigh}} = 4.004\ldots\sqrt{m_0} \approx 4\sqrt{m_0}, \tag{11}$$

which typically overestimates the significant wave height as determined from time records (see Figure 5). For the scaled Rayleigh distribution (equation (7)), the result is obviously

$$H_{1/3} \approx H_{1/3,\,\mathrm{Rayleigh,\,scaled}} = 4\alpha\sqrt{m_0}. \tag{12}$$

For the other theoretical distributions, a closed analytical expression for the significant wave height seems to be available only for *Tayfun* [1990, equations (48), (49), and (52)]. If required in the following, the significant wave height is determined numerically from the computed theoretical probability density functions. For the two-parameter Weibull distribution, the significant wave height, estimated as the mean of the one-third highest waves in the Weibull distribution, is as follows:

$$H_{1/3} \approx H_{1/3,\,\mathrm{Weibull}} = 3b^{1/a}\Gamma(1+1/a, \ln 3), \tag{13}$$

in which $\Gamma(1+1/a, \ln 3) = \int_{\ln 3}^\infty t^{1/a}e^{-t}dt$ is the incomplete Γ function.

[28] The cumulative distribution function of the maximum crest height or maximum wave height in a record of N waves is readily obtained, assuming that crest heights or wave heights are independent (for large N, the results are fairly insensitive to this assumption). The expressions are identical for the wave height and the crest height. Representing either of these as x gives the following equation:

$$\Pr\{\underline{x}_{\max} < x_{\max}\} = P(x_{\max}) = [P(x)]^N. \tag{14}$$

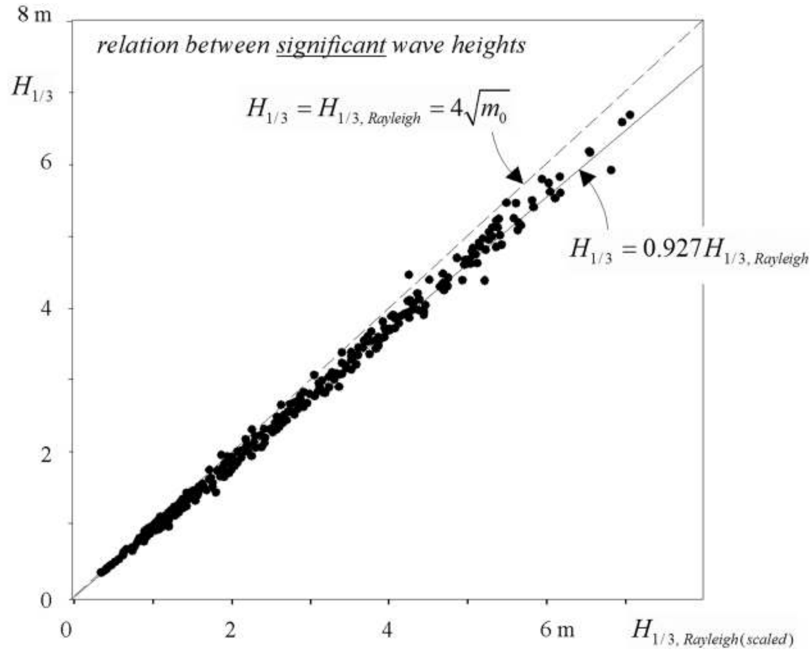

Figure 5. The significant wave height $H_{1/3}$ estimated directly from the time records of the waves, compared with the theoretical estimate $H_{1/3,\mathrm{Rayleigh}} = 4\sqrt{m_0}$ for each record. Buoy observations from location K13 in the southern North Sea (53.13°N, 03.13°E) during December 2003 [from *Holthuijsen*, 2007]. Solid line is the least squares best fit of the data.

If the (scaled) Rayleigh probability density function ($\beta = 1$ in Table 2) is written as follows:

$$p(x) = \frac{x}{\delta/2}\exp\left(-\frac{x^2}{\delta}\right), \tag{15}$$

then the expected value of the maximum [*Longuet-Higgins*, 1952; *Forristall*, 1978] is

$$E\{x_{\max}\} \approx \left(1 + \frac{\gamma_E}{2\ln N}\right)\sqrt{\delta \ln N}, \tag{16}$$

where $\gamma_E = 0.5772$ is Euler's constant. For the conventional Rayleigh distribution $\delta = 2$ or $\delta = 8$ for the crest height and wave height, respectively, and $\delta = 8\alpha^2$ for the scaled Rayleigh distribution for the wave height. For the Rayleigh-like distributions of Table 2 with $f(\tilde{H}) = 1$, *Tayfun and Fedele* [2007] additionally replace N in our equation (16) with βN.

3.3. The Joint North Sea Wave Project (JONSWAP) Spectrum

[29] For an average JONSWAP spectrum [*Hasselmann et al.*, 1973] (with peak-enhancement factor $\gamma = 3.3$ and peak-width parameters $\sigma_a = 0.07$ and $\sigma_b = 0.09$), the value of ν is $\nu = 0.382$. This gives a scaling factor in the approach of *Longuet-Higgins* [1980] $\alpha = 0.945$, which is slightly higher than the scaling factors of 0.925 and 0.927 suggested by Figures 4 and 5 but in good agreement with the suggestion of *Forristall* [1978] that $\alpha = 0.942$ for these observations. For the same spectrum, $\rho = -0.727$ in the *Naess* [1985] approach, with a corresponding scaling factor $\alpha = 0.929$. These numbers, together with those for *Vinje** [1989], *Tayfun* [1990], and *Boccotti** [1989, 2000] are summarized in Table 3. Note that the correlations ρ and ρ'' in the approach of *Vinje** [1989], *Tayfun* [1990], and *Boccotti** [1989, 2000] not only affect the scaling –but also enhance the probability density for high wave heights (without specifying what happens to the lower wave heights). This high-end scaling is illustrated here with the scaling as it applies to the significant wave height, i.e., the ratio of the significant wave height from these modified Rayleigh distributions and from the conventional Rayleigh distribution ($\alpha = 1$), indicated with $H_{1/3,\mathrm{theory}}/H_{1/3,\mathrm{Rayleigh}}$ (Table 3). The fact that the value of $H_{1/3,\mathrm{theory}}/H_{1/3,\mathrm{Rayleigh}}$ of *Naess* [1985]

Table 3. The Relevant Parameters and the Scaling $H_{1/3,\mathrm{theory}}/H_{1/3,\mathrm{Rayleigh}}$ for the Various Approaches for a JONSWAP Spectrum[a]

JONSWAP Spectrum	Parameter	α	β	$H_{1/3,\mathrm{theory}}/H_{1/3,\mathrm{Rayleigh}}$
Longuet-Higgins [1980]	$\nu = 0.382$	0.945	1	0.945
Naess [1985]	$\rho_{NB} = -0.727$	0.929	1	0.929
*Vinje** [1989][b]	$\rho_{VT} = -0.753$	0.936	1.079	0.953
Tayfun [1990]	$\rho_{VT} = -0.753$	0.936	1.079	0.958
*Boccotti** [1989, 2000][b]	$\rho_{NB} = -0.727$, $\rho'' = 0.479$	0.929	1.150	0.962
Forristall [1978]	–	–	–	0.942

[a]All correlations are taken from the observed correlation function (obtained as the fast Fourier transform of the observed spectrum) at $\tau = 1/2T_{m01}$.
[b]With modified correlation coefficients.

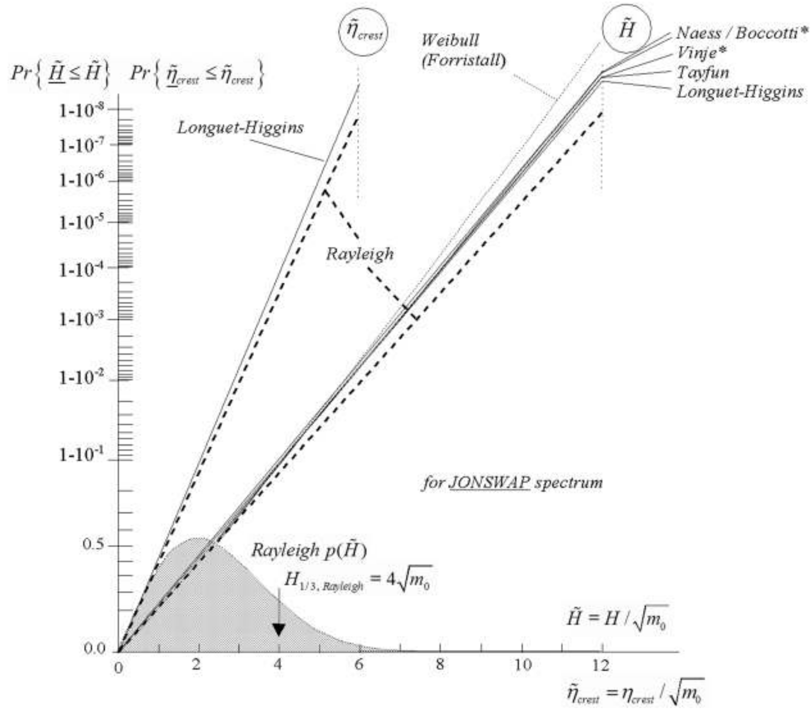

Figure 6. The theoretical distributions for the normalized wave height and crest height for an average JONSWAP spectrum (coefficients from Table 3) and the empirical Weibull distribution of *Forristall* [1978]. The distributions of *Vinje** [1989], *Boccotti** [1989, 2000], and *Tayfun* [1990] are shown for $\tilde{H} > 2.5$. Silhouette of Rayleigh probability density function $p(\tilde{H})$ is added for reference.

is closest to the above average observed values of 0.925 and 0.927 (see Figures 4 and 5) is not indicative of any superiority of this approach as the spectrum of these observations is not necessarily a JONSWAP spectrum. The significant wave height obtained with the empirical Weibull approach of *Forristall* [1978] is approximately equal to the significant wave height in these theoretical approaches (Table 3). It is obvious that for a JONSWAP spectrum, all approaches predict a significant wave height 4%–7% lower than that predicted by the conventional Rayleigh distribution. The corresponding distribution functions are shown in Figure 6 (with Rayleigh scales; see Figure 4). Note that under these (academic) conditions, the theoretical distributions are almost indistinguishable from each other (differences <1.7% in wave heights for $\tilde{H} \approx 12$).

[30] For this academic case of a JONSWAP spectrum, the Weibull distribution of *Forristall* [1978] is almost equal to the theoretical distributions for $\tilde{H} < 5$, say, but predicts much lower wave heights for high wave heights. The difference with the conventional Rayleigh approach is 6% for $\tilde{H} \approx 4$ but 12% for $\tilde{H} \approx 12$. The fact that the significant wave height with the Weibull approach of *Forristall* [1978] is approximately equal to the significant wave height in the other approaches in Table 3 illustrates that this high-end tail of the distribution barely affects the estimation of the significant wave height.

[31] The corresponding expected values of the maximum wave height and of the maximum crest height as a function of the number of waves are shown in Figure 7. Strictly speaking, the values of the maximum wave height for the approaches of *Vinje** [1989], *Tayfun* [1990], and *Boccotti** [1989, 2000] cannot be estimated analytically from the distribution (because the integral of the distribution is not unity). Still, as noted earlier, if $f(\tilde{H}) = 1$, *Tayfun and Fedele* [2007] estimate the expected value of the maximum wave height by replacing N in our equation (16) with βN. But we verified numerically that for waves with an average JONSWAP spectrum and for the approaches with $f(\tilde{H}) = 1$ and $\beta \neq 1$ (i.e., *Vinje** [1989] and *Boccotti** [1989, 2000]; see Table 2), this replacement overestimated the expected values, varying from 3.5% for $N = 10^3$ to 1.5% for $N = 10^7$. We therefore estimated for these three approaches the expected values numerically from the zeroth and the first order moment of the probability density function of the maximum wave height $p(x_{max}) = d[p(x)]^N/dx$ (see equation (14)). This was also done for the Weibull distribution of *Forristall* [1978].

[32] Note that, consistent with the above, the expected maxima in Figure 7 for the theoretical distributions are nearly identical. The results for the Weibull distribution of *Forristall* [1978] are much lower than for all the other approaches, also consistent with the above. The results for the maximum crest heights are based on the conventional Rayleigh distribution ($\alpha = 1$) and the theory of *Longuet-Higgins* [1980]. The other theories mentioned do not apply to the crest height.

4. Buoy Observations

4.1. Wave Height and Crest Height

[33] The large number of waves in our buoy observations allows us to show the distribution of the normalized wave

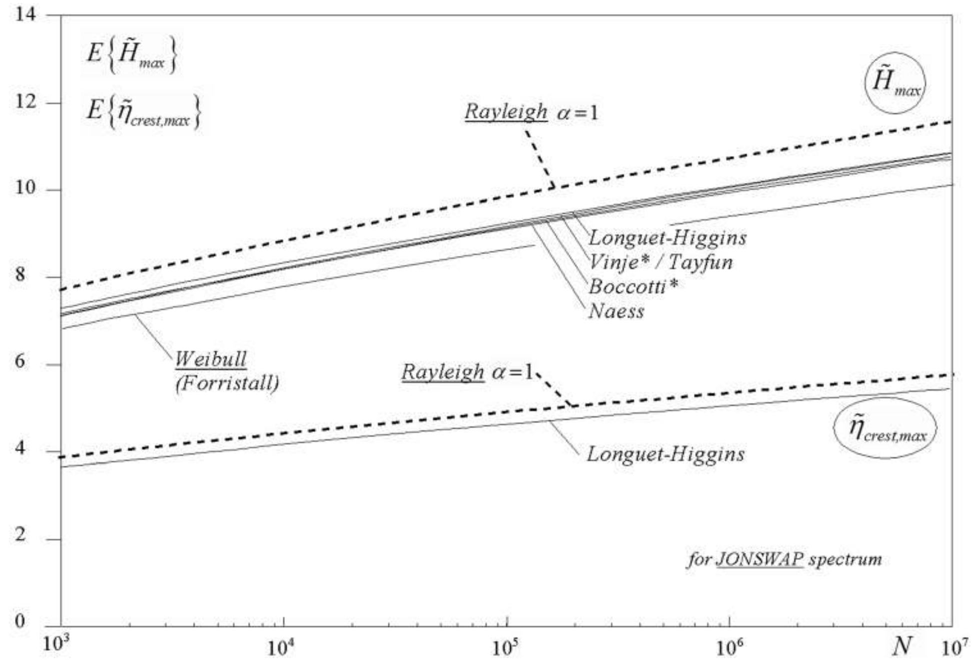

Figure 7. The expected value of the maximum normalized wave height and crest height in N waves for the various Rayleigh-based approaches for an average JONSWAP spectrum (coefficients from Table 3) and for the Weibull distribution of *Forristall* [1978].

heights to rather high values, in fact, up to $H/\sqrt{m_0}$ = 10.35 (after correcting for the sampling-rate bias). To do this, we ranked the individual wave or crest heights and used the plotting position suggested by *Goda* [1988] for Rayleigh distributions:

$$\Pr\{\underline{x} < x\} \approx 1 - (i - a_G)/(N + b_G), \qquad (17)$$

in which x is either the normalized crest height, trough depth, or wave height and the values of the coefficients in this expression are a_G = 0.20 + 0.27/$\sqrt{C}$ and b_G = 0.20 + 0.23$\sqrt{C}$ with C = 2 for a Rayleigh distribution. For the sake of clarity, we show only the results that are located nearest to multiples of 1/2$\sqrt{m_0}$ for the wave heights and nearest to multiples of 1/4$\sqrt{m_0}$ for crest heights and trough depths. We added the highest value in each distribution to show the full range of the observations. The results are shown in Figure 8.

[34] For the normalized crest heights, the agreement between the conventional Rayleigh distribution (α = 1) and the observations is reasonable because the observations cluster well along a straight line and the (least squares) best fit gives a scaling factor α = 1.030. The trough depth is Rayleigh distributed with nearly perfect scaling (α = 0.995 or 0.5% discrepancy with the conventional Rayleigh distribution).

[35] The distribution of the normalized observed wave heights deviates, as expected because of the (de-)correlation between crest height and trough depth, considerably from the conventional Rayleigh distribution (α = 1). These normalized observed wave heights are significantly lower, and the shape of the distribution shows a gentle S-curve (when plotted with Rayleigh scales as in Figure 8), with higher waves at low values and lower waves at high values (at the same levels of probability). A similar S-curve is noticeable in the observations of *Forristall* [1978; Figure 4]. In fact, Figure 8 shows that the Weibull distribution of *Forristall* [1978] (α = 2.155 and b = 8.42) fits our data nearly perfectly (Figure 8), the maximum difference with the observations being ~1.5%. The S-curve is so gentle that the observations cluster reasonably enough around a straight line (the value of α is close enough to α = 2) to say that the normalized observed wave heights are nearly Rayleigh distributed, but with lower values than given by the conventional Rayleigh distribution (α = 1). A least squares best fit with a scaled Rayleigh distribution (equation (6)) gives α = 0.905. Still, the difference in normalized wave height (at the same level of probability) between this best fit scaled Rayleigh distribution and the Weibull distribution as suggested by Forristall is 7% for $\tilde{H} \approx 2$, 5% near $\tilde{H} \approx 4$, and 2.5% near $\tilde{H} \approx 10$ (the Weibull distribution fitting the observations better). These differences also occur in the altimeter observations (see section 6) and may well be significant in an engineering context. We therefore do not support the conclusion of *Longuet-Higgins* [1980] that the introduction of the two-parameter Weibull distribution offers no obvious advantage (even if the parameters of this Weibull distribution cannot as yet be obtained from the spectrum with a predictive theory).

[36] The observed values of the parameters that are relevant for the various theories, the spectral width ν (determined from the observed spectra), and the correlations ρ_{NB}, ρ_{VT}, and ρ'' (determined from the observed surface correlation functions obtained as the Fourier transforms of the observed spectrum multiplied by f^2 when required, at τ = 1/2T_{m01}), are shown in the histograms of Figure 9. The average values are $\overline{\nu}$ = 0.415, $\overline{\rho}_{NB}$ = −0.567, $\overline{\rho}_{VT}$ = −0.618, and $\overline{\rho''}$ = 0.335, respectively (note that this value of $\overline{\nu}$ is almost

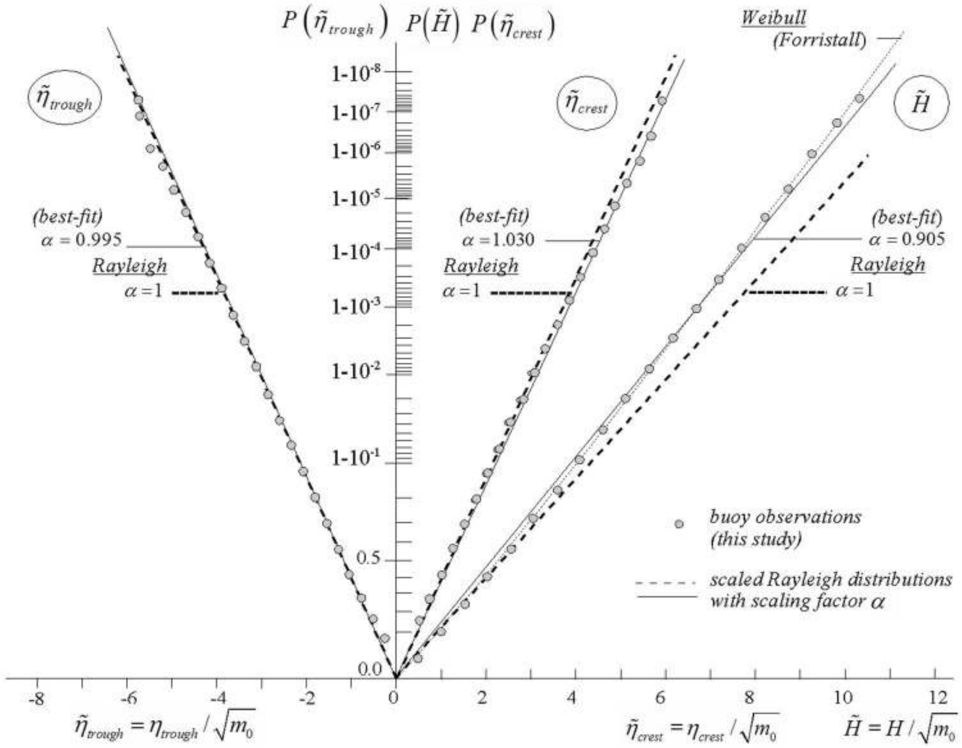

Figure 8. The distributions of the normalized wave height, crest height, and trough depth as observed by the buoys in the present study, the best fit (scaled) Rayleigh distributions and the empirical (Forristall) Weibull distribution. The conventional Rayleigh distributions ($\alpha = 1$) have been added for reference. Only data points nearest to multiples of $1/2\sqrt{m_0}$ or $1/4\sqrt{m_0}$ are shown. The highest normalized observed crest height, wave height, and deepest trough depth are added (top value in each distribution) to show the full range of the observations.

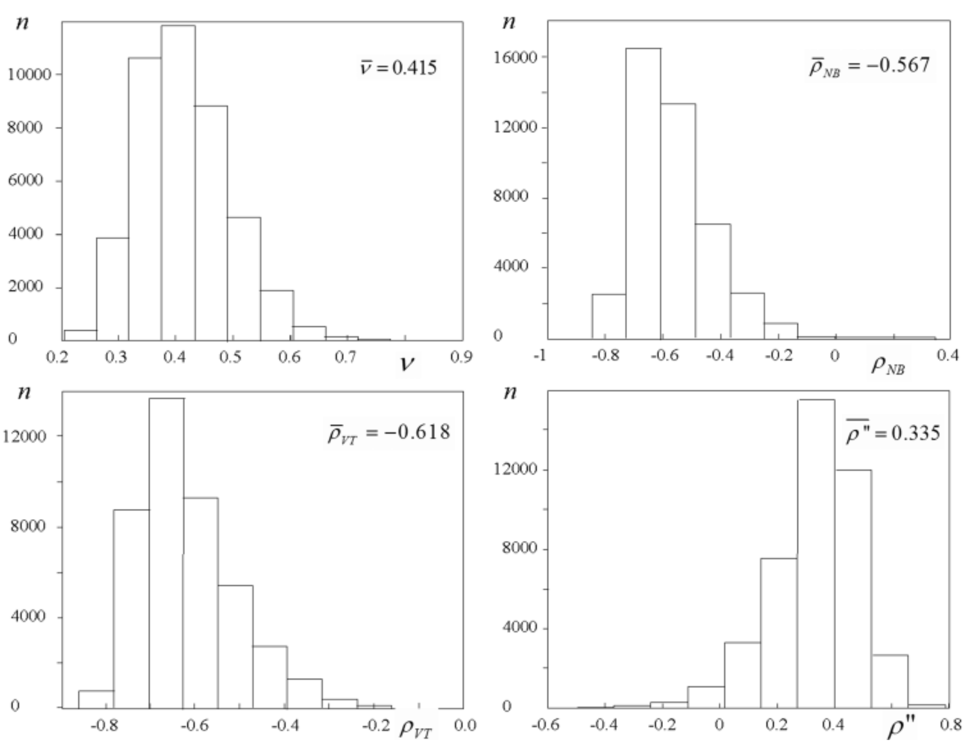

Figure 9. The histograms of the parameters observed in this study (buoys only) that are relevant for the various theories (n is the number of observations in each interval of parameter values).

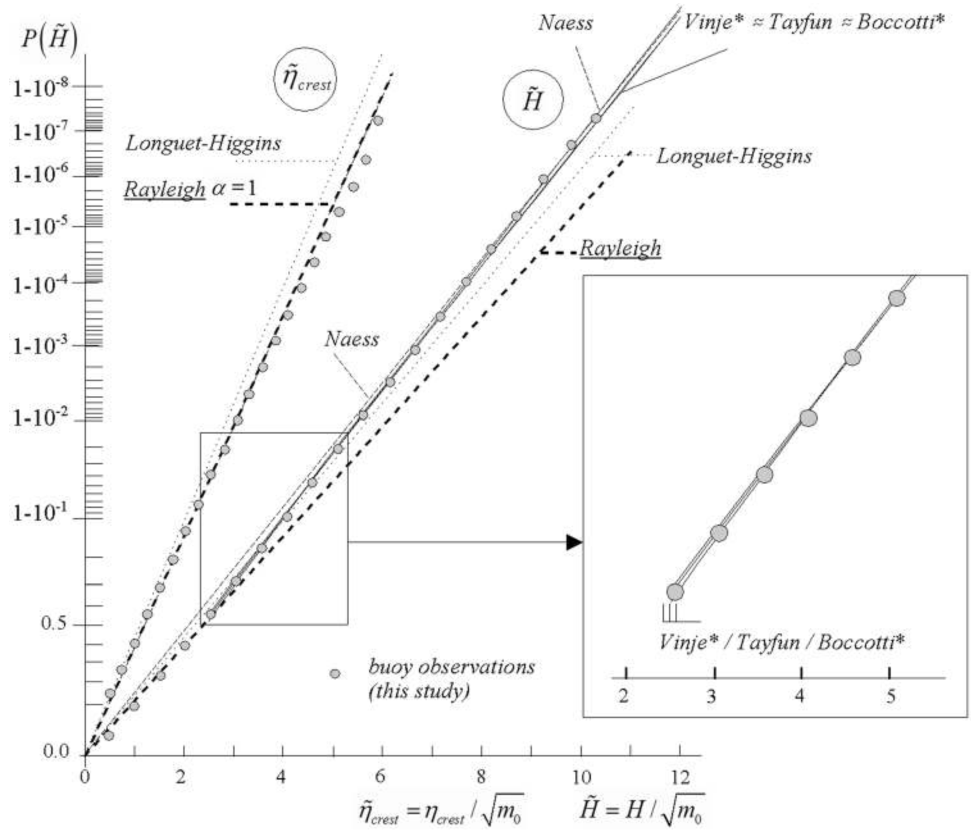

Figure 10. The observed normalized crest height and wave height distribution (same as in Figure 8) and the theoretical distributions for the average observed parameters of Figure 9. The distributions of *Tayfun* [1990], *Vinje** [1989] and *Boccotti** [1989, 2000] are shown for $\tilde{H} > 2.5$ and enlarged in the box for $2.5 < \tilde{H} < 5.0$. Only data points nearest to multiples of $1/2\sqrt{m_0}$ or $1/4\sqrt{m_0}$ are shown. The highest normalized observed crest height and wave height are added (top value in each distribution) to show the full range of the observations.

identical to the value $v = 0.416$ for a Pierson-Moskowitz spectrum [*Pierson and Moskowitz*, 1964]. The corresponding theoretical distributions are compared with the observed distributions in Figure 10. The crest height distribution of *Longuet-Higgins* [1980] underpredicts the normalized observed crest heights by ~9.5%. The three theoretical distributions of the normalized wave height of *Vinje** [1989], *Boccotti** [1989, 2000], and *Tayfun* [1990] agree very well with the observations for $\tilde{H} > 2.5$ (e.g., maximum 1.5% difference). In fact, they reproduce the S-shape of the observed distribution very well (i.e., show the proper S-curve of Figure 8 in this range; the distribution of *Tayfun* [1990] slightly better for low values of $\tilde{H}$ and the distribution of *Vinje** [1989] slightly better for high values of $\tilde{H}$). With the close agreement between these three theoretical distributions and the Weibull distribution of *Forristall* [1978] on the one hand, and the observations on the other, it is no surprise that these four distributions are almost identical to each other, at least for these observations (differences between the Weibull distribution and the distribution of *Tayfun* [1990] is <1.5% over the range of observations and $\tilde{H} > 2.5$). The distribution of *Naess* [1985] underestimates the normalized observed wave heights by 2 – 10% for $\tilde{H} < 8$ but is as good as the other three theoretical distributions for $\tilde{H} > 8$.

4.2. Maximum Wave Height and Crest Height

[37] In this section, we consider the maximum normalized crest height and maximum normalized wave height in a sequence of N consecutive individual normalized waves. More precisely, we consider the average of these maxima from n such sequences, each obtained by concatenating normalized observed records $\eta(t)/\sqrt{m_0}$ of nominal duration. The results for this and for the Weibull distribution and the best fit scaled Rayleigh distributions are given in Figure 11. The 95% confidence interval of the average is estimated as for a Gaussian distribution with standard deviation [*Cartwright*, 1958] $\sigma_{\eta_{\text{crest, max}}} = \sqrt{\text{var}_{\eta_{\text{crest, max}}}/n}$ in which

$$\text{var}_{\eta_{\text{crest, max}}} = E\left(\eta^2_{\text{crest, max}}\right) - E^2\left(\eta_{\text{crest, max}}\right) \approx \frac{1}{2\ln(N)}\left(1.6449 - \frac{2.1515}{\ln(N)}\right). \quad (18)$$

The standard deviation for the average maximum wave height would be twice as high (assuming $H \approx 2\eta_{\text{crest}}$).

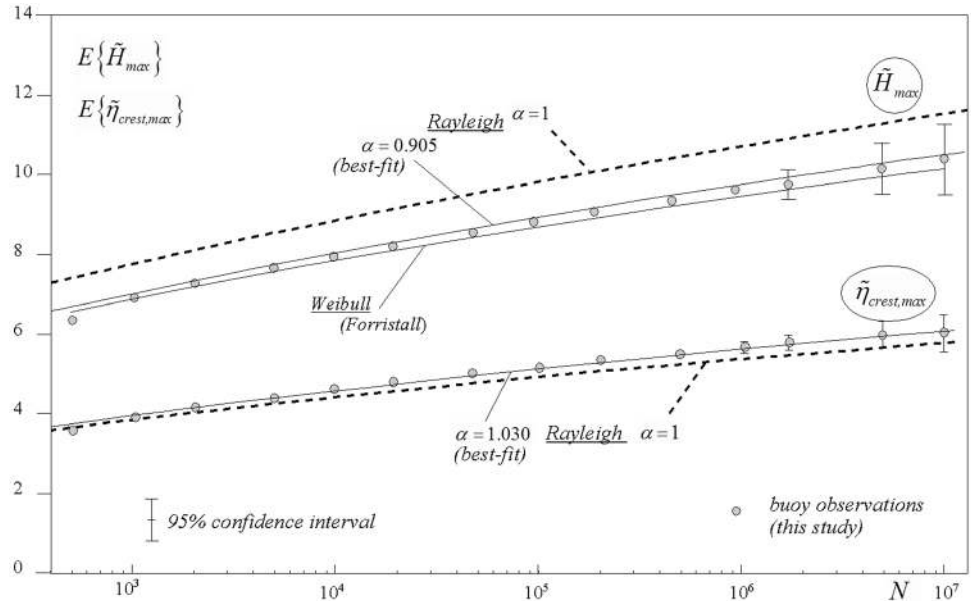

Figure 11. The average maximum wave height and crest height (normalized) observed by the buoys as a function of the number N waves in a sequence and the predictions from the best fit scaled Rayleigh distributions and the (Forristall) Weibull distribution (see Figure 8).

[38] The conventional Rayleigh approach (α = 1) for predicting the expected maximum normalized *crest* heights (lower set of data in Figure 11) does not agree well with the observations for $N > 2000$: the deviations are as large as 5% (underprediction). The same conventional Rayleigh approach (α = 1) also does not properly predict the normalized maximum wave heights (top observations in Figure 11). But using α = 0.905 of the above best fit scaled Rayleigh distribution gives very reasonable predictions (deviations <2%, except near N = 500, where the average normalized observed maximum wave height is 5% lower than predicted). The predictions by the best fit Weibull distribution are just as good (but the errors are of opposite sign).

[39] The theoretical predictions using the above observed average parameter values $\overline{\nu}$ = 0.415, $\overline{\rho}_{NB}$ = −0.567, $\overline{\rho}_{VT}$ = −0.618, and $\overline{\rho''}$ = 0.335 are shown in Figure 12. The predictions of the normalized maximum wave heights (top observations in Figure 12) by *Longuet-Higgins* [1980] are obviously systematically slightly too high (5%, except at $N \approx 500$, where it is 10%). The agreement for the other theories is almost perfect, with a maximum error of 1%

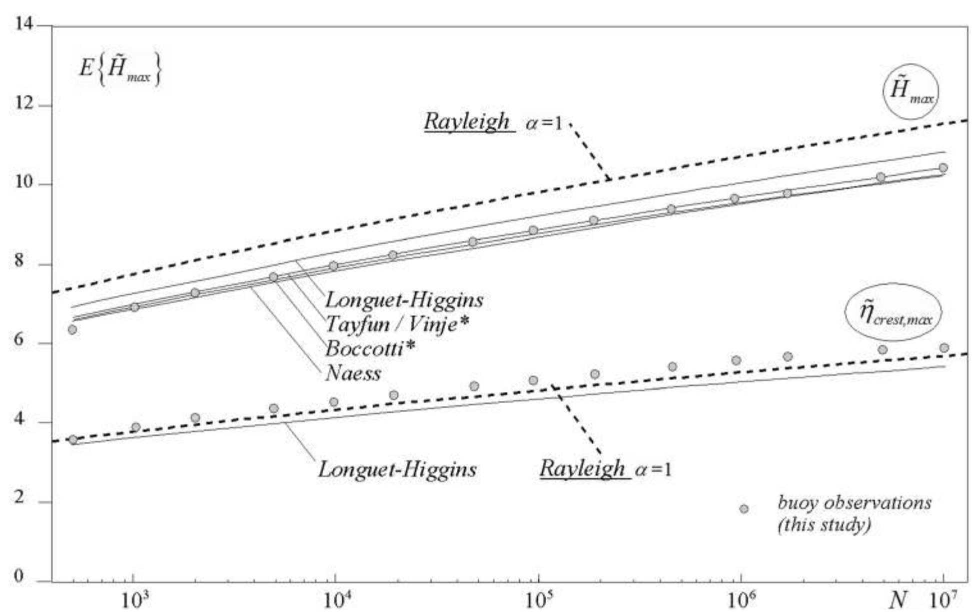

Figure 12. The average maximum crest height and wave height (normalized) observed by the buoys as a function of the number N of waves in a sequence and the theoretical predictions.

Table 4. The Relevant Parameters and the Scaling of the Significant Wave Height for the Various Approaches for the Observations of the Present Study[a]

	Parameter	α	β	$H_{1/3,\mathrm{theory}}/H_{1/3,\mathrm{Rayleigh}}$	$H_{1/3,\mathrm{theory}}/H_{1/3,\mathrm{obs}}$
Rayleigh [*Longuet-Higgins*, 1952]	–	–	–	1	1.042
Longuet-Higgins [1980]	$\overline{\nu} = 0.415$	0.935	1	0.935	0.974
Naess [1985]	$\overline{\rho}_{NB} = -0.567$	0.885	1	0.885	0.922
Vinje [1989][b]	$\overline{\rho}_{VT} = -0.618$	0.899	1.144	0.930	0.969
Tayfun [1990]	$\overline{\rho}_{VT} = -0.618$	0.885	1.144	0.936	0.975
Boccotti [1989, 2000][b]	$\overline{\rho}_{NB} = -0.567$, $\rho'' = 0.335$	0.899	1.303	0.944	1.000
Forristall [1978]	$a = 2.126$, $b = 8.42$	–	–	0.942	0.982
Observed time series (this study)				$H_{1/3,\mathrm{obs}}/H_{1/3,\mathrm{Rayleigh}}$, 0.957	$H_{1/3,\mathrm{obs}}/H_{1/3,\mathrm{obs}}$, 1

[a]All correlations taken from the observed correlation functions (obtained as the fast Fourier transforms of the observed spectrum multiplied with f^2 when required) at $\tau = 1/2T_{m01}$. The value for $H_{1/3,\mathrm{theory}}/H_{1/3,\mathrm{Rayleigh}} = 1/1.045 = 0.957$ has been taken from Figure 13 (top left).
[b]With modified correlation coefficients.

(overprediction) for the approaches of *Boccotti** [1989, 2000] and *Naess* [1985] near $N = 10^6$ and virtually no errors (<1%) for *Vinje** [1989] and *Tayfun* [1990], except near $N =$ 500. In contrast to the predictions of the normalized maximum wave heights, the approach of *Longuet-Higgins* [1980] underpredicts the normalized maximum crest heights by as much as 7% at $N = 10^7$ (the other theories do not apply).

4.3. The Significant Wave Height

[40] The scale of the wave heights is usually represented by the significant wave height, which can be estimated either directly from the wave record $H_{1/3}$ or with one of the above theories from the spectrum or with the empirical Weibull distribution. With the above estimated Rayleigh, Rayleigh-like, and Weibull distributions (see Table 4), we computed these estimates (analytically when possible: Rayleigh; Weibull; *Longuet-Higgins* [1980]; *Naess* [1985]; *Tayfun* [1990, equations (48), (49) and (52)]; and numerically with the required moments of the distributions when not analytically possible: *Vinje** [1989] and *Boccotti** [1984, 2000]) and from the observed distribution of Figure 8 (for which the best fit scaled Rayleigh distribution gives $\tilde{H}_{s,\ \mathrm{obs}} = 3.62$). The results are given in Table 4, from which it is obvious that the estimate of the significant wave height by the conventional Rayleigh distribution is too high (by 4%). The estimate with *Naess* [1985] is too low by 7.8%, whereas the estimate with *Boccotti** [1989, 2000] is nearly perfect.

[41] To inspect the variability of the ratio $H_{1/3,\ \mathrm{theory}}/H_{1/3,\ \mathrm{obs}}$, we also computed this ratio for each individual wave record separately (only for the analytical estimates, i.e., without *Vinje** [1989] and *Boccotti** [1989, 2000]), with the observed values of the significant wave height corrected for the sampling-rate bias with the expressions of *Tayfun* [1993] given in section 2.1 (i.e., with the mean wave period and sample rate for each individual record). The scatter diagrams of these values are shown in Figure 13. These results are reasonably consistent with the results of Table 4 in the sense that the slope of the least squares best fit linear relationship is reasonably close to the above values of $H_{1/3,\ \mathrm{theory}}/H_{1/3,\ \mathrm{obs}}$. They also show that the theory of *Tayfun* [1980] gives the best estimate of the significant wave height (slope closest to unity, highest correlation, and smallest scatter index; see figure legend for definitions).

4.4. The Dependency on Correlation

[42] The above theories indicate that the statistical characteristics addressed depend on the correlation ρ between crest height and trough depth (such correlation is strongly related to the width of the spectrum, implying that the statistics are equally dependent on ν). Such dependency should then be evident as stratification in the above results. To verify this, we ranked all records in ascending order of correlation (we used ρ_{VT}) and divided them into five groups (each containing 20% of the total number of records). The result for the observed distribution functions for the normalized wave height is shown in Figure 14, where the labels $\overline{\rho}_{10}$, $\overline{\rho}_{30}$, $\overline{\rho}_{50}$, $\overline{\rho}_{70}$, and $\overline{\rho}_{90}$ represent the groups in ascending order of (absolute) correlation. It is obvious that the observed values are well organized in the sense that at each value of $\tilde{H}$ for $\tilde{H} > 1.5$, the observed wave heights increase monotonically with increasing (absolute) correlation (the only exceptions being the highest normalized wave height in the group $\overline{\rho}_{10} = -0.447$, which has been labeled as an outlier in Figure 14, and the highest normalized wave in the group $\overline{\rho}_{90} = -0.743$). As expected, the group with the highest (absolute) correlation $\overline{\rho}_{90} = -0.743$ corresponds to the highest normalized wave heights (at the same level of probability) and is closest to the conventional Rayleigh distribution. For $\tilde{H} < 1.5$, the stratification is similarly monotonic but with an opposite trend (the lines in Figure 14 cross near $\tilde{H} = 1.5$), which is also consistent with the group with highest (absolute) correlation being closest to the Rayleigh distribution, because the Rayleigh distribution crosses the observations near $\tilde{H} = 1.5$. To verify that the linear theories can reproduce this variation, we compare the normalized observations with the predictions of *Tayfun* [1990], which combines the availability of analytical expressions with overall good agreement with our observations] for the groups with the highest and the lowest (absolute) correlation in Figure 14. This comparison shows that the variation is very well predicted (see Figure 14). We did the same analysis for the average normalized observed maximum wave height (Figure 15) with basically the same result, except that the monotonic character of the stratification is not fully maintained above $N = 7 \times 10^4$.

5. Altimeter Observations

[43] We mentioned in the introduction that we do not consider nonlinear theories, because our observations were carried out with buoys, which by their nature do not properly represent the nonlinear character of the waves. However, to cursorily inspect the effect of nonlinearities, we

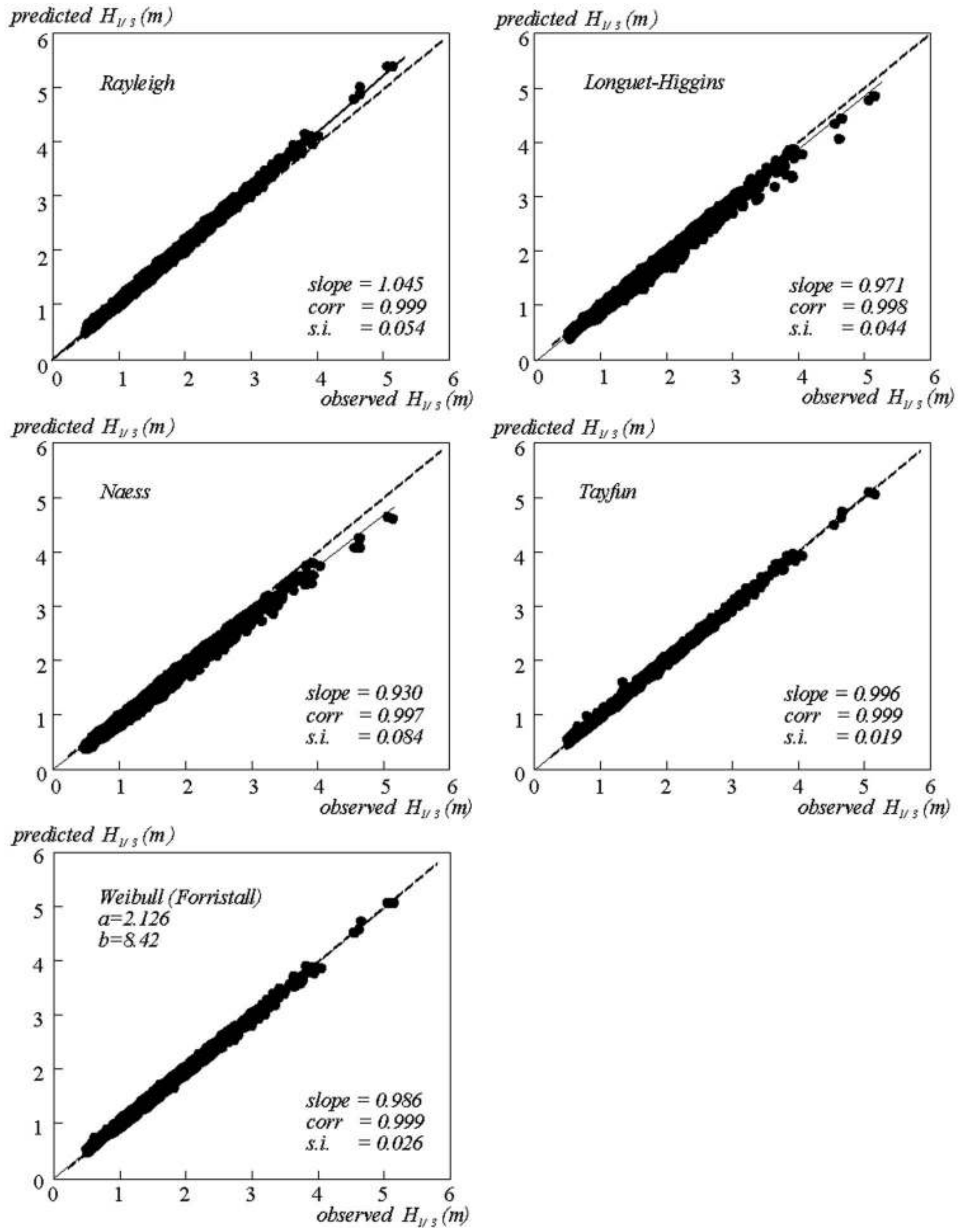

Figure 13. The significant wave height estimated directly from the observed time records and predicted by the various approaches (using the spectral width and the correlation from the time records; analytical predictions only). The solid line is the best fit line, and the dashed line shows perfect agreement. (slope) Slope of best fit line; (corr) correlation coefficient of data; (s.i.) scatter index, or rms-error normalized with average of observations.

compared the results of our buoy observations with similar results obtained with the (fixed) laser altimeters described earlier. The analysis of the altimeter data is identical to the above analysis of the buoy data. The results are compared with those of the buoys in Figures 16 and 17 over the range of the altimeter observations (which is much shorter than that of the buoy observations). It must be noted that these two data sets are taken in very different circumstances. Differences may therefore be due not only to different instruments but also to different physical conditions. The basis for the comparison is that both data sets represent the normalized surface elevation of wind-generated waves in deep water.

[44] Given this caveat, it is remarkable that the distribution of the normalized wave heights as observed using the altimeters is almost identical to that observed by the buoys. It shares the same gentle S-curve (see Figure 8) with the buoy data (except for the highest altimeter data, which seem to be mildly anomalous, i.e., at $\tilde{H} \approx 7$). This implies that the two-parameter Weibull distribution and the distributions of *Vinje** [1989], *Boccotti** [1989, 2000], and *Tayfun* [1990], which fitted the buoy observations very well, provide an

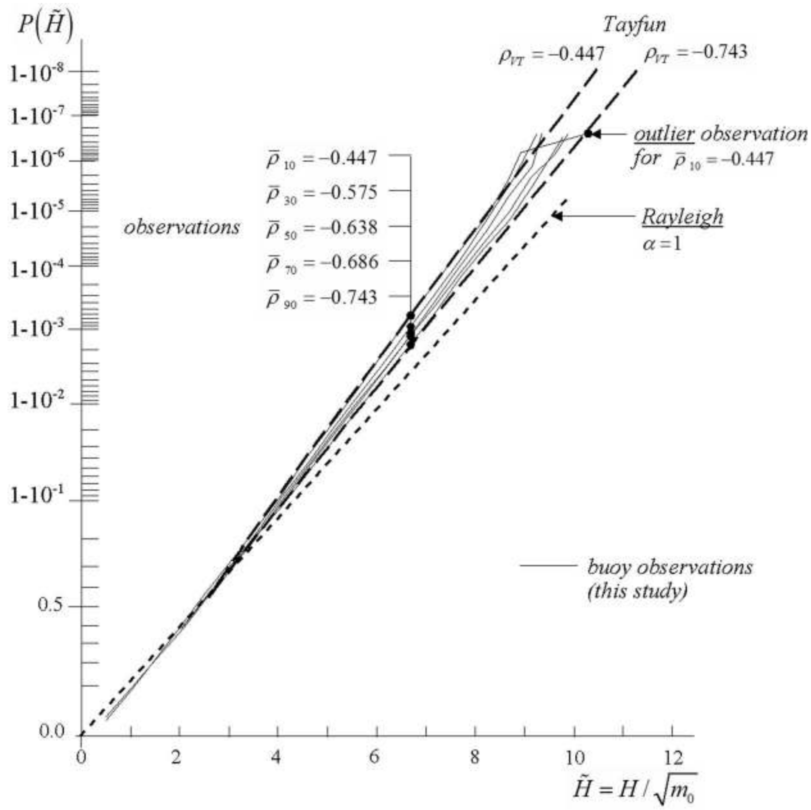

Figure 14. The distributions of the normalized wave height as observed by the buoys in the present study, for five groups of correlation and their variation theoretically predicted using the method of *Tayfun* [1990] over the range of validity ($\tilde{H} > 2.5$). One outlier is labeled as such. Highest normalized observed wave height in each group is added (top value in each distribution) to show the full range of the observations.

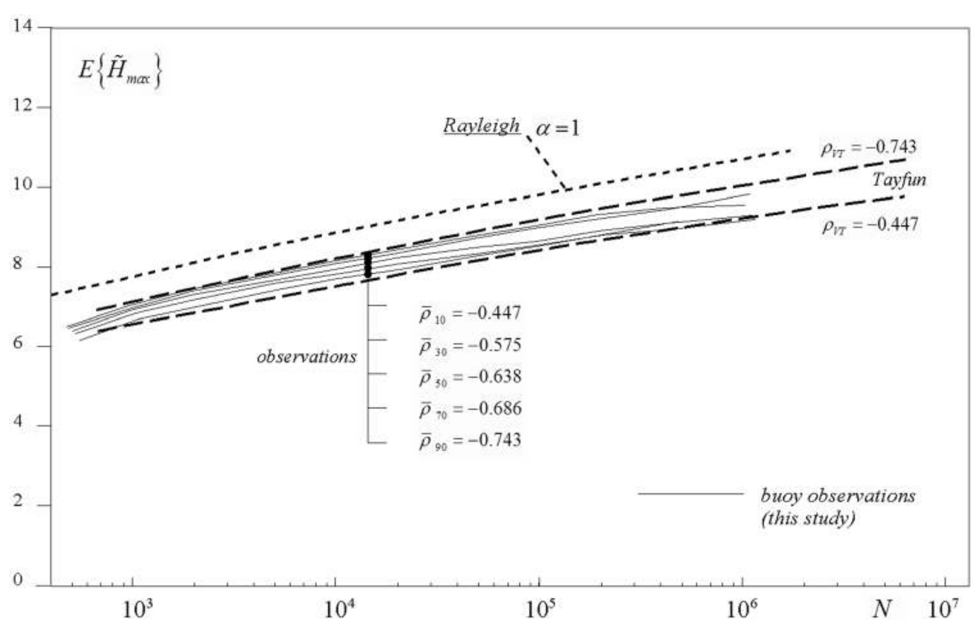

Figure 15. The average values of the maximum wave height (normalized) observed by the buoys in the present study, for five groups of correlation and the theoretical predictions of *Tayfun* [1990].

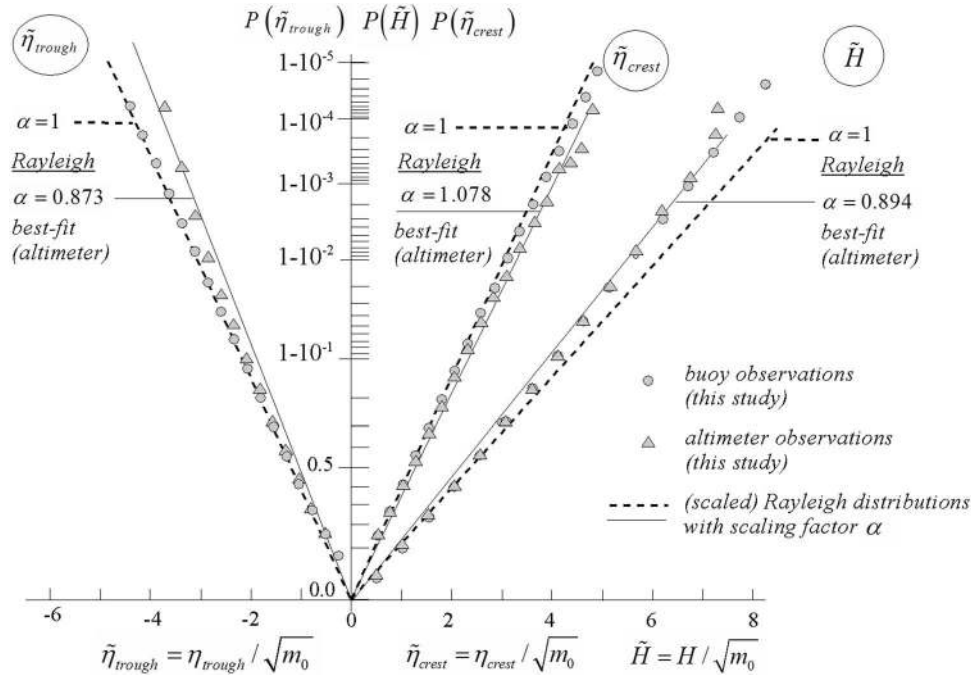

Figure 16. The distributions of the normalized wave height and crest height as observed by the buoys and the altimeters in the present study (over the range of the altimeter observations) and the Rayleigh distributions that best fit the altimeter observations (compare with Figure 8). Only data points nearest to multiples of $1/2\sqrt{m_0}$ or $1/4\sqrt{m_0}$ are shown. Highest observed crest height and wave height in each laser altimeter distribution added (top value in each distribution) to show the full range of the observations.

equally good fit to the altimeter observations. The coincidence of the two observed data sets indicates that the distribution of the wave heights does not seem to be severely affected by the type of measuring instrument used: either a moving buoy or a fixed altimeter. Over all, the conventional Rayleigh distribution (α = 1) overpredicts these normalized altimeter wave height observations by 12%, i.e., the best fit scaled Rayleigh distribution for these observations gives α = 0.894 (the inverse of which is 1.12). This is marginally more than in the buoy data (where α = 0.905 or 10% overprediction). Without the one highest normalized altimeter data point in Figure 16, the overprediction is slightly less

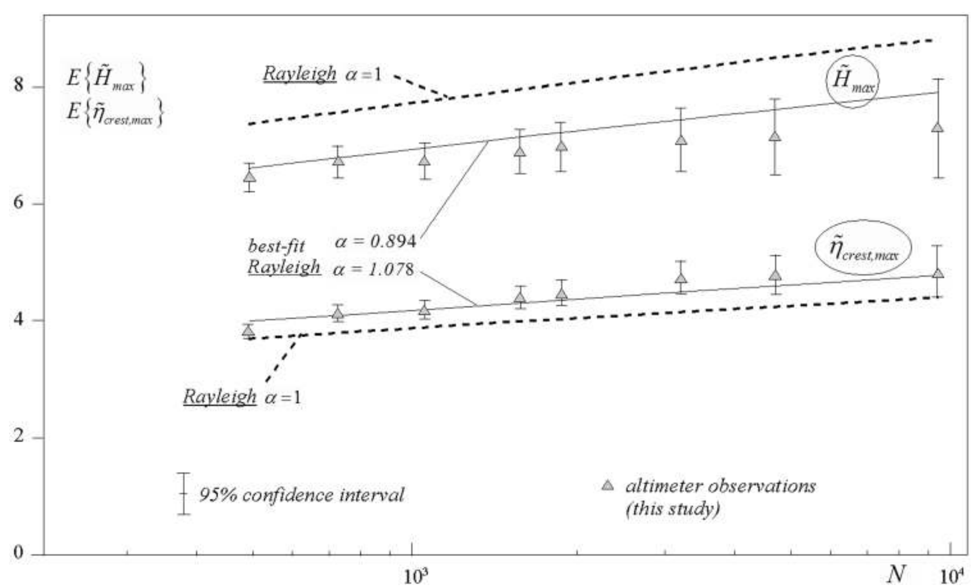

Figure 17. The average values of the altimeter observed maximum wave height and crest height (normalized) as a function of the number of waves in a sequence. Vertical bars indicate 95% confidence interval of observed average value.

(α = 0.908). The normalized crest heights observed by the altimeters are generally slightly higher than those observed by the buoys, with α = 1.078 for the best fit scaled Rayleigh distribution. The normalized trough depths observed by the altimeters are overall significantly shallower than observed by the buoys, with α = 0.873 for the best fit scaled Rayleigh distribution (versus α = 0.995 for the buoy observations). This difference in scaling confirms the idea that a laser altimeter observes the nonlinear character of the waves better than a buoy does. The good agreement of the buoy observations with the Rayleigh distribution is therefore not valid for the altimeter observations.

[45] The only linear theory that predicts a (scaled Rayleigh) distribution for the crest heights is the theory of *Longuet-Higgins* [1980]. For the normalized altimeter observations, for which ν = 0.485, the corresponding theoretical scaling parameter would be α = 0.910, considerably lower than the observed value α = 1.078. This further supports the idea that the altimeter observations show nonlinear effects in the crests heights, which are obviously not included in any linear theory.

[46] The average maximum normalized crest height and wave height in a sequence of N waves observed by the altimeters are given in Figure 17. These values are reasonably consistent with the scaling of the distributions as shown by the solid lines, but the deviation is somewhat larger for the maximum wave heights for large sequences (N > 3000, say), still within the 95% confidence interval.

6. Discussion and Conclusions

[47] In this study, we have analysed the short-term statistics of 10 million individual deep-water wind-generated waves in open sea by normalizing and concatenating >40,000 wave records obtained during a period of 15 years with four buoys located along the Catalan coast of Spain on the Mediterranean Sea. We could thus inspect the statistical characteristics of these waves over a range of normalized wave heights of $0 < \tilde{H} < 10$ (wave heights defined with downward zero crossings of the surface elevation). The main limitation of these observations is that buoys tend to "linearize" waves, i.e., they tend to suppress the nonlinear character of the waves. To inspect this effect, we supplemented our data with some 10,000 individual waves observed with laser altimeters in the North Sea. The range of these observations is $0 < \tilde{H} < 7$. We upscaled the observed normalized crest heights, trough depths, and wave heights with 3% (approximately) to account for a sampling-rate bias.

[48] The normalized crest heights observed by the buoys are typically 3% higher than predicted with the conventional Rayleigh distribution (scaling parameter α = 1), whereas the altimeter observations are even higher: typically 5% higher than observed with the buoys. The scaled Rayleigh distribution of *Longuet-Higgins* [1980] predicts lower crest heights and thereby underestimates the normalized crest heights observed by the buoys by ~9.5%. The normalized trough depths observed by the buoys are very well predicted by the conventional Rayleigh distribution (scaling parameter α = 1; 0.6% underestimation). This is not the case for the altimeter observations, for which the normalized trough depths are 12% shallower than observed with the buoys. These differences between the results of the buoys and the laser altimeters supports the idea that a buoy tends to linearize the waves.

[49] The differences between the buoy-observed and predicted maximum normalized crest heights are consistent with these findings (i.e., the expected value of the maximum normalized crest height in a sequence of many waves).

[50] The distribution of the normalized wave heights observed by the buoys is very well approximated over the entire range of observations by a two-parameter Weibull distribution with the parameter values suggested by *Forristall* [1978], the maximum difference from the normalized observations being ~1.5%. This is also the case for the altimeter observations, for which the distribution of normalized observed wave heights is almost identical to the distribution of the buoy-observed normalized wave heights over the range of these observations $0 < \tilde{H} < 7$ (except for one or two mildly anomalous altimeter data points). This supports the conclusion of earlier studies [e.g., *Tayfun*, 1983; *Vinje*, 1989; *Massel*, 1996; *Tayfun and Fedele*, 2007] that nonlinear effects do not seem to affect the statistics of individual wave heights significantly (but they do affect the statistics of the crest heights). This conclusion contradicts the conclusion of theories that try to predict the occurrence of freak waves [e.g., *Mori and Yasuda*, 2002; *Janssen*, 2003; *Mori and Janssen*, 2006; *Onorato et al.*, 2009], but these conclusions are being adapted (see *Dysthe et al.* [2008], *Socquet-Juglard et al.* [2005], *Onorato et al.* [2009], *Gramstad and Trulsen* [2007], *Waseda* [2006], *Onorato et al.* [2009], and other nonlinear theories [e.g., *Longuet-Higgins*, 1963]). The superior fit of the Weibull distribution compared to that of the scaled Rayleigh distribution is an empirical finding without theoretical substantiation and consequently without theoretical predictions of its coefficients. However, the close agreement between the results obtained with different types of instruments (buoys, wave staffs, and laser altimeters) and different physical conditions (hurricanes in the Gulf of Mexico, all weather off the Catalan coast, and storms in the central North Sea) suggest a reasonably general applicability of this Weibull distribution. The conventional Rayleigh distribution does not fit the normalized wave height distributions observed in the present study nearly as well: neither the buoy observations nor the altimeter observations. It overpredicts the normalized wave heights by ~10% and the observed significant wave height by 4%. The latter finding would be close to the 7.5% and 7.3% of *Longuet-Higgins* [1980] and *Holthuijsen* [2007], respectively, if we had not corrected our observations for the sample rate bias (bringing our value from 4% to 6.9%) or vice versa. The Rayleigh-like distributions of *Boccotti** [1989, 2000], *Vinje** [1989], and *Tayfun* [1990] are almost identical to the above Weibull distribution and agree very well (maximum 1.5% difference) with the buoy observations (and therefore also the altimeter observations) for $\tilde{H} > 2.5$ (below this range, the theories are not valid). The scaled Rayleigh distributions of *Longuet-Higgins* [1980] and *Naess* [1985] do not provide such a good fit. The first distribution underpredicts the lower normalized wave heights (by as much as 8% near $\tilde{H} = 2.5$), and it overpredicts the higher normalized wave heights by as much as 5% near $\tilde{H} = 10$. The second distribution underpredicts the normalized wave heights by ~5% or more for

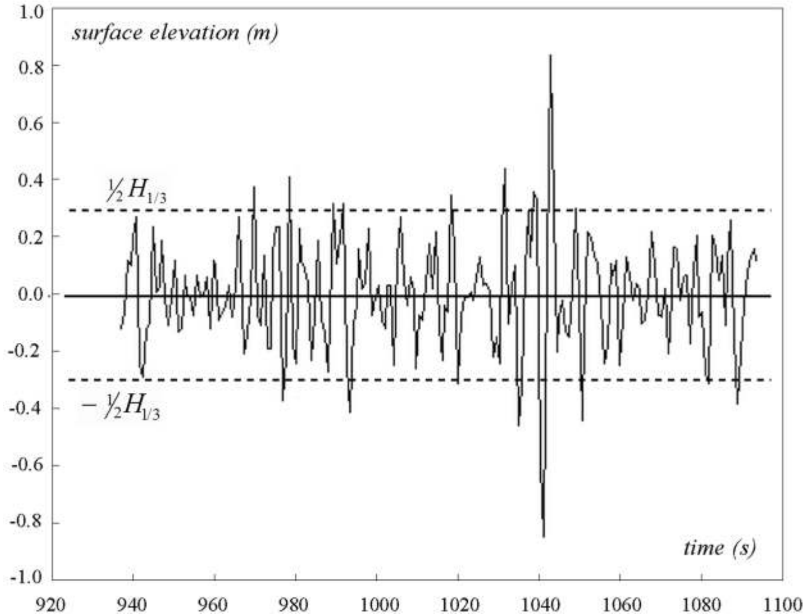

Figure 18. The one wave in the buoy observations that may qualify as a freak wave in the sense that it seems to be a statistical outlier.

$\tilde{H} < 4$. The conclusion for this aspect of our study is therefore that the Weibull distribution and all theoretical distributions, except the ones of *Longuet-Higgins* [1980] and *Naess* [1985], fit the observations very well. The theoretically predicted dependency of the buoy-observed normalized wave height distribution on the (de-)correlation between crest height and trough depth has been confirmed by stratifying our data. This dependency is very well predicted by the theory of *Tayfun* [1990] and probably also by the theories of *Boccotti** [1989, 2000] and *Vinje** [1989]. These theories are therefore more generally applicable than the Weibull distribution of *Forristall* [1978].

[51] As a final comment, we want to add that having 10 million individual waves at our disposal gives us the opportunity to try and identify freak waves in our data. We were careful to formulate our criteria for the initial selection such that we would not filter out very large waves. However, we did reject complete records on the basis of these criteria, and such a record may have contained an extremely large wave. In addition, as mentioned in section 1, a buoy may swerve around a steep wave crest. It may thus inherently fail to observe a freak wave (this argument obviously does not apply to the laser altimeters). The following comments are therefore qualified.

[52] Unfortunately, the term *freak wave* is not well defined. Any definition in terms of somewhat higher than twice the significant wave height, i.e., a freak wave is a wave with a height $H > \beta H_{m0}$ (where β is a constant $2 \leq \beta \leq 2.5$ [e.g., *Onorato et al.*, 2002; *Kharif and Pelinovsky*, 2003; *Mori and Janssen*, 2006; *Liu and MacHutchon*, 2006; *Dysthe et al.*, 2008]) implies that we found thousands of such waves, hardly the rare event that a freak wave is supposed to be. An alternative definition and one that we favor is that a freak wave is an unexpected wave the height of which is a statistical high outlier, i.e., inconsistent with the statistics of the ambient sea state. We did not carry out a formal procedure to detect such outliers, but a visual inspection of the observed distribution of the normalized wave heights (the data underlying Figure 8, 14, and 16) and of the normalized extreme wave heights (the data underlying Figures 11, 15, and 17) showed one obvious outlier (identified as such in Figure 14, the highest normalized wave in the group with lowest crest-trough correlation; it is the highest normalized wave in our buoy observations). In an absolute sense, it is a rather low wave, only 1.69 m, but compared to the significant wave height of this record, $H_{1/3} = 0.59$ *m* and $H_{m0} = 0.67$ *m*, it is relatively high, with $H = 10.05\sqrt{m_0}$ (or $H = 10.35\sqrt{m_0}$ when corrected for the sample rate bias). This wave may therefore qualify as a freak wave (shown in Figure 18). However, it must be noted that (1) this wave is an outlier only in a selected group of wave records (lowest crest-trough correlation), and (2) this wave seems to fit the statistics of the complete data set of the buoys perfectly well (the highest normalized wave in Figure 8). The latter comment nicely illustrates the statement of *Forristall* [2005, p. 34]: "An unusually large wave will always stand out as a rogue wave in a short record. Yet it may fit standard statistics perfectly well if the statistics from many hours are combined." We do not want to suggest that freak waves do not occur. To the contrary, sufficient instrument observations and eyewitness accounts exist to support the contention that they do occur, not only in computer simulations or in the laboratory but also at open sea [e.g., *Draper*, 1965; *Mallory*, 1974; *Smith*, 1976; *Atkins*, 1977; *Haver and Andersen*, 2000; *Buckley*, 2005; *Holthuijsen*, 2007].

Appendix A: Benjamin-Feir Index and Large Waves

[53] The definition of the Benjamin-Feir Index (BFI; *Benjamin and Feir* [1967]) is basically the ratio of wave steepness to randomness of the waves [*Mori and Janssen*,

2006]. If the steepness of the waves is expressed as $\varepsilon = k_{mean}\,\sigma_\eta$ (the mean wave number $k_{mean} = 2\pi/L_{mean}$, where L_{mean} is the mean wave length and σ_η is the standard deviation of the random surface elevation) and the degree of randomness μ is expressed in terms of the peakedness Q_p of the wave spectrum $E(\omega)$ (the inverse of spectral width relative to mean frequency [*Goda*, 1976]), $\mu^{-1} = Q_p\sqrt{\pi} = 2\sqrt{\pi}\int_0^{+\infty}\omega E^2(\omega)\,d\omega/(\int_0^{+\infty}E(\omega)d\omega)^2$, then the Benjamin-Feir Index may be defined as BFI = $\varepsilon\sqrt{2}/\mu$. To estimate the probability of encountering a large wave height in a storm, we define a reference storm as a stationary wave condition with a typical number of waves (N = 2000) and a typical value for the BFI(= 0.8). To arrive at these numbers, consider a reasonable duration of a storm of 6 hours and a mean wave period of 10 s. The number of waves then is $N \approx 2000$. The outcome of the following analysis is fairly insensitive to the precise value, e.g., N = 4000 gives $n_{0.01} = 2.90$ rather than $n_{0.01} = 2.83$. To arrive at a typical value for the BFI, consider a typical spectrum and a typical wave steepness in a storm. The spectrum is the well-established JONSWAP spectrum [*Hasselmann et al.*, 1973]. The value of the peakedness parameter for this spectrum is $Q_{p,\mathrm{JONSWAP}} = 3.1437$ (peak enhancement factor $\gamma = 3.3$, spectral width parameters $\sigma_a = 0.07$ and $\sigma_b = 0.09$). The wave steepness in a storm is typically [*Kahma and Calkoen*, 1992] $\varepsilon = 0.1$, so that BFI = 0.8 in the reference storm. The probability that an arbitrarily chosen (normalized) wave height $\underline{n} = \underline{H}/H_s$ in the reference storm is lower than n is 1 - $P(n)$, where, in the present approach, $P(n)$ is the Rayleigh-Edgeworth distribution [*Mori and Janssen*, 2006] $P(n) = e^{-2n^2}[1 + 2\pi n^2(n^2 - 1)\,\mathrm{BFI}^2/(3\sqrt{3})]$. The probability that the maximum wave height in the storm (i.e., of N statistically independent waves) is higher than n is $1 - [1 - P(n)]^N$. For the reference storm, with N = 2000 and BFI = 0.8, this probability is 0.01 for $n = n_{0.01} = 2.83$. In other words, in the context of the theory of *Mori and Janssen* [2006], the maximum wave height per (reference) storm exceeds the value $2.83H_s$ once in 100 (reference) storms (on average). It must be emphasized that the theory utilized here applies only to unidirectional waves. Any short-crestedness in the waves would greatly reduce the effects mentioned [e.g., *Waseda*, 2006], so that the above represents an upper limit.

[54] **Acknowledgments.** We are grateful to the Xarxa d'Instruments Oceanogràfics i Meteorològics de la Generalitat de Catalunya and the participants in the WADIC experiment for their permission to use their data. We also thank Joan Pau Sierra Pedrico of the Technical University of Catalonia (UPC), Jesús Gómez Aguar of UPC, and Steven Barstow of FUGRO-OCEANOR for their kind and very helpful assistance in accessing the data. We gratefully acknowledge the comments of the reviewers of our paper, in particular regarding the effect of the sampling-rate error.

References

Allender, J., T. Audunson, S. F. Barstow, S. Bjerken, H. E. Krogstad, P. Steinbakke, L. Vartdal, L. E. Borgman, and C. Graham (1989), The WADIC project: A comprehensive field evaluation of directional wave instrumentation, *Ocean Eng.*, *16*(5/6), 505–536.

Atkins, J. E. (1977, January), Special reports on freak waves, *Mar. Obs.*, *47*(255), 32–35.

Benjamin, T. B., and J. E. Feir (1967), The disintegration of wavetrains on deep water, Part 1: Theory, *J. Fluid Mech.*, *27*, 417–430.

Boccotti, P. (1989), On mechanics of irregular gravity waves, *Atti Accademia Nazionale dei Lincei*, Memorie VIII, *19*, 111–170.

Boccotti, P. (2000), *Wave Mechanics for Ocean Engineering*, 496 pp., Elsevier, Amsterdam.

Bolaños, R., G. Jorda, J. Cateura, J. Lopez, J. Puigdefabregas, J. Gomez, and M. Espino (2009), The XIOM: 20 years of regional coastal observatory in the Spanish Catalan coast, *J. Mar. Syst.*, *77*(3), 237–260, doi:10.1016/j.jmarsys.2007.12.018.

Buckley, W. H. (2005), Extreme waves for ship and offshore platform design, *Tech. Res. Rep. 57*, 65 pp, Soc. of Nav. Architects and Maritime Eng., Jersey City, N. J.

Cartwright, D. E. (1958), On estimating the mean energy of sea waves from the highest waves in a record, *Proc. R. Soc. London, Ser. A*, *247*, 22–48.

Cartwright, D. E., and M. S. Longuet-Higgins (1956), The statistical distribution of the maxima of a random function, *Proc. R. Soc. London, Ser. A*, *237*, 212–232.

Draper, L. (1965), "Freak" ocean waves, *Mar. Obs.*, *35*, 193–195.

Dysthe, K., H. E. Krogstad, and P. Müller (2008), Oceanic rogue waves, *Annu. Rev. Fluid Mech.*, *40*, 287–310.

Forristall, G. Z. (1978), On the statistical distribution of wave heights in a storm, *J. Geophys. Res.*, *83*(C5), 2353–2358, doi:10.1029/JC083iC05p02353.

Forristall, G. Z. (1984), The distribution of measured and simulated wave heights as a function of spectral shape, *J. Geophys. Res.*, *89*(C6), 10,547–10,552, doi:10.1029/JC089iC06p10547.

Forristall, G. Z. (2005), Understanding rogue waves: Are new physics really necessary?, in *Rogue Waves, Proc. 14th Aha Huliko: A Hawaiian Winter Workshop*, pp. 29–35, University of Hawaii, Honolulu, HI.

Goda, Y. (1976), On wave groups, in *Proc. First International Conference on Behaviour of Offshore Structures BOSS'76*, vol. 1, pp. 115–128, Norw. Inst. of Technology, Trondheim.

Goda, Y. (1988), On the methodology of selecting design wave height, in *Proc. 21st International Conference on Coastal Engineering*, *Malaga*, *1*, 899–913, Am. Soc. Civ. Eng., New York.

Gramstad, O., and K. Trulsen (2007), Influence of crest and group length on the occurrence of freak waves, *J. Fluid Mech.*, *582*, 463–472.

Haring, R. E., A. R. Osborne, and L. P. Spencer (1976), Extreme wave parameters based on continental shelf storm wave records, in *Proc. 15th Conference on Coastal Engineering*, *Honolulu*, *1*, 151–170, Am. Soc. Civ. Eng., New York.

Hasselmann, K., et al. (1973), Measurements of wind-wave growth and swell decay during the Joint North Sea Wave Project (JONSWAP), *Dtsch. Hydrogr. Z.*, *12*(Suppl., A8), 95 pp.

Haver, S., and O. J. Andersen (2000), Freak waves: Rare realisations of a typical population or typical realisations of a rare population?, in *Proc. 10th Offshore and Polar Engineering Conference*, *Seattle*, *3*, 123–130, Int. Soc. of Offshore and Polar Engineers (ISOPE), Cupertino, Cal.

Holthuijsen, L. H. (2007), *Waves in Oceanic and Coastal Waters*, 387 pp., Cambridge Univ. Press, Cambridge, U. K.

James, I. D. (1986), A note on the theoretical comparison of wave staffs and wave rider buoys in steep gravity waves, *Ocean Eng.*, *13*(2), 209–214.

Janssen, P. A. E. M. (2003), Nonlinear four-wave interactions and freak waves, *J. Phys. Oceanogr.*, *33*(4), 863–884.

Kahma, K. K., and C. J. Calkoen (1992), Reconciling discrepancies in the observed growth of wind-generated waves, *J. Phys. Oceanogr.*, *22*(12), 1389–1405.

Kharif, C., and E. Pelinovsky (2003), Physical mechanisms of the rogue wave phenomenon, *Eur. J. Mech. B*, *22*, 603–634.

Larsen, L. H. (1981), The influence of bandwidth on the distribution of heights of sea waves, *J. Geophys. Res.*, *86*(C5), 4299–4301, doi:10.1029/JC086iC05p04299.

Liu, P. C., and K. R. MacHutchon (2006), Are there different kinds of rogue waves?, *Proc. 25th Conference on Offshore Mechanics and Arctic Engineering, Hamburg, Germany*, paper 92619, Am. Soc. Mech. Eng., New York.

Longuet-Higgins, M. S. (1952), On the statistical distributions of the heights of sea waves, *J. Mar. Res.*, *11*(3), 245–265.

Longuet-Higgins, M. S. (1963), The effect of nonlinearities on statistical distributions in the theory of sea waves, *J. Fluid Mech.*, *17*, 459–480.

Longuet-Higgins, M. S. (1975), On the distribution of the periods and amplitudes of sea waves, *J. Geophys. Res.*, *80*(18), 2688–2693, doi:10.1029/JC080i018p02688.

Longuet-Higgins, M. S. (1980), On the distribution of the heights of sea waves: some effects of nonlinearity and finite band-width, *J. Geophys. Res.*, *85*(C3), 1519–1523, doi:10.1029/JC085iC03p01519.

Magnusson, A. K., M. A. Donelan, and W. M. Drennan (1999), On estimating extremes in an evolving wave field, *Coastal Eng.*, *36*, 147–163.

Mallory, J. K. (1974), Abnormal waves in the south east coast of South Africa, *Int. Hydrographic Rev.*, *51*, 99–129.

Massel, S. R. (1996), *Ocean Surface Waves: Their Physics and Prediction*, *11*, World Scientific, Singapore.

Mori, N., and P. A. E. M. Janssen (2006), On kurtosis and occurrence probability of freak waves. *J. Phys. Oceanogr.*, *36*(7), 1471–1483.
Mori, N., and T. Yasuda (2002), A weakly non-Gaussian model of wave height distribution for random wave train, *Ocean Eng.*, *29*, 1219–1231.
Naess, A. (1985), On the distribution of crest to trough wave heights, *Ocean Eng.*, *12*(3), 221–234.
Nolte, K. G., and F. H. Hsu (1979), Statistics of larger waves in a sea state, *J. Waterway, Port, Coastal and Ocean Eng.*, *105*(WW4); 389–404.
Onorato, M., A. R. Osborne, and M. Serio (2002), Extreme wave events in directional, random oceanic sea states, *Phys. Fluids*, *14*(4), L25–L28.
Onorato, M., et al. (2009), Statistical properties of mechanically generated surface waves: A laboratory experiment in a three-dimensional wave basin, *J. Fluid Mech.*, *627*, 235–257.
Pierson, W. J., and L. Moskowitz (1964), A proposed spectral form for fully developed wind seas based on the similarity theory of S. A. Kitaigorodskii, *J. Geophys. Res.*, *69*(24), 5181–5190, doi:10.1029/JZ069i024p05181.
Smith, R. (1976), Giant waves, *J. Fluid Mech.*, *77*, 417–431.
Socquet-Juglard, H., K. Dysthe, K. Trulsen, H. E. Krogstad, and J. Liu (2005), Probability distributions of surface gravity waves during spectral changes, *J. Fluid Mech.*, *542*, 195–216.
Stansell, P. (2004), Distributions of freak wave heights measured in the North Sea, *Appl. Ocean Res.*, *26*, 35–48.
Stansell, P. (2005), Distributions of extreme wave, crest and trough heights measured in the North Sea, *Appl. Ocean Res.*, *32*, 1015–1036.
Tayfun, M. A. (1983), Effects of spectrum band width on the distribution of wave heights and periods, *Ocean Eng.*, *10*(2), 107–118.
Tayfun, M. A. (1990), Distribution of large wave heights, *J. Waterway, Port, Coastal and Ocean Eng.*, *116*(6), 686–707.
Tayfun, M. A. (1993), Sampling-rate errors in statistics of wave heights and periods, *J. Waterway, Port, Coastal and Ocean Eng.*, *119*(2), 172–192.
Tayfun, M. A. (1994), Distribution of envelope and phase in weakly nonlinear random waves, large wave heights, *J. Eng. Mech.*, *120*(4), 1009–1025.
Tayfun, M. A. (2004), Statistics of wave crests in storms, *J. Waterway, Port, Coastal and Ocean Eng.*, *130*(4), 155–161.
Tayfun, M. A. (2006), Statistics of nonlinear wave crests and groups, *Ocean Eng.*, *33*, 1589–1622.
Tayfun, M. A., and F. Fedele (2007), Wave-height distributions and nonlinear effects, *Ocean Eng.*, *34*, 1631–1649.
Tucker, J. W., and E. G. Pitt (2001), *Waves in Ocean Engineering*, 521 pp, Elsevier, Amsterdam.
Vinje, T. (1989), The statistical distribution of wave heights in a random seaway, *Appl. Ocean Res.*, *11*(3), 143–152.
Waseda, T. (2006), Impact of directionality on the extreme wave occurrence in a discrete random wave system, *Ninth International Workshop on Wave Hindcasting and Forecasting*, Environment Canada, U.S. Army Engineer Research and Development Center's Coastal and Hydraulics Laboratory, WMO/IOC Joint Technical Commission for Oceanography and Marine Meteorology (JCOMM), Victoria, B.C., Canada, paper 8.

M. Casas-Prat, Escola Tècnica Superior d' Enginyers de Camins, Canals i Ports de Barcelona, Technical University of Catalonia, Jordi Girona 1-3, E-08034 Barcelona, Spain. (merce.casas@upc.edu)

L. H. Holthuijsen, Faculty of Civil Engineering and Geosciences, Delft University of Technology, Stevinweg 1, NL-2628CN Delft, Netherlands. (l.h.holthuijsen@ct.tudelft.nl)