

# Extreme Sea State Prediction at E.C.M.W.F.

Jean-Raymond Bidlot and Peter A.E.M. Janssen  
European Centre for Medium range Weather Forecasts,  
Shinfield Park, Reading, RG2 9AX, U.K.  
email: jean.bidlot@ecmwf.int

## 1 Introduction.

Recently, there has been considerable progress in the understanding of the occurrence of freak waves. The notion of freak waves was first introduced by Draper (1965). Freak waves are waves that are extremely unlikely as judged by the Rayleigh distribution of wave heights (Dean, 1990). In practice this means that when one studies wave records of a finite length (say of 10-20 min), a wave is considered to be a freak wave if the wave height  $H$  (defined as the distance from crest to trough) exceeds the significant wave height  $H_S$  by a factor 2. It should be clear that it is hard to collect evidence on such extreme wave phenomena because they occur so rarely. Nevertheless, observational evidence from time series collected over the past decade does suggest that for large surface elevations the probability distribution for the surface elevation may deviate substantially from the one that follows from linear theory with random phase, namely the Gaussian distribution (Wolfram and Linfoot, 2000). Furthermore, there are now a number of recorded cases which show that the ratio of maximum wave height and significant wave height may be as large as three (Stansell, 2005).

The increased understanding of the generation of freak waves follows from the present-day ability to simulate these extreme events by means of the Zakharov equation (Zakharov, 1968, Janssen, 2003 (hereafter referred to as J2003)). This is an approximate evolution equation which is obtained from the exact equations for surface gravity waves in the limit of small wave steepness. Yasuda et al. (1992), Trulsen and Dysthe (1997) and Osborne et al. (2000) studied simplified versions of the Zakharov equation and it was found that these waves can be produced by nonlinear self modulation of a slowly varying wave train. An example of nonlinear modulation or focussing is the instability of a uniform narrow-band wave train to side-band perturbations. This instability, known as the side-band, modulational or Benjamin-Feir (1967) instability, will result in focusing of wave energy in space and/or time as is illustrated by the experiments of Lake et al. (1977).

Therefore, in the context of the deterministic approach to wave evolution there seems to be a reasonable theoretical understanding of why in the open ocean freak waves occur. In ocean wave forecasting practice one follows, how-

ever, a stochastic approach because the phases of the individual waves are unknown. Clearly, in the context of wave forecasting only statements of a probabilistic nature can be made. As freak waves imply considerable deviations from the Normal, Gaussian probability distribution function (pdf) of the surface elevation, the main question therefore is whether the pdf of the surface elevation can be determined in a reliable manner. Traditionally, it is known that the surface elevation pdf deviates from the Normal distribution because the actual shape of the ocean waves deviates from the sinusoidal form. However, there is also a dynamical cause for deviations from Normality. J2003 showed that the deviations from the Normal pdf of the surface elevation are also related to the presence of resonant and nonresonant four-wave interactions. In fact, the kurtosis, which vanishes for a Gaussian distribution and is a measure for extreme events, was found to be related to a six-dimensional integral involving the action density to the third power.

As a first step towards validation of Janssen's approach, the kurtosis was evaluated from the theoretical expression and for uni-directional, narrow-band spectra it was found that the dynamical part of the kurtosis depends on the square of the Benjamin-Feir Index (BFI). Here, the BFI is the ratio of the wave steepness to the spectral bandwidth. This dependence on the BFI was confirmed by recent experimental work done by Onorato et al. (2009).

For operational implementation the expression for the kurtosis is far too involved, and clearly some simplification is desirable. It is assumed that freak wave events most likely only occur for narrow band wave trains. This corresponds to situations where both the frequency and angular distribution of the waves is narrow. In the narrow-band approximation it is possible to simplify and evaluate the six-dimensional integral.

The general result for the kurtosis and its relation to the wave spectrum was originally derived for deep-water waves, but Janssen and Onorato (2007) have shown how to extend it to shallow water. Finite-amplitude deep-water waves are subject to modulational instability which results in a nonlinear energy transfer among the components in the wave spectrum, which eventually can lead to the formation of extreme waves. However, in shallow water, finite-amplitude surface gravity waves generate a current and deviations from the mean surface elevation. This stabilizes the modulational instability, and as a consequence, in a fairly wide range around  $kD = 1.363$  the nonlinear transfer becomes small. In addition, while for  $kD > 1.363$  there is nonlinear focussing giving the possibility of the formation of extreme waves, in the opposite case the process of nonlinear focussing ceases to exist. For narrow-band spectra, it is then straightforward to parametrize the stabilizing effects of shallow water.

This paper presents how these latest developments lead to the operational implementation of the extreme sea state prediction system at the European Centre for Medium range Weather Forecasts (ECMWF). A short description of the system is given. The details of the theoretical derivations are referred to the technical memorandum by Janssen and Bidlot (2009, hereafter JB09)). As pointed out, a common practice to describe freak waves is in term of ratio of maximum wave height to significant wave height. Therefore, in this latest extension of its extreme sea state prediction system, ECMWF introduced two new parameters, namely maximum wave height  $H_{max}$  and the corresponding maximum period  $T_{max}$ . The derivation of  $H_{max}$  and an attempt at verification is presented in Section 2 (for  $T_{max}$  see JB09).

## 2 Extension of freak wave warning system.

Before starting with a detailed calculation of the kurtosis of the sea surface and its dependence on the wave spectrum, it is briefly mentioned that the starting point of ocean waves dynamics is the Hamiltonian formulation of the nonlinear water wave equations. Assuming that the waves are weakly nonlinear and applying a *canonical* transformation which removes most of the contributions by non-resonant interactions, one arrives at the well-known Zakharov equation for the free wave part of the action variable. The properties of the Zakharov equation have been studied in great detail by, for example, Crawford et al. (1981) for deep-water waves and by Janssen and Onorato (2007) for shallow-water waves. It describes all the known properties of weakly nonlinear waves in deep and shallow water and is therefore a good starting point for further analysis.

Based on the above theoretical development it should be clear that the expression of the kurtosis of the pdf of the surface elevation consists of two additive contributions. The first one was derived by Janssen (2003) and reflects the effects of resonant and non-resonant four-wave interactions, while the second contribution stems from the canonical transformation and reflects the contribution from asymmetries in the shape of the waves. However, the contribution of the canonical transformation yields a lengthy expression and only for narrow band wave trains its form is known explicitly (JB09). In the following, the definition of kurtosis used in this work is introduced. Then the general expression of the contribution to the kurtosis by the dynamics of the waves is presented and the limit of a narrow-band wave train is taken. The total kurtosis then consists of the sum of the 'dynamics' contribution and the 'wave-shape' contribution.

### 2.1 Kurtosis for narrow-band ocean waves.

There are several definitions of kurtosis possible. Here, it is defined in such a way that it is directly related to the fourth cumulant of the pdf of the surface elevation  $\eta$ . Hence, the kurtosis  $C_4$  is defined as

$$C_4 = \frac{\langle \eta^4 \rangle}{3\langle \eta^2 \rangle^2} - 1. \quad (1)$$

The advantage of this definition of kurtosis (some call it the 'excess' kurtosis) is that for a Gaussian pdf  $C_4$  vanishes since for a Gaussian  $\langle \eta^4 \rangle = 3\langle \eta^2 \rangle^2$ . Hence,  $C_4$  measures deviations from the Gaussian sea state. In other words, when  $C_4 > 0$  the probability of extreme events is higher than expected from the Normal distribution, while when  $C_4 < 0$  the probability of extreme events is lower than 'Normal'. On the other hand, as shown in Janssen (2004), the four-wave interactions only occur because the fourth cumulant is finite, hence there is a direct connection between the changes in the wave spectrum caused by nonlinear four-wave interactions and extreme sea states.

J2003 obtained an expression for the 'dynamics' part of the kurtosis  $C_4$  in terms of the action density spectrum  $N$  (cf. Eq. (29) of J2003). Denoting the variance of the surface elevation by  $m_0 = \langle \eta^2 \rangle$ , one finds

$$C_4 = \frac{4}{g^2 m_0^2} \int d\mathbf{k}_{1,2,3,4} T_{1,2,3,4} \delta_{1+2-3-4} (\omega_1 \omega_2 \omega_3 \omega_4)^{\frac{1}{2}} G(\Delta\omega, t) N_1 N_2 N_3, \quad (2)$$

where the transfer function  $G$  is given by

$$G(\Delta\omega, t) = \frac{1 - \cos(\Delta\omega t)}{\Delta\omega}. \quad (3)$$

Here,  $\Delta\omega = \omega_1 + \omega_2 - \omega_3 - \omega_4$ ,  $T_{1,2,3,4}$  is a complicated, homogeneous function of the four wave numbers  $\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4$  which because of the  $\delta$ -function enjoy the resonance condition  $\mathbf{k}_1 + \mathbf{k}_2 = \mathbf{k}_3 + \mathbf{k}_4$ . In addition, the angular frequency  $\omega(\mathbf{k})$  obeys the dispersion relation  $\omega(\mathbf{k}) = \sqrt{gkT_0}$ , with  $k$  the magnitude of the wave number vector  $\mathbf{k}$  and  $T_0 = \tanh(kD)$ , where  $D$  is the water depth. Here only the deep-water limit,  $D \rightarrow \infty$ , will be discussed. The shallow water extension was addressed in JB09.

Eq. (2) is valid for arbitrary two-dimensional action density spectra. Although, strictly speaking, the determination of the kurtosis involves an eight-dimensional integral in wave number space, the resonance conditions restrict the evaluation to a six-dimensional subspace only. Nevertheless, for operational purposes this is still far too time-consuming, simplifying assumptions have to be made. Here, the so-called narrow-band approximation is assumed which basically implies almost unidirectional waves that have a sharply peaked frequency spectrum. In practice, around the peak of the spectrum this is usually a valid approximation.

Introducing the frequency spectrum

$$E(\omega, \theta) d\omega d\theta = \frac{\omega N(\mathbf{k})}{g} d\mathbf{k},$$

and performing the integration over  $\mathbf{k}_4$ , then Eq. (2) becomes

$$C_4 = \frac{4g}{m_0^2} \int d\omega_1 d\omega_2 d\omega_3 d\theta_1 d\theta_2 d\theta_3 T_{1,2,3,4} \sqrt{\frac{\omega_4}{\omega_1 \omega_2 \omega_3}} G(\Delta\omega, t) E_1 E_2 E_3. \quad (4)$$

with,

$$\omega_4 = \Omega(\mathbf{k}_4) = \sqrt{g|\mathbf{k}_1 + \mathbf{k}_2 - \mathbf{k}_3|}.$$

Now the narrow-band approximation is applied, i.e. the spectrum is mainly concentrated at  $\omega = \omega_0$  and  $\theta = \theta_0$ , and falls off rapidly, much faster than the other terms in the integrand of Eq. (4). In that event, the transfer coefficient  $T_{1,2,3,4}$  can be approximated by its narrow-band value  $k_0^3$ . In addition,  $\omega_4$  is approximated. Denoting the width of the frequency spectrum by  $\sigma_\omega$  and the angular width by  $\sigma_\theta$  one may write for angular frequency and direction

$$\omega_1 = \omega_0(1 + \delta_\omega \nu_1), \quad \theta_1 = \theta_0 + \delta_\theta \phi_1,$$

where in the narrow-band approximation the parameters  $\delta_\omega$  and  $\delta_\theta$ , defined as

$$\delta_\omega = \frac{\sigma_\omega}{\omega_0}, \quad \delta_\theta = \sigma_\theta, \quad (5)$$

are small. The angular frequency  $\omega_0$  may be defined in several ways. For example, one could take it as the peak frequency. Here, for convenience it is defined by means of the first moment

$$\omega_0 = \int d\omega d\theta \omega E(\omega, \theta) / m_0.$$

Expanding  $\omega_4$  in the small parameters  $\delta_\omega$  and  $\delta_\theta$  one finds up to third order an expression for  $\omega_4$ . As a consequence, the frequency mismatch  $\Delta\omega$  becomes

$$\Delta\omega = \delta_\omega^2 \omega_0 \{ (\nu_3 - \nu_1)(\nu_3 - \nu_2) - R(\phi_3 - \phi_1)(\phi_3 - \phi_2) \} \quad (6)$$

where the parameter  $R$  has been introduced which measures the importance of the angular width with respect to the frequency width,

$$R = \frac{1}{2} \frac{\delta_\theta^2}{\delta_\omega^2}.$$

Introducing the integral steepness parameter

$$\epsilon = k_0 \sqrt{m_0}.$$

and applying the narrow-band approximation to  $C_4$  yields

$$C_4^{dyn} = 4\epsilon^2 \omega_0 \int d\nu_1 d\nu_2 d\nu_3 d\phi_1 d\phi_2 d\phi_3 G(\Delta\omega, t) \hat{E}_1 \hat{E}_2 \hat{E}_3. \quad (7)$$

where  $\Delta\omega$  is given by Eq. (6), and the spectrum  $E$  is now regarded as a function of  $\nu$  and  $\phi$ . Also, the spectrum has been normalised in such a way that  $m_0 = 1$ , hence  $\hat{E}_1 = E(\nu_1, \phi_1)/m_0$ .

It is important to realise that according to Eq. (7), the kurtosis depends on the square of the BFI, as expressed by the ratio of  $\epsilon$  and  $\delta_\omega$ , and on the ratio of the directional width and frequency width through the parameter  $R$ .

Eq. (7) is the general expression for the dynamics part of the kurtosis of a narrow-band wave train (for this reason the label 'dyn' is temporarily added). As explained in the beginning of this section, there is also a contribution due to the asymmetrical shape of the waves related to the canonical transformation. For a narrow-band wave train one can write down the canonical transformation explicitly and the resulting kurtosis may be evaluated. As a result one finds (Janssen, 2009)

$$C_4 = C_4^{dyn} + 6\epsilon^2.$$

Therefore, for a narrow-band wave train the wave-shape contribution to the kurtosis is known in terms of the moments of the spectrum, and it is straightforward to evaluate its contribution.

## 2.2 Operational Implementation of kurtosis calculation.

Even though, according to (7), the dynamical part of the kurtosis can be explicitly evaluated, for operational use, a suitable approximation was found. Based on numerical simulations with the two-dimensional, Nonlinear Schrödinger Equation, Mori (private communication, 2007) found the following fit for the maximum of the kurtosis

$$C_4^{dyn} = \frac{0.031}{\delta_\theta} \times \frac{\pi}{3\sqrt{3}} BFI^2, \quad (8)$$

where  $BFI$  is the Benjamin-Feir Index, defined as

$$BFI = \frac{\epsilon\sqrt{2}}{\delta_\omega}. \quad (9)$$

therefore, finite directional width  $\delta_\theta$  is seen to give a considerable reduction in kurtosis  $C_4^{dyn}$ . Including the contribution from the shape of the waves the total kurtosis becomes now

$$C_4 = C_4^{dyn} + \alpha\epsilon^2. \quad (10)$$

where for deep-water  $\alpha = 6$ .

This result holds for deep-water waves. The extension to shallow water is achieved by means of a redefinition of the Benjamin-Feir Index.<sup>1</sup> (JB09).

### 2.2.1 Determination of the *BFI* and $\delta_\theta$ .

The estimation of the Benjamin-Feir Index requires knowledge of the significant steepness  $\epsilon$  and the spectral width  $\delta_\omega$  in frequency space. In addition, an estimate of the directional width  $\delta_\theta$  is required as well. Here, a description is given of a robust method to estimate the *BFI* for modelled and observed spectra. In particular, the estimation of the width of observed frequency spectra is not a trivial task, because observed spectra show considerable noisy behaviour around the peak of the spectrum (which is frequently ill-defined).

Janssen and Bouws (1986) developed a robust method to estimate the width of observed spectra, which was applied to frequency spectra obtained from a waverider located at IJmuiden over a fifteen year period. Following Goda these authors used the peakedness factor  $Q_p$  defined as

$$Q_p = \frac{2}{m_0^2} \int_{\mathcal{D}} d\omega \omega E^2(\omega)$$

where Janssen and Bouws (1986) chose, after extensive experimentation, as integration domain  $\mathcal{D}$  all frequencies for which  $E(\omega) > 0.25E(\omega_p)$ . The advantage of this integral measure is that, because of the dependence on the square of the frequency spectrum, peaks in the spectrum are emphasized. Janssen and Bouws (1986) also explored alternative integral measures such as one based on the second moment of the wave spectrum, but these alternatives give more emphasis to the high-frequency part of the spectrum and are therefore more sensitive to high-frequency noise.

Janssen and Bouws (1986) checked from the observed spectra that to a good approximation the spectra are symmetrical around the peak and that the Gaussian approximates the observed spectral shape well. In the narrow-band approximation one finds to high accuracy

$$Q_p = \frac{1}{\delta_\omega \sqrt{\pi}} \quad (11)$$

where  $\delta_\omega$  is the relative width defined in Eq. (5). A robust method to estimate the relative spectral width now is to determine the spectral  $Q_p$  and to invert Eq. (11), hence

$$\delta_{\omega,obs} = \frac{1}{Q_{p,obs} \sqrt{\pi}}$$

---

<sup>1</sup>Also the parameter  $\alpha$  needs adjustment for the shallow water case, but this has not been introduced yet

As a consequence, the observed BFI becomes

$$BFI = k_0 m_0^{1/2} Q_{p,obs} \sqrt{2\pi}$$

The modelled BFI is calculated in an identical fashion through the peakedness factor  $Q_p$  and the integral steepness  $\epsilon$ .

The directional width  $\delta_\theta$  at the peak of the spectrum could be estimated by the usual approach, i.e.

$$\delta_\theta = \sqrt{2(1 - M_1)}$$

where  $M_1 = I_1/m_0$  and  $I_1 = \int d\omega d\theta \cos(\theta) E(\omega, \theta)$ , but it won't always provide the sharpest estimate of directional width near the peak.

An alternative approach to estimate the frequency and directional width of the two-dimensional model spectrum is to fit the one-dimensional frequency and directional spectra with a parabola thus giving sharp estimates for  $\delta_\omega$  and  $\delta_\theta$ . In fitting the parabola also a sharper estimate of the peak period  $T_p$  may be provided as up to now the peak period did correspond to the maximum of the one-dimensional frequency spectrum so  $T_p$  could only assume discrete values because of the discretization of the wave spectrum in frequency space. However, occasionally the fitting procedure may fail because, e.g., the peak of the spectrum is erratic. Therefore the widths are determined by taking the minimum value from the integral method, i.e.  $Q_p$  and  $M_1$ , and from the fitting procedure. Nevertheless, because of the relatively coarse discretization of the spectrum, narrow spectra are too wide in the present version of the wave model (24 directions). To accomodate for this, the constant has been increased in the expression for the kurtosis, Eq. (8), by a factor of two from 0.031 to 0.062.

### 3 Maximum wave height.

It is common to define as a freak wave a wave whose height is at least 2.2 times the significant wave height. This is a very discrete and singular approach, which is in practice not easy to verify. Nevertheless, it is desirable to be able to quantify extreme sea states and to be able to validate them against observations in a meaningful manner. It is then natural to consider the concept of maximum wave height, a concept which is well-known in engineering practice. It should be realized, as also pointed out extensively by Mori and Janssen (2006), that the maximum waveheight  $H_{max}$  not only depends on the shape of the probability distribution function of the sea surface, but also on the number of waves at hand. Consider now a time series of wave heights of length  $T$  involving a number of  $N$  waves. A good estimate of the maximum wave height is the expectation value for maximum wave height denoted by  $\langle H_{max} \rangle$ . As an extension of Goda's work for Gaussian sea states,  $\langle H_{max} \rangle$  will be determined for a pdf with finite kurtosis and the result will be compared with observations of maximum wave height from buoys. The agreement is good, and therefore this measure for maximum wave height was introduced into the operational ECMWF wave forecasting system in June 2008.

Before proceeding it is mentioned that there is an important caveat. It is well-known that for narrow-band wave trains the probability density function (pdf) of wave height is the Rayleigh distribution. This was shown a long time ago

by Longuet-Higgins (e.g. 1957). He noted that it is in general straightforward to obtain the statistical properties of the envelope of a wave train, even for broad-band wave trains. For a Gaussian sea state the pdf of the envelope is found to be the Rayleigh distribution. The statistical properties of waveheight are much harder to obtain. For narrow-band wave trains it can be argued that waveheight is twice the envelope and thus wave height will then follow the Rayleigh distribution as well. However, for broad-banded wave trains the pdf of wave height is not known.

One may wonder why it is so difficult to obtain the pdf of wave height for general spectra of finite width. An important reason for this is that, at least in a theoretical context, wave height is an *ill-defined* quantity, in contrast to, for example, the envelope of a wave train. Analyzing a time series it is fairly easy (JB09) to construct at any point in time the envelope of a wave train, however, this is not possible for the wave height of a wave train (except of course in the narrow-band approximation). In practice, researchers obtain the wave height distribution by means of the zero-crossing method. This is a very elegant method, which is easily implemented: Search for two consecutive zero-upcrossings in the time series and determine the wave height from the difference of the maximum and the minimum of the surface elevation  $\eta$  in the corresponding time interval. Thus, wave height is determined by sampling with the zero-crossing frequency  $(m_2/m_0)^{1/2}$  (with  $m_n$  the  $n^{th}$  moment of the wave spectrum). However, what about sampling with other frequencies, corresponding to different (spatial) scales. For higher sampling frequency, wave heights are expected to be reduced compared to lower sampling frequency because one would expect that at smaller scales wave heights are smaller. Therefore wave height depends on the choice of spatial and temporal scale, and hence the wave height pdf will depend on the way one samples the time series.

For the envelope distribution there is much less of a problem, because the envelope is a continuous function of time. By sampling at a sufficiently high frequency one simply gets the 'usual' pdf for envelope. In fact, in JB09 a review of the derivation of the pdf of the envelope is given and it was shown that for linear waves the pdf is always Rayleigh, despite claims by Longuet-Higgins (1983) to the contrary.

Finally, one may wonder why one is interested so much more in the wave-height distribution rather than the envelope probabilities. If one is interested in extreme forces on structures such as oil rigs or ships than one would expect that the quantity of interest is something like the energy of the waves, which is closely related to the square of the envelope. For extreme cases the square of the wave height would underestimate the force on structures (as the pdf of wave height falls below the Rayleigh distribution, while the pdf of the envelope is Rayleigh). In other words, there is a case to concentrate on the envelope distribution rather than the wave height distribution. Alternative arguments to use the envelope rather than wave height are presented in Longuet-Higgins (1984).

Therefore, the theoretical developments will all concern the (statistical) properties of the envelope of a wave train and wave height is defined as twice the envelope. Details of the theoretical development and its verification against Monte Carlo simulations was presented in JB09. In order to obtain an expression for the expectation value of maximum wave height the work of Mori and Janssen (2006) was followed closely. One may then take the following steps

1. Start from the pdf of surface elevation  $\eta$ , which is the well-known Gram-Charlier expansion, i.e. pdf depends on skewness and kurtosis, which are assumed to be small.
2. Obtain the pdf of 'wave height' defined as twice the envelope. Here the envelope  $\rho$  follows implicitly by writing the surface elevation signal as

$$\eta = \rho \cos \phi$$

with  $\phi$  the local phase of the wave train. Local wave height is then defined as  $h = 2\rho$  and the wave height distribution in terms of  $H$ , the wave height normalized with the significant wave height becomes:

$$p(H) = 4H \exp(-2H^2) [1 + C_4 A_H(H)] \quad (12)$$

where

$$A_H(H) = 2H^4 - 4H^2 + 1$$

Note that because of symmetries the pdf of  $H$  does not contain skewness.

3. The maximum wave height distribution is obtained by simply writing down the probability that for given number of independent waves  $N$  the maximum wave height has a certain chosen value. The maximum wave height distribution  $p_m(H_{max})$  becomes

$$p_m(H_{max}) = N [1 - P(H_{max})]^{N-1} p(H_{max})$$

where, with  $B_H(H) = 2H^2 (H^2 - 1)$ ,

$$P(H) = \int_H^\infty dh p(h) = \exp(-2H^2) (1 + C_4 B_H(H))$$

is the exceedence probability of wave height,  $N$  is the number of waves, and  $p(H_{max})$  follows from Eq. (12). In the continuum limit this becomes

$$p_m(H_{max}) = N p(H_{max}) \times \exp[-NP(H_{max})] \quad (13)$$

Notice that the maximum wave height distribution involves a double exponential function.

4. The expectation value of maximum wave height follows from

$$\langle H_{max} \rangle = \int_0^\infty dH_{max} H_{max} p_m(H_{max}) \quad (14)$$

Notice that  $H_{max} = F[C_4(BFI, R), N]$ , where  $N = T_D/T_p$  with  $T_p$  the peak period and  $T_D$  the duration of the timeseries. By making this choice for the number of waves  $N$  it is tacitly assumed that two successive 'waves' are uncorrelated. This assumption is hard to justify because the correlation between two following waves may be of the order of 50 %. It would be more appropriate to correct for this correlation thereby either reducing the number of degrees of freedom or reducing the variance of the pdf.

The integral in (14) may be evaluated in an approximate fashion for large  $N$  and small  $C_4$  (JB09). The main result becomes

$$\langle H_{max} \rangle = \sqrt{\langle z \rangle}, \quad (15)$$

where

$$\langle z \rangle = \hat{z}_0 + \frac{\gamma}{2} + \frac{1}{2} \log \left[ 1 + C_4 \left\{ 2\hat{z}_0(\hat{z}_0 - 1) - \gamma(1 - 2\hat{z}_0) - \frac{1}{2}(\gamma^2 + \frac{\pi^2}{6}) \right\} \right], \quad (16)$$

with  $\hat{z}_0 = \frac{1}{2} \log N$  and  $\gamma = 0.5772$  is Euler's constant. An estimate of the sharpness of the estimate for the expectation value of maximum wave height may be given as well. This follows immediately from the width  $\sigma$  of the maximum wave height distribution. For linear waves its width  $\sigma$  is approximately (JB09)

$$\frac{\sigma}{\langle H_{max} \rangle} \simeq \frac{\pi}{2\sqrt{6} (\log N + \frac{1}{2}\gamma)}, \quad (17)$$

and clearly, the longer the time series of independent events, the sharper the estimate for maximum wave height becomes.

**ECMWF Analysis VT:Saturday 10 February 2007 00UTC Surface:  
Maximum wave height (m)**

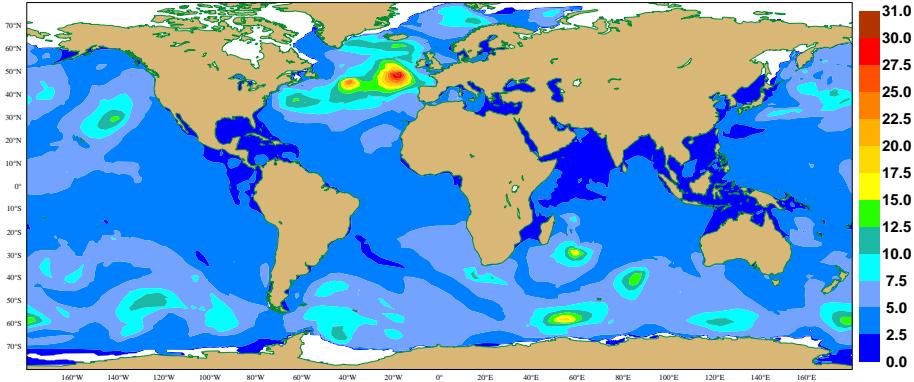


Figure 1: Map of analyzed maximum wave height for the 10<sup>th</sup> of February 2007.

Let us now discuss some characteristic properties of the new freak wave warning system. In Fig. 1 an example of a maximum wave height map is shown for a big storm in the North Atlantic that occurred on the 10<sup>th</sup> of February 2007. Here, the maximum wave height refers to time series with a duration  $T_D$  of 3 hrs and the number of waves  $N$  follows from the relation  $N = T_D/T_p$ , where  $T_p$  is the peak period. The maximum of significant wave height in the North Atlantic was 15.9 m at that time while the extremum in maximum wave height is found to be 31.6 m. Notice, however, the dependence of the estimate of the maximum wave height on the number of waves in the time series of duration  $T_D$ . Although according to Eq. (16) it only depends on the logarithm of  $N$ , nevertheless for  $T_D = 20$  min maximum wave height will decrease on average by about 20% giving an extreme value of 26.5 m. Inspecting the kurtosis map shown in Fig. 2, however, it is found that regarding maximum wave height, the

ECMWF Analysis VT: Saturday 10 February 2007 00UTC Surface: Wave spectral kurtosis Kurtosis

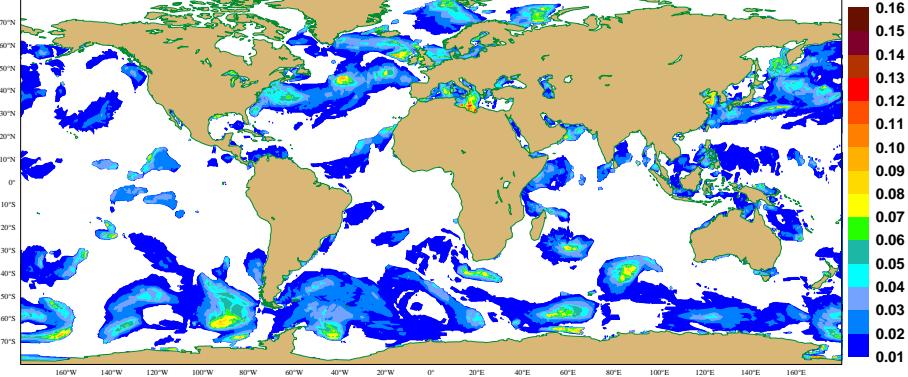


Figure 2: Map of analyzed kurtosis  $C_4$  for the 10<sup>th</sup> of February 2007.

extreme event in the North Atlantic was not exceptional as the kurtosis  $C_4$  was only about 0.06 corresponding to a normalized maximum wave height  $H_{max}/H_S$  of only 1.95. In order to appreciate that such a condition is not exceptional the left panel of Fig. 3 shows the relation between  $C_4$  and  $BFI$  obtained from the global field for February 10, 2007, 00 UTC. Typically, maximum values of kurtosis are around 0.2 at values of  $BFI$  of the order 1. It is also of interest to study under what kind of meteorological conditions exceptional waves may occur. Some information on this is provided by the right panel of Fig. 3, which shows kurtosis plotted against the wave age parameter  $c_p/U_{10}$ . In particular for young windsea with  $c_p/U_{10} < 1$  large values of kurtosis, and hence abnormal sea states, are possible according to the present approach. Young windseas typically occur in fetch-limited conditions, when the wind just start blowing or during the passage of a front when the wind turns by a significant amount.

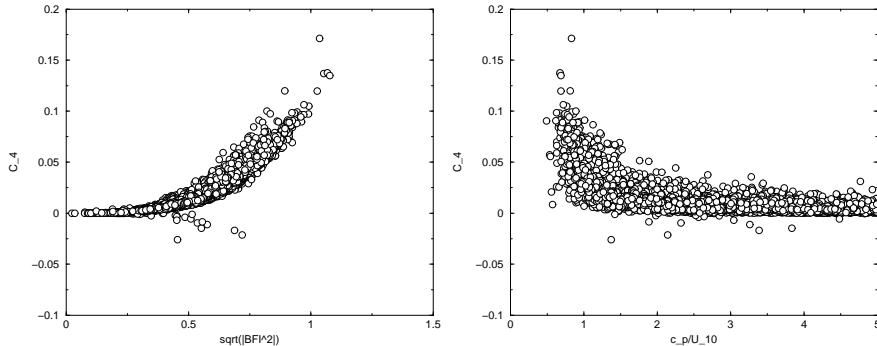


Figure 3: The left panel shows the dependence of kurtosis  $C_4$  on the Benjamin-Feir Index, while the right panel shows the dependence of  $C_4$  on the wave age parameter  $c_p/U_{10}$ .

According to Eqns. (15)-(16) the normalised maximum wave height depends on two parameters namely the number of waves  $N$  and the kurtosis parameter  $C_4$ . Fig. 4 shows the dependence of kurtosis on these two parameters as

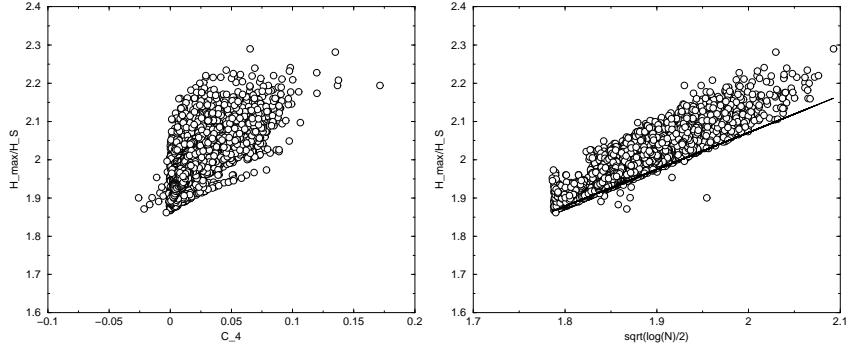


Figure 4: The left panel shows the kurtosis dependence of the expectation value of normalised maximum wave height  $\langle H_{max} \rangle$  while the right panel shows the dependence of  $\langle H_{max} \rangle$  on the number of waves  $N$  in the timeseries of duration of 3 hrs through the parameter  $\sqrt{\log N/2}$ . The full line shows the relation between  $H_{max}$  and the number of waves for vanishing kurtosis.

obtained from the global field of Fig. 1. In particular, the figure in the right panel, which shows normalised maximum wave height as function of  $\sqrt{\log N/2}$ , is illuminating. A comparison with the corresponding relation for vanishing kurtosis immediately shows the importance of nonlinearity on the estimate of maximum wave height. While for this synoptic case the full line never meets the criterium for freak waves to occur (recall the condition for freak waves is  $H_{max}/H_S > 2.2$ ), when effects of nonlinearity through a finite value of kurtosis are included there *are* a number of cases that meet the criterion for extreme events. The question now is how realistic is the ECMWF freak wave warning system.

### 3.1 Verification aspects and maximum wave height verification.

It is clear that for operational applications a choice for the length of the time-series needs to be made. Buoy time series are typically 20-30 minutes long so initially it was thought that, in order to validate the model results against buoy data, it would make sense to take this period as the length of the time series. However, for practical application a timescale related to the changes in the synoptic conditions seems more appropriate. This would mean a much longer duration of say 3 hrs. A compromise was found by choosing a duration of 3 hrs, while for validation purposes 6 consecutive buoy observations were collected making up an observed duration of about 3 hrs. The observed maximum wave height is then the maximum of the 6 consecutive maximum wave height observations.

In the data set currently used in the ECMWF wave verification system (Bidlot et al., 2005; Bidlot et al., 2007) only Canada (Meds) and Norway (Oceanor) supply buoy observations of maximum waveheight. Inspecting the distributions for normalized maximum wave height of MEDS buoys and Oceanor buoys it was found that they belong to two different populations: the mean value of normalized maximum wave height of the Oceanor buoys was considerably smaller

Comparison against Canadian (MEDS) and Norwegian (Oceanor) buoys:

All buoys 20060202 to 20080131

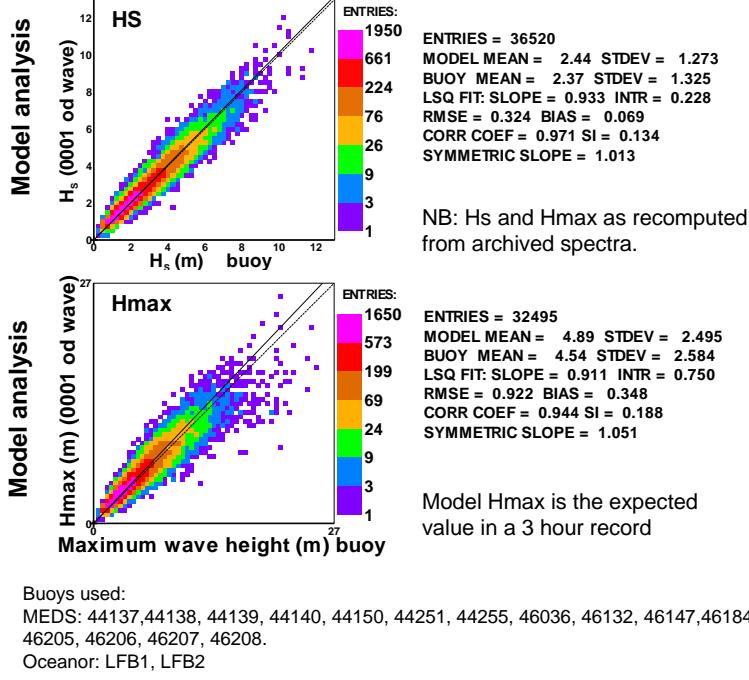


Figure 5: Validation of analyzed maximum wave against observed maximum wave height from a number of buoys that report maximum wave height (the buoy list is shown as well). Period is February 2006 until January 2008. For a comparison of the quality of the  $H_{max}$  estimates the validation of model wave height against buoy data is shown as well.

than the mean value from the MEDS buoys. It is suspected that this is related to a different length of the time series used (17.5 min. (Oceanor) versus 30 min. (MEDS)) and possibly to a different procedure to obtain an estimate of maximum wave height. Because the majority of maximum wave height measurements is from MEDS, only the latter data will be considered for the validation of the probability distribution function, although for the verification of maximum wave height all data will be used.<sup>2</sup> The MEDS buoys have a single accelerometer and the maximum wave height is obtained by taking twice the maximum of a surface elevation timeseries obtained at all the times where acceleration is minimal. This procedure does not give the maximum of envelope wave height but there is no other routinely observed information on maxima available. Nevertheless, this may give rise to problems in the interpretation of the comparison between model and observations.

First results of a comparison of modelled and observed maximum wave height are shown in Fig. 5. For a first comparison the agreement between modelled and observed maximum wave height is quite impressive. The relative positive bias

<sup>2</sup>The MEDS data have the additional advantage that also one-dimensional spectra are reported. These are needed later to determine the *BFI*.

is about 5% while the scatter index is about 19%. For comparison the scatter index for significant wave height for the same set of buoys and period is about 13%. This impressive agreement is puzzling, because for starters actually apples and pears are being compared, since the model value is an expectation while the buoy value is instantaneous. This puzzle was solved when it was realized that the pdf of maximum wave height is fairly narrow. For linear waves its width  $\sigma$  is approximately given by Eq. (17). Clearly, the longer the length of the time series the sharper the estimate of maximum wave height becomes. For a 3 hour duration and a peak period of 10 s one finds  $\sigma/\langle H_{max} \rangle \simeq 0.08$ , therefore the maximum wave height distribution is indeed fairly narrow as the scatter index has the much larger value of 19%.

### 3.2 Verification of the probability density function.

Nevertheless, it is emphasized that apples and pears *are* being compared. This is clearly visible in the plot of the geophysical<sup>3</sup> distribution of normalised (by significant wave height) expectation value and a comparison with the graph of the distribution of the actual, observed value of the normalised maximum wave height, as shown in the left panel of Fig. 6. The width of the modelled maximum wave height distribution, being about 0.05, is much smaller than the width of the observed distribution, which is about 0.16 and it is evident that there is no resemblance between the two distributions. The reason for this discrepancy is most likely that the observed distribution is a single realisation which is not necessarily representative for the area of interest, while the modelled distribution is based on the expectation value of the normalised maximum wave height.

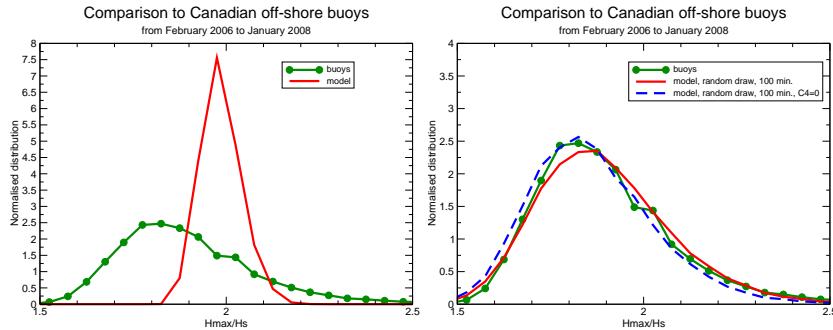


Figure 6: The left panel shows the comparison between observed  $H_{max}/H_S$  distribution and the modelled distribution of the expected normalised maximum wave height. The right panel shows instead of the distribution of the expected maximum wave height the model distribution obtained by a random draw of  $H_{max}$  for given number of waves and given kurtosis. The right panel also shows the impact of nonlinearity on the maximum wave height distribution by means of a plot of the case of zero kurtosis. The length of the timeseries is 100 min. which is thought to match the length of the buoy time series.

---

<sup>3</sup>There is a need now to make a distinction between the maximum wave height pdf and the geophysical distribution of maximum wave height. In principle the geophysical distribution follows from the combination of the maximum wave height pdf and the geophysical distribution of the number of waves  $N$  and the kurtosis  $C_4$ . Only when the latter distributions are much more narrow than the maximum wave height pdf the geophysical distribution will coincide with the maximum wave height pdf. For brevity the adjective geophysical will be dropped

The question now arises whether it is possible to simulate the observed distribution of normalised maximum wave height. This turns out to be possible indeed and in order to understand the method that will be followed, it is important to return to the basic mechanism of freak wave generation. As already discussed in the Introduction, freak waves are regarded to be the result of a nonlinear focussing phenomenon but it should be realized, as pointed out in J2003, that the focussing is the most efficient when the phases of the waves involved in the focussing are chosen appropriately (constructive interference). However, in the field there is no knowledge of the phases and for practical purposes the phases are chosen in an *almost* random manner. Nonlinearity will give rise to a certain degree of correlation between the waves and for this reason the adjective almost, and the effects of small nonlinearity on the pdf are given in Eqns. (12) and (13).

A way to simulate the observed distribution of maximum waveheight is therefore to start from the theoretical pdf of maximum wave height (13), and to generate from this pdf for given number of waves  $N$  and given kurtosis  $C_4$  a random draw of normalised maximum wave height. Basically one obtains a random draw of maximum waveheight from the condition that the cumulative distribution is a random number between 0 and 1. For duration, a 100 min period has been chosen as this is thought to match the length of the buoy time series appropriately, despite the fact that according to the data provider the length of the time series is 30 min.<sup>4</sup> The resulting modelled distribution function is plotted in the right panel of Fig. 6 and the very good agreement with the observed distribution is to be noted, in particular in the extremes. For reference, also the model distribution according to linear theory (i.e.  $C_4 = 0$ ) is plotted and although linear theory gives a reasonable agreement with the observations it is noted that extremes are underestimated by linear theory. This underestimation of the extremes has some practical consequences. It is common to define a freak wave as an event with  $H_{max}/H_S > 2.2$ . Integrating the nonlinear and the linear distribution from 2.2 until infinity one finds that according to linear theory 4.5% of the cases are freak wave events while according nonlinear theory 7.5% of the cases are freak waves which amounts to an increase of 60%. According to the observations 8.5% of the cases are freak waves, therefore nonlinear theory underestimates the number of freak waves somewhat.

The slight underestimation by nonlinear theory is more pronounced when a plot of the logarithm of the distribution is made as shown in Fig. 7 and is compared to the logarithm of the observed distribution.<sup>5</sup> It is evident that the really extreme events with  $H_{max}/H_S > 2.5$  are seriously underestimated by the present nonlinear theory, although in the range of 1.9 until 2.5 there is good agreement. The reason for the discrepancy between model and observations is not clear at present. Noting that this is a first, preliminary comparison, a

---

<sup>4</sup>Note that according to Fig. 5 the model overestimates maximum wave height by 5%. This overestimation can be removed by reducing the number of degrees of freedom  $N$  or equivalently by shortening the length of the timeseries from 180 min. to 100 min. This reduction in the number of degrees of freedom is in qualitative agreement with the correlation between two successive waves.

<sup>5</sup>This comparison was restricted to cases with a significant wave height larger than 2 m because buoys might have problems with accurately representing low sea states. This is also evident in the next section where buoys are not representing high frequencies very well. This reduces the number of collocations from 32,000 to 16,000. Nevertheless there are still about 1,300 cases that satisfy the freak wave criterion of  $H_{max}/H_S > 2.2$ .

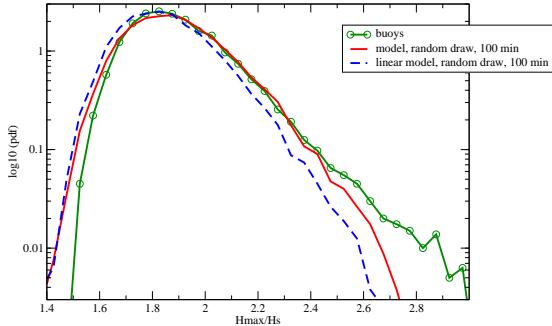


Figure 7: The logarithm of maximum wave height distribution obtained by a random draw of  $H_{max}$  for given number of waves and given kurtosis as compared to the observed maximum wave height distribution.

number of detailed studies of the buoy time series need to be carried out. A first look at the time series for maximum wave height suggests that these really extreme events are present only for a very short time. However, at present there is no criterion to decide whether these cases can be regarded as outliers or not. Also, the buoys are giving maximum wave height based on twice the crest value which may be an overestimate of envelope wave height. Conversely, the assumption made to derive the model maximum wave height might not be fully adequate as discussed in JB09.

## 4 Conclusions.

This paper describes an update of the ECMWF freak wave warning system and its first, still preliminary validation against observations of maximum wave height. This version became operational in June 2008.

The freak wave warning system has been extended by including effects of directionality in the estimation of the kurtosis of the surface elevation pdf, while also the contribution of bound waves to the kurtosis has been introduced. Furthermore, a parametrisation of shallow water effects in the kurtosis calculation has been introduced. Next, two new output parameters were introduced, namely maximum wave height and the corresponding period, which provide some simple measures for extreme sea states. The maximum wave height pdf, which includes nonlinear effects, was obtained following the work of Mori and Janssen (2006).

A preliminary validation of the maximum wave height product was performed as well. The present system is capable of giving realistic estimates of extreme ocean wave events. However, because of the nature of these events, only probabilistic statements can be issued. This is evident from the validation of the modelled maximum wave height distribution function against individual observed events as a random draw from the theoretical pdf was required in order to get a good match with the observed pdf.

The main output of the warning system is the expectation value of maximum wave height over a three hour time interval. Unfortunately, we cannot validate the quality of this parameter as no observations of the expectation value over a three hour interval are available to us. Nevertheless, one can make the compromise to consider the expectation value of normalised maximum wave height

over the much shorter period of 30 mins. The observed estimate for the expectation value of maximum wave height now follows from the average of the 6 successive observations (rather than taking the maximum of the 6 observations as done in section 3.2). Again it is suspected that correlation effects are relevant and therefore the number of degrees of freedom in the model pdf is reduced by 40%. This choice provides an unbiased estimate of modelled maximum wave height. The resulting comparison between modelled and observed maximum wave height is shown in Fig. 8 while the comparison between modelled and observed geophysical distribution of normalised maximum wave height is shown in Fig. 9. Again for the much shorter time series there is a good agreement between modelled and observed maximum wave height, while, as expected, the averaging procedure applied to the observations results in a much sharper geophysical distribution function. No doubt, if there would have been more independent observations available at the relevant synoptic times this would have resulted in an even sharper distribution function. Therefore, the expectation value of maximum wave height over the shorter time interval seems to be a valuable product, and by extrapolation it is expected that the same holds true for the present operational product, which is the expectation value of maximum wave height over a three hour interval.

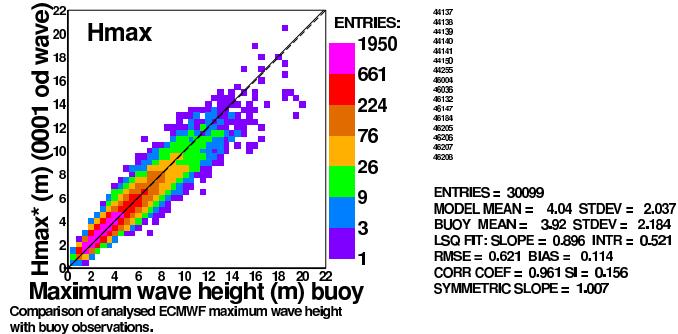


Figure 8: Comparison of observed and modelled expectation value of maximum wave height. Time interval for the model pdf is 18 mins, consistent with a 40% reduction of the number of degrees of freedom. Period is February 2006 until January 2008.

For a first validation, it is believed that some promising results have been obtained. Nevertheless, a number of issues need to be clarified. For example, the effects of correlation between successive waves on the probability distribution function of maximum wave height have to be estimated. Presently it is assumed that two wave events are not correlated, but this assumption is hard to justify as the correlation between two successive waves may be of the order of 50 %. However, to estimate effects of correlation is not a trivial task. A first step was taken by Kimura (1980) and Longuet-Higgins (1984) who, following the work of Uhlenbeck (1943) and Rice (1945), studied the joint probability distribution  $p(\rho_1, \rho_2)$  of the envelope  $\rho_1$  at time  $t$  and the envelope  $\rho_2$  at time  $t + \tau$  and its dependence on correlation. One of the interesting conclusions from their work is that for finite correlation  $\kappa$  the variance of the pdf, usually given by  $m_0$ , is reduced by the factor  $\sqrt{1 - \kappa^2}$ . Although the effect of correlation is only

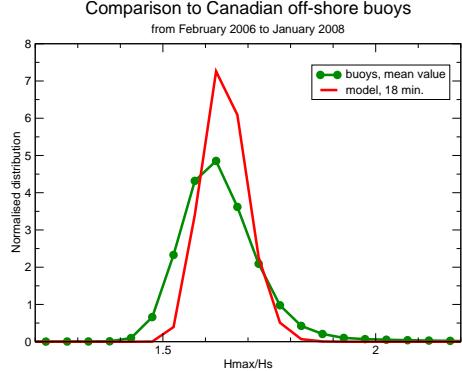


Figure 9: Comparison of observed and modelled expectation value of normalised maximum wave height distribution. The model time interval is 18 mins.

of second order, this still may give a considerable shift in the maximum wave height pdf of the order of 5 – 10% towards lower normalised maximum wave height. The task to estimate effects of correlation is, unfortunately, nontrivial as the joint pdf of  $N - 1$  somewhat correlated events is required.

Furthermore, it is required to study in what manner the Gram-Charlier expansion for the pdf of the surface elevation may be extended into the regime of very extreme events. The Gram-Charlier expansion is an expansion of the pdf in terms of the Gaussian distribution and its derivatives. Although this set of basis functions is orthogonal it is by no means certain that this gives a uniformly valid expansion for extreme values. Furthermore, for large values of the kurtosis the pdf may become negative, which is a highly undesirable property of the expansion.

Also, and this is work still in progress, more realistic estimates of the canonical part of the kurtosis need to be developed. Presently, the narrow-band approximation is used where the canonical part of the kurtosis is given by  $6\epsilon^2$  (see Eq. (10)), but it is already known from Janssen (2009) that for realistic spectra the contribution of bound waves to the kurtosis may increase by a factor of two.

Finally, according to the buoy observations there are freak waves in 8.5% of the cases, while according to nonlinear theory there are freak waves in 7.5 % of the cases. This does not imply, of course, that this is the frequency of “monster waves” as one still needs to multiply this number by the frequency of occurrence of large significant wave height events. Adopting as criterion of an extreme event that significant wave height should be larger than 8 m, then according to the available information from altimeter satellite data and first-guess wave model results the probability that on a global scale significant wave height is larger than 8 m equals 0.003. Therefore, the probability of having “monster waves” somewhere on the globe is about 0.00024.

## References

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

- Bidlot J.-R., P.A.E.M. Janssen, and S. Abdalla, 2005: On the importance of spectral wave observations in the continued development of global wave models. *Proc. Fifth Int. Symposium on Ocean Wave Measurement and Analysis WAVES2005*, 3rd-7th July 2005, Madrid, Spain.
- Bidlot J.-R., J.-G. Li, P. Wittmann, M. Faucher, H. Chen, J.-M, Lefevre, T. Bruns, D. Greenslade, F. Arduin, N. Kohno, S. Park and M. Gomez, 2007: Inter-Comparison of Operational Wave Forecasting Systems. *Proc. 10th International Workshop on Wave Hindcasting and Forecasting and Coastal Hazard Symposium*, North Shore, Oahu, Hawaii, November 11-16, 2007.
- Burgers, G., F. Koek, H. de Vries and M. Stam, 2008. Searching for factors that limit observed extreme maximum wave height distributions in the North Sea. *Extreme Ocean Waves*, E. Pelinovsky and C. Kharif (eds), Springer Science+Business Media B.V., pp. 127-138.
- Crawford, D.R., B.M. Lake, P.G. Saffman and H.C. Yuen, 1981. Stability of weakly nonlinear deep-water waves in two and three dimensions. *J. Fluid Mech.* **105**, 177-191.
- Dean, R.G., 1990. Freak waves: A possible explanation. In A. Torum & O.T. Gudmestad (Eds.), *Water Wave Kinematics* (pp. 609-612), Kluwer.
- Draper, L., 1965. 'Freak' ocean waves. *Marine Observer* **35**, 193-195.
- Janssen, P.A.E.M., 2003. Nonlinear Four-Wave Interactions and Freak Waves. *J. Phys. Oceanogr.* **33**, 863-884.
- Janssen, P.A.E.M., 2009. Some consequences of the canonical transformation in the Hamiltonian theory of water waves. *J. Fluid Mech.*, doi:10.1017/S0022112009008131
- Janssen, P.A.E.M., and E. Bouws, 1986. On the minimum width of a gravity wave spectrum, KNMI-OO Memorandum OO-86-01, KNMI, De Bilt, The Netherlands.
- Janssen, P.A.E.M., and M. Onorato, 2007. The Intermediate Water Depth Limit of the Zakharov Equation and Consequences for Wave Prediction. *J. Phys. Oceanogr.* **37**, 2389-2400.
- Janssen P.A.E.M., and J.-R. Bidlot, 2009: On the extension of the freak wave warning system and its verification. *ECMWF Tech. Memo.*, **588**. ECMWF, Reading, U.K., 42 pp., available online at: <http://www.ecmwf.int/publications/>
- Kimura, A., 1980. Statistical properties of random wave groups. *Proc. 17th Int. Conf. on Coastal Engng, Sydney*, pp. 2955-2973. New York: Am. Soc. Civ. Engrs.
- Lake B.M., H.C Yuen, H. Rungaldier, and W.E. Ferguson, 1977. Nonlinear deep-water waves: Theory and experiment. Part 2. Evolution of a continuous wave train, *J. Fluid Mech.* **83**, 49-74.
- Longuet-Higgins, M.S., 1957. The statistical analysis of a random, moving surface, *Phil. Trans. R. Soc. Lond. A* **249**, 321-387.
- Longuet-Higgins, M.S., 1983. On the joint distribution of wave periods and amplitudes in a random wave field. *Proc. Roy. Soc. London A* **389**, 241-258.
- Longuet-Higgins, M.S., 1984. Statistical properties of wave groups in a random sea state. *Phil. Trans. R. Soc. Lond. A* **312**, 219-250.
- Mori N. and P.A.E.M. Janssen, 2006. On kurtosis and occurrence probability of freak waves. *J. Phys. Oceanogr.* **36**, 1471-1483.
- Onorato, M., T. Waseda, A. Toffoli, L. Cavalieri, O. Gramstad, P.A.E.M. Janssen, T. Kinoshita, J. Monbaliu, N. Mori, A.R. Osborne, M. Serio, C.T. Stansberg, H. Tamura, and K. Trulsen, 2009. Statistical properties of directional ocean waves:

- the role of the modulational instability in the formation of extreme events. *Phys. Rev. Lett.* **102**, 114502/1-4.
- Osborne, A.R., M. Onorato, and M. Serio, 2000. The nonlinear dynamics of rogue waves and holes in deep water gravity wave trains. *Phys. Lett. A* **275**, 386-393.
- Rice, S.O., 1945. The mathematical analysis of random noise. *Bell Syst. Tech. J.* **24**, 46-156.
- Stansell, P., 2005. Distributions of extreme wave, crest and trough heights measured in the North Sea. *Ocean Engineering* **32**, 1015-1036.
- Trulsen K., and K. Dysthe, 1997. Freak Waves-A Three-dimensional Wave Simulation. in *Proceedings of the 21st Symposium on naval Hydrodynamics*(National Academy Press), pp 550-558.
- Uhlenbeck, G.E., 1943. Theory of random process. M.I.T. Radiation Lab. Rep. 454, October 1943.
- Wolfram J., and B. Linfoot, 2000. Some experiences in estimating long and short term statistics for extreme waves in the North Sea. Abstract for the *Rogue waves 2000 workshop*, Ifremer, Brest.
- Yasuda, T., N. Mori, and K. Ito, 1992. Freak waves in a unidirectional wave train and their kinematics. *Proc. 23rd Int. Conf. on Coastal Engineering*, Vol. 1, Venice, Italy, American Society of Civil Engineers, 751-764.
- Zakharov, V.E., 1968. Stability of periodic waves of finite amplitude on the surface of a deep fluid. *J. Appl. Mech. Techn. Phys.* **9**, 190-194.