CHALMERS | UNIVERSITY OF GOTHENBURG

PREPRINT 2009:41

Note on the Distribution of Extreme Wave Crests

ANASTASSIA BAXEVANI
OSKAR HAGBERG
IGOR RYCHLIK

Department of Mathematical Sciences
Division of Mathematical Statistics
CHALMERS UNIVERSITY OF TECHNOLOGY
UNIVERSITY OF GOTHENBURG
Göteborg Sweden 2009

Preprint 2009:41

Note on the Distribution of Extreme Wave Crests

Anastassia Baxevani, Oskar Hagberg, Igor Rychlik

Department of Mathematical Sciences
Division of Mathematical Statistics
Chalmers University of Technology and University of Gothenburg
SE-412 96 Göteborg, Sweden
Göteborg, October 2009

Preprint 2009:41
ISSN 1652-9715

Matematiska vetenskaper
Göteborg 2009

NOTE ON THE DISTRIBUTION OF EXTREME WAVE CRESTS

ANASTASSIA BAXEVANI, OSKAR HAGBERG, AND IGOR RYCHLIK

Abstract. The sea elevation at a fixed point is modeled by means of a second order model, which is a smooth algebraic function of a vector valued Gaussian process. Asymptotic methods, presented first in [1], are used to estimate the mean upcrossing intensity $\mu^+(h)$. The intensity is then used to determine the density of crest height in a second order sea. Numerical examples illustrate the method. The proposed approximation is used to estimate the design crest height for a specified return period.

Introduction

It is a common practice in oceanography, to model the sea surface elevation at a fixed point as a Gaussian process which, during a limited period of time (20 min - 3 hours), can also be considered stationary. The statistical properties of the sea surface elevation under stationary conditions are called a *sea state*. The sea state in the case of an underlying Gaussian process is identified by the power spectral density S, the mean sea level m, called still water level and usually set to be zero, and the water depth d. In the Gaussian model the individual cosine wave trains superimpose linearly (add) without interaction and therefore the model is also called the *linear sea model*.

However, it is well known that for steep waves in deep water, or as the water depth decreases, the sea surface profile departs from the Gaussian assumption. Under these conditions the wave profile becomes asymmetric with higher and steeper crests and shallower and flatter troughs, due to interaction between individual cosine waves. Consequently, the linear model can lead to underestimation of wave crests which increases in severity as the wave energy increases.

Adding a quadratic correction term to the linear model results to a more accurate description of the wave asymmetry. The resulting process presented in Appendix I, is non Gaussian and usually referred to as a second order sea model. A detailed description of the model can be found in Appendix II. The probabilistic properties of both the linear and second order model, and hence of the sea states, are uniquely determined by the power spectral density S (which is identical for both models), the mean water level m and the water depth d. In the case the second order correction term cannot be neglected, the estimation of the spectral density S using wave measurements becomes a non-trivial mathematical problem, since the effect of the quadratic term has to be subtracted from the measured signal before estimating the spectrum, for a detailed description of the procedure see [3].

In reliability analysis of ocean structures when studying the fatigue damage accumulated by a material during stationary weather conditions (sea state), the distribution of wave

Key words and phrases. Crest distribution, non-Gaussian sea, Rice's formula, FORM-, SORM-method.

crest height, A_c, is required. When the observation period consists of more than one sea states, the total fatigue damage is defined as the damage accumulated over the different sea states. Hence, it is sufficient to develop a method for estimating the distribution of A_c under stationary conditions. Although the exact form of this distribution is not known, in the case of a Gaussian sea, it is well approximated by means of the Rayleigh distribution. A discussion on the distribution of A_c is presented in section 1.

The design crest is defined as the wave crest that is exceeded in one year with very low probability of the order of 10^{-2} or 10^{-4}. Estimation of the design crest involves knowledge of the variability of the sea states over long periods of time. This is presented in section 2. Finally, in section 3, we present a numerical illustration of the different approximation methods in the estimation of the design wave.

1. Crest height distribution

An apparent wave is defined as the part of the sea record between two successive upcrossings of the still water level, m. The wave crest height, A_c, is the maximum of the apparent wave, i.e., is the maximum value the process $X(t)$ attains between two successive upcrossings of the level m. A discussion on the distribution of different characteristics of apparent waves can be found in [23].

Knowledge of the distribution of wave crest height, when the wave spectrum and water depth are known, is essential to a variety of problems. Unfortunately though, there is still considerable uncertainty concerning the form of the A_c distribution. The long-run distribution of A_c is defined as the histogram of the observed A_c's as the observational time increases and although it is not so difficult to estimate the empirical distribution of A_c using wave crest measurements, the empirical evidence tends to be confusing since different instruments provide with different results.

On the other hand, the theoretical derivation of the distribution of A_c is a difficult problem especially when the non-linearities have to be considered. Despite the difficulties, several approximations do exist and we present some of them in this section. For simplicity we assume that the still water level equals zero, i.e. $m = 0$.

We begin with a general result valid for any stationary random sea model $X(t)$. In [27], it has been shown that the following relation is valid

$$P(A_c > h|S) \leq \frac{\mu^+(h|S)}{\mu^+(0|S)}, \tag{1}$$

where $\mu^+(h|S)$ is the upcrossing intensity of the level h during the sea state S and with a slight abuse of notation we can write $P(A_c > h|S)$ for the dependence of the crest distribution on the sea state. The intensity $\mu^+(h|S)$ can be computed using the Rice formula (5). Here we should note that although this bound is true for all stationary random processes, is particularly accurate for non clustering upcrossings of high levels h.

1.1. Rayleigh Bound for Gaussian sea (linear case).

Consider a sea state S, that is characterized by the significant wave height, h_s, and the average wave period, t_z. Assume also that the sea elevation is a Gaussian process with mean value zero. Then it is well known that the upcrossing intensity $\mu^+(h|S)$ can be computed by means of the Rice formula, [26] to give

$$\mu^+(h|S) = \frac{1}{t_z} e^{-8\left(\frac{h}{h_s}\right)^2}. \tag{2}$$

Inserting (2) in (1) we obtain the following bound

$$P(A_c > h|S) \leq e^{-8\left(\frac{h}{h_s}\right)^2}, \tag{3}$$

which is very accurate for high levels h. A direct consequence of (3), is the following, rather conservative, Rayleigh approximation for the wave crest height $A_c \approx \frac{h_s}{4} \cdot R$, where R is distributed as a Rayleigh random variable ($P(R > r) = \exp(-r^2/2)$). This approximation is widely used in the oceanographic literature, and is motivated by means of a heuristic argument for narrow band processes.

1.2. **Approximations of crest height distribution for second order sea.** In a second order sea, the interaction between individual waves is non negligible and the Rayleigh bound discussed in section 2.1, is often non conservative.

In the last decade, the crest height distribution for second order seas has been the subject of intensive studies. Both [9] and [24] have developed models for the crest height distribution using long simulations of the second order sea at a fixed point. The resulting distributions are parametric, depend on the spectrum S and as reported in [16] and [10] give accurate approximations in many cases. A review of the different approximation methods can be found in [25]. An alternative approach using the 5^{th} order Stokes correction and a narrow band argument, is presented in [6].

1.2.1. *Forristall's approximation.* In this presentation of Forristall's model, we follow the approach adopted in [16]. Forristall's model is based on a Monte-Carlo study, using the second order sea model, (see [9] and references therein) and statistical fitting of the Weibull distribution to the resulting record of wave crests. To be specific, Forristall simulated the ocean surface at a fixed point using a directional spectrum $S(\omega, \theta)$ and fitted a two-parameter Weibull distribution,

$$P(A_c > h|S) = e^{-\left(\frac{h}{a h_s}\right)^b}, \tag{4}$$

to the wave crest heights extracted from the simulation record. The fitted distribution obviously depends on the sea state which is characterized by the spectrum. The parameters a and b are given in terms of s_1, which is a measure of steepness and the Ursel number U_r, which is a measure of the impact of water depth on the non-linearity of waves.

In the case of a second order long-crested (uni-directional) sea, Forristall postulated that

$$\begin{aligned} a &= 1/\sqrt{8} + 0.2892 \cdot s_1 + 0.106 \cdot U_r, \\ b &= 2 - 2.1597 \cdot s_1 + 0.0968 \cdot U_r^2, \end{aligned}$$

where

$$U_r = \frac{g^2 \lambda_0^2 h_s}{\lambda_1^2 d^3}, \quad s_1 = \frac{4}{2\pi\, g} \frac{\lambda_1}{\sqrt{\lambda_0}}.$$

In the case of a second order sea with a directional spectrum the parameters a and b read

$$\begin{aligned} a &= 1/\sqrt{8} + 0.2568 \cdot s_1 + 0.08 \cdot U_r, \\ b &= 2 - 1.7912 \cdot s_1 - 0.5302 \cdot U_r + 0.2824 \cdot U_r^2. \end{aligned}$$

Note that for $a = 1/\sqrt{8}$ and $b = 2$, the Weibull distribution in (4) reduces to the Rayleigh distribution. Concluding, the Weibull distribution parameters depend on the spectrum S in a rather complicated fashion, although the two spectral moments λ_0, λ_1 are sufficient for computing the distribution.

Forristall's model is similar to the model proposed by Prevosto, see [25]. Due to somewhat different parameterizations, the comparison between the two models is not straightforward, but they appear to give similar results.

1.2.2. *Dawson's approximation.* An alternative approximation based on the Rayleigh law and the 5th order Stokes correction is presented in [6].

$$P(A_c > h|S) = \exp(-8x^2 + 8rx^3 - 4r^2x^4 + 14r^3x^5/3 - 117r^4x^6/24),$$

with

$$x = h/h_s$$

and $r = \left(\frac{2\pi}{t_z}\right)^2 h_s/g$ be the characteristic steepness.

1.2.3. *Tails of crest distribution for second order sea.* The problem of approximating the tails of the crest height distribution is actually equivalent to the problem of estimating the upcrossing intensity, for which as we show next, is sufficient to know the joint density $f_{X(0),\dot{X}(0)}(h,z)$.

In the case of a stationary process, the upcrossing intensity $\mu^+(h|S)$ is computed using the Rice formula; see [18] for a proof. That is,

$$\mu^+(h|S) = \int_0^{+\infty} z f_{X(0),\dot{X}(0)}(h,z)\,dz, \tag{5}$$

where $f_{X(0),\dot{X}(0)}(h,z)$ denotes the joint density of $X(0), \dot{X}(0)$.

Estimation of the joint density $f_{X(0),\dot{X}(0)}$ is rather difficult, since for most of the non Gaussian processes this density does not have an explicit form. In the case of a second order process the joint density exists; it is a function of the vector valued Gaussian process characterized by the sea state S. Several methods have been proposed for computing this joint density, see [22], [5], [19], [13]. All these methods rely on the fact that the moment generating function of the joint density has an explicit algebraic expression and hence an inverse Fourier transform may be used to evaluate the density.

In this work, we are mainly interested in the crest distribution for high levels h. Hence, the asymptotic method, called the Breitung method, proposed in [1] and further investigated in [12] and [13] is employed. A related method given in [21], provides with an asymptotic formula for $\mu^+(h|S)$ as $h \to \infty$, opposed to the Breitung method which provides with an asymptotic result only in the extreme cases of either a purely linear or a purely second-order process. However, as argued in Appendix III the Breitung method probably balances better the influences of the linear and second-order terms than the asymptotic results derived in [21].

In Appendix III we show that a good approximation for the upcrossing intensity is the following,

$$\mu^+(h|S) \approx c(\beta_h)\exp(-\beta_h^2/2). \tag{6}$$

An explicit formula for the constant $c(\beta)$ can also be found in Appendix III. The increasing function β_h is the so-called Hasofer-Linds safety index and is estimated using standard numerical methods, see Appendix III for a more detailed discussion.

Hence, (1) can be written as

$$P(A_c > h|S) \approx \frac{c(\beta_h)}{\mu^+(0|S)}\exp(-\beta_h^2/2). \tag{7}$$

The computations of the factor $c(\beta)$ involve the curvature of the response surface at the design point, and hence the approximation in (7) is called SORM (Second-Order-Reliability-Method). For the estimation of the zero-upcrossing intensity $\mu^+(0|S)$, there are a few options. We may use Monte-Carlo simulations or the saddle point method or other approximations. Here, we have decided to adopt an alternative approach. It is our experience, that the wave intensity in a second order sea does not differ much from the wave intensity in the Gaussian sea, hence we may set $\mu^+(0|S) \approx 1/t_z$. Hence, (7) becomes

$$P(A_c > h|S) \approx t_z c(\beta_h)\exp(-\beta_h^2/2). \tag{8}$$

For a second order sea, $c(\beta)$ for positive values of β and h, is close to $1/t_z$ and often $|c(\beta_h)t_z - 1| < 0.2$. By replacing $t_z c(\beta_h)$ by 1 we obtain the somewhat simpler approximation

$$P(A_c > h|S) \approx \exp(-\beta_h^2/2). \tag{9}$$

This type of approximation was also used in [29]. Since this approximation depends only on the index β_h, it is called the FORM (First-Order-Reliability-Method) approximation.

In the examples presented in section 2.2.4, the differences between the FORM and SORM approximations are negligible. However, since application of the SORM method is not numerically more intensive, it should be preferred over the FORM method, since it is never known if there exist a special type of sea state for which $c(\beta)$ is not close to $1/t_z$, which would lead to considerable differences between the two approaches.

Comparison of accuracy of the tail approximations for crest height distribution. We turn now to a comparison of the tails of the proposed distributions. We assume that the tails of the distribution computed using the SORM method are the *true* ones.

We consider four different sea states, two uni-directional and two directional. The frequency spectrum $S(\omega)$ used in all different sea states is from the Pierson-Moskowitz family. We consider two different defining sets of parameters, $(h_s, t_p, \gamma, \sigma_a, \sigma_b) = (20, 16.8, 1, 0.07, 0.09)$ and $(h_s, t_p, \gamma, \sigma_a, \sigma_b) = (24, 18, 1, 0.07, 0.09)$. The spreading function $G(\omega, \theta)$, is of the $\cos 2s$-type with parameter $s = 10$. The resulting tail distributions are gathered in Fig. 2. The top plots are computed using the spectrum defined by the first set of parameters and the bottom

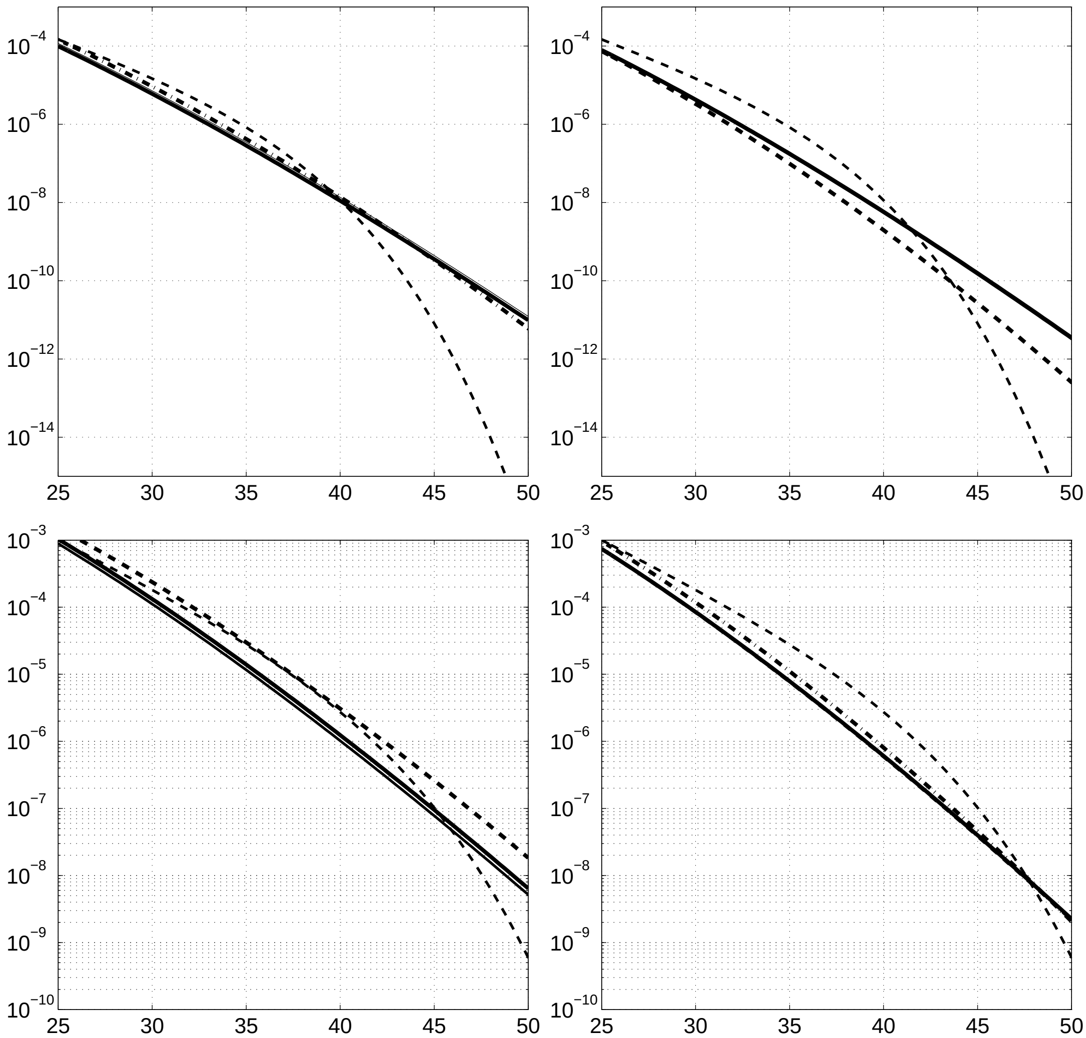

FIGURE 1. *Solid lines are for the FORM and SORM methods, light dashed line is for the Dawson model, and thick dashed line is for the Forristall model.*

plots using the second set of parameters. The first and third plots are for the uni-directional sea and the second and fourth for the directional one.

The tails of the distributions computed using the FORM and SORM methods (solid lines) almost coincide. The dashed lighter line is for the Dawson model. This model is based on the 5th order Stokes correction and as a result the line is bending down. Finally the thicker dashed line is the distribution computed using the Forristall model. The tails of this distribution seem to follow reasonably well the solid lines and taking into account that we are looking far into the tails, i.e., long outside the region at which the Forristall model was fitted to the data, we may conclude that the Forristall model gives quite accurate results.

2. Design wave

An important parameter regarding safety regulations is the height of the maximum wave crest over a certain period of time and at a specific location or over some region. For example, Norwegian regulations require that the probability of a wave reaching the deck of an oil rig over a year does not exceed 10^{-4}. Of course, one should also account for the wave-structure interactions, which may lead to an increase of the wave crest height of the order 20-30%, for tide and storm surge, which under certain conditions may add a few meters to the crest height, see [15]. The contribution of most such effects may be estimated with great accuracy by means of rather simple statistical methods. However, the most important quantity that has to be estimated correctly is the wave crest height. In this section we investigate the problem of predicting the response value h^{crt}, corresponding to an annual exceedance probability p_0, i.e. we wish to solve the following equation

$$P(max_{0\le t\le T}X(t) > h^{crt}) = p_0$$

for h^{crt}. If $p_0 = 10^{-4}$, the level h^{crt} is called the 10 000 year wave crest.

For small probabilities p_0, the critical level h^{crt} takes large values and therefore at least in principle standard statistical methods may be employed. The Peaks Over Threshold (POT) method, requires equidistant measurements (for example, daily) of the maximum crest height over a large period of time. Such measurements are not available for most of the regions which makes the POT method difficult to apply. Another approach consists in assuming the wave crests are independent, so we may write,

$$P(max_{0\le t\le T}X(t) \le h) \approx P(A_1^c \le h)P(A_2^c \le h)\cdots P(A_N^c \le h), \tag{10}$$

where A_i^c denotes the crest height of the i^{th} wave and N is the average number of waves during the year. Obviously, the A_i^c's are not identically distributed. We turn now to a new method for estimating the design wave.

2.1. **Distribution of design crest - Rice method.** Let $N_T^+(h)$ be the number of upcrossings of the level h by the process $X(t)$ during the time period $[0,T]$. It is easy to see that for any fixed time $t_0 \in [0,T]$,

$$P(\max_{0\le t\le T} X(t) > h) = P(X(t_0) > h) + P(N_T^+(h) > 0, X(t_0) \le h)$$

and since $P(N_T^+(h) > 0, X(t_0) \le h) \le P(N_T^+(h) > 0) \le \mathsf{E}[N_T^+(h)]$ we also have that

$$P(\max_{0\le t\le T} X(t) > h) \le P(X(t_0) > h) + \mathsf{E}[N_T^+(h)]. \tag{11}$$

If $X(t)$ was stationary during the period T, then $P(X(t_0) > h) = P(X(0) > h)$ is negligible compared to $\mathsf{E}[N_T^+(h)] = T\mu^+(h)$ and hence we can write $P(\max_{0\le t\le T} X(t) > h) \le T\mu^+(h)$.
$\mathsf{E}[N_T^+(h)]$ and $P(X(t) > h)$.

The Rice formula provides us with the tools to compute

$$\mathsf{E}[N_T^+(h)] = \int_0^T \mu_t^+(h)\,dt = \int_0^T \int_0^{+\infty} z f_{X(t),\dot X(t)}(h,z)\,dz\,dt, \tag{12}$$

where $N_T^+(h)$ is the number of upcrossings of the level h during $[0,T]$ and $\mu_t^+(h)$ is the intensity of upcrossings of the level h by the process $X(t)$. The last equality is obviously true when the joint density of $X(t)$ and $\dot X(t)$ exists. The joint density of the process and its derivative, $f_{X(t),\dot X(t)}(h,z)$, includes both sources of variability of the sea surface, the variable sizes of the sea waves during a sea state as well as the evolution of the sea states with time.

If we denote by $S_t = S_t(\omega, \theta)$ the random sequence of sea states and assume the conditional density $f_{X(t),\dot{X}(t)|S_t}(h, z)$, of the process $X(t)$ and its derivative $\dot{X}(t)$ on the sea state at time t, S_t, is well defined, we may write

$$\int_0^{+\infty} z f_{X(t),\dot{X}(t)}(h, z)\, dz = \mathsf{E}[\int_0^{+\infty} z f_{X(t),\dot{X}(t)|S_t}(h, z)\, dz], \tag{13}$$

where the expectation is taken over the whole sea state sequence. However since the change rate of the sea state process is much slower than of the sea elevation, we may locally approximate the density $f_{X(t),\dot{X}(t)|S_t}(h, z)$ by that of a second order sea. Therefore,

$$\int_0^{+\infty} z f_{X(t),\dot{X}(t)|S_t}(h, z)\, dz \approx \mu^+(h|S_t),$$

where $\mu^+(h|S_t)$ was defined in (5), and consequently $\mu_t^+(h) \approx \mathsf{E}[\mu^+(h|S_t)]$. The intensity $\mu^+(h|S_t)$ is estimated using (6).

Even after having estimated $\mu^+(h|S_t)$, it is not obvious how to evaluate $E[\mu^+(h|S_t)]$. One possibility is to model the sea state S_t in a parametric fashion, say S_t depends on a vector of parameters $\mathbf{a} = (a_1, ..., a_n)$. Then, the sea state process can be identified to the evolution of the parameter vector, i.e. the vector valued process $\mathbf{a}(t) = \{a_1(t), \ldots, a_n(t)\}$. Formally we may write

$$\mu^+(h|S_t) = \mu^+(h|a_1(t), \ldots, a_n(t)).$$

If additionally assume that the joint probability density of the vector process$\mathbf{a}(t)$ exists and denote it by $f_t(a_1, \ldots, a_n)$, the following is true,

$$\mu_t^+(h) \approx \mathsf{E}[\mu^+(h|S_t)] = \int \mu^+(h|a_1, \ldots, a_n) f_t(a_1, \ldots, a_n)\, da_1 \ldots da_n. \tag{14}$$

A change in the order of integration in (6) and (14) allows us to write,

$$\begin{aligned} \mathsf{E}[N_T^+(h)] &= \int_0^T \mu_t^+(h)\, dt = \\ &= T \int \mu^+(h|a_1, \ldots, a_n) \left(\frac{1}{T} \int_0^T f_t(a_1, \ldots, a_n)\, dt \right) da_1 \ldots da_n = \\ &= T \int \mu^+(h|a_1, \ldots, a_n) f(a_1, \ldots, a_n)\, da_1 \ldots da_n, \end{aligned} \tag{15}$$

where the density

$$f(a_1, \ldots, a_n) = \frac{1}{T} \int_0^T f_t(a_1, \ldots, a_n)\, dt \tag{16}$$

describes the variability of the parameters that define the spectrum S_t at an arbitrary time t of the year.

Similarly, for any fixed t, we may write

$$P(X(t) > h) \approx \int P(X(t) > h|a_1, \ldots, a_n) f_t(a_1, \ldots, a_n)\, da_1 \ldots da_n. \tag{17}$$

Since the process $X(t)$ conditionally on the sea state process $a_1(t), \dots, a_n(t)$ is a second order process the probability (17) can be estimated using the FORM approximation, e.g. formula (9). Indeed,

$$P(X(t) > h|a_1, \dots, a_n) \approx 1 - \Phi(\beta_h) \leq \frac{1}{\beta_h \sqrt{2\pi}} \exp(-\beta_h^2/2). \tag{18}$$

2.2. **3-hour design wave crest.** For comparison reasons we consider the so-called 3-hour maximum approach, see formula (20) in [15].

In practice a common choice of parameters is $a_1 = h_s$ and $a_2 = t_p$. Hence the joint density $f(a_1, a_2)$ defined in (16), is the so called long term density of the significant wave height and wave period, $f(h_s, t_p)$. Let A^c_{3h} be the maximum crest height observed during a stationary period 3 hours long and under the assumption the wave crest distribution is uniquely characterized by h_s and t_p. A common practice is to assume that the wave crests are independent, see (10),

$$P(A^c_{3h} \leq h|h_s, t_p) \approx P(A^c \leq h|h_s, t_p)^N, \quad N = 3 \cdot 3600 \cdot \mu^+(0|h_s, t_p),$$

where N is the expected number of waves in a 3 hour period. Now assuming the 3-hour period was chosen at random

$$P(A^c_{3h} \leq h) = \int\int P(A^c \leq h|h_s, t_p)^N f(h_s, t_p)\, dh_s\, dt_p.$$

An estimate of the design wave crest h_{crt} is the solution to the equation

$$1 - P(A^c_{3h} \leq h_{crt}) = p_0/2920,$$

where 2920 is the number of 3 hour periods during the year and $p_0 = P(max_{0\leq t\leq T} X(t) > h^{crt})$.

We will now show that the presented method is asymptotically equivalent to the Rice method if the probability distribution of the wave crest is close to the bound given in (1). Indeed, it is clear that for high values of h the following approximation is valid

$$P(A^c \leq h|h_s, t_p)^N \approx 1 - N(1 - P(A^c \leq h|h_s, t_p))$$

and hence, by combining the last two equations we have the following approximation for high values of h,

$$P(\max_{0\leq t\leq T} X(t) > h) \approx T \int\int \mu^+(0|h_s, t_p) P(A^c > h|h_s, t_p) f(h_s, t_p)\, dh_s\, dt_p. \tag{19}$$

This formula will be used in the sequel to compare tails of $P(\max_{0\leq t\leq T} X(t) > h)$ derived from the distribution of individual crest heights to the method based on the Rice formula for the upcrossing intensity $\mu^+(h)$.

If the probability $P(A^c > h|h_s, t_p)$ is close to the bound given by (1), the following approximation is true

$$P(\max_{0\leq t\leq T} X(t) > h) \approx T \int\int \mu^+(h|h_s, t_p) f(h_s, t_p)\, dh_s\, dt_p. \tag{20}$$

3. Model application

In this section we will illustrate the computations needed to estimate the crest height of the **design wave**, and compare the accuracy of the different methods.

3.1. Sea state parametrization.

Let us set the water depth d at 500 meters, which in practice is equivalent to deep water. In order to fully characterize the sea state we need to specify the spectral densities encountered at each location. In the case of a Gaussian sea, the spectrum and hence the sea state is fully characterized by the significant wave height h_s and the average wave period t_z. We turn now to the description of a family of spectral densities that will be used in this example for describing the encountered sea states.

Let $S(\omega)$ denote a JONSWAP frequency spectrum. This is a parametric spectrum that is fully characterized by the following set of parameters:

$$(\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5) = (h_s, t_p, \gamma, \sigma_a, \sigma_b).$$

The parameter γ is a factor determining the concentration of the spectrum around the peak frequency, t_p, and depends on both h_s and t_p. The parameters σ_a and σ_b, are spectral width parameters and in this case are set to be $\sigma_a = 0.07$ and $\sigma_b = 0.09$. For the directional spectrum we need a spreading function $G(\omega, \theta)$ of the $\cos 2s$-type with parameter $s = 10$. Here we should emphasize that this particular choice of parameters is for illustration reasons only.

If the joint long run density $f(h_s, t_p)$ is known, the sea state distribution can be determined. Modeling the joint density of h_s, t_p has been the subject of elaborated research by several authors, see e.g. [15] or [20] and references therein. Most often though, we have instead of $f(h_s, t_p)$, an estimate of the joint density of h_s and t_z. Recovering the joint density of h_s, t_p although possible is not trivial.

For a JONSWAP spectrum, the peak period can be evaluated by means of the following relation

$$t_p = t_z \cdot (1.30301 - 0.01698 \cdot \gamma + 0.12102/\gamma). \tag{21}$$

In the present example, since our intention is to illustrate the different methods and compare their accuracy, we will further simplify the sea state model by assuming a regression relation between the mean period t_z and the significant wave height h_s. Labeyrie in [17] proposed the following relation, based on data from the Frigg Field,

$$t_z = E[T_z | H_s = h_s] = \sqrt{a h_s + b}, \tag{22}$$

with $a = 8$ and $b = 21$. Krogstad in [16] proposed another relation between t_z and h_s derived from a Haltenbanken buoy data set consisting of about 12 000 data records

$$t_z = 4.27 \cdot h_s^{0.37}. \tag{23}$$

We shall consider four different spectral densities. In the first two cases the spectrum is a uni-directional spectrum and as such we employ the JONSWAP spectrum introduced previously. We use two different choices for γ. First we set $\gamma = 3.3$, and use the relation in (22) to obtain t_p as a function of h_s,

$$t_p = 1.2836\sqrt{8h_s + 21}.$$

Then, we set $\gamma = 1$, which is a special case of JONSWAP spectrum called Pierson-Moskowitz spectrum, and the relation in (23) to obtain the following relation

$$t_p = 6.008 \cdot h_s^{0.37}.$$

The other two spectra we use are directional. They are constructed as the product of the JONSWAP frequency spectrum for $\gamma = 3.3$ and $\gamma = 1$ with a spreading function of the $\cos 2s$ type. In all four cases, the spectrum is fully characterized by only one parameter, the significant wave height h_s. That is, in order to evaluate the distribution in (15), the only thing that remains to be estimated is the density $f(h_s)$. Note that we do not argue on the validity of these models, we merely wish to demonstrate the performance of the different algorithms.

3.2. **LONG TERM DISTRIBUTION FOR THE SIGNIFICANT WAVE HEIGHT.** The significant wave height random process $H_s(t)$, is often accurately described by means of log-normal models, see [2], [4] and references therein.

Estimates of the significant wave height from the TOPEX-Poseidon satellite, over an area in the North Atlantic called the North Atlantic Route (NAr) collected during the period from 1992 until 1999, were used in the study presented in [4]. The authors suggest that the variability of the significant wave height during a fixed time point, should be modeled locally as

$$\ln(H_s(t)) = \beta_0 + \beta_1 \cos(\phi t) + \beta_2 \sin(\phi t) + \sigma\epsilon(t),$$

where $\epsilon(t)$ is stationary correlated Gaussian noise with mean zero and variance one, and $\phi = 2\pi/365.2$. The parameters $\beta_0, \beta_1, \beta_2$ and σ are locally estimated and depend on the geographical location. In [4] , a collection of the model parameter estimates for the different stationary regions in the North Atlantic, is presented.

In this example we shall use the models from two different areas along the North Atlantic route, one close to the coast of North America and one close to Europe. The model parameter estimates in the location close to North America, lying between -54 and -52 degrees in longitude and between 42 and 46 degrees in latitude, are $\sigma^2 = 0.1529$, $\beta_0 = 0.8674$, $\beta_1 = 0.3836$ and $\beta_2 = 0.0635$. The location close to Europe that lies between -12.8 and -10 degree in longitude and between 48 and 52 degrees in latitude, has parameter estimates $\sigma^2 = 0.1667$, $\beta_0 = 1.0364$, $\beta_1 = 0.4294$ and $\beta_2 = 0.0652$.

3.3. **Estimation of the design wave crest.** In this section we compare the design wave crest h_{crt} for levels p_0 between 10^{-2} and 10^{-4}, estimated using the Rice approach (20)to the design crests derived using the method that depends on the crest height distribution of individual waves (19), where the crest height distributions were approximated using the methods proposed by Forristall and Dawson and the FORM approximation.

For illustration reasons, we shall compute the design wave height distribution at each location using the different spectra. For the location close to the coast of America, we shall use the two spectra derived from the Frigg field data with $\gamma = 3.3$. For the other location close to Europe, we shall use the two spectra derived from the Haltenbanken buoy and the Pierson-Moskowitz spectrum. The resulting distributions are plotted in Fig. 2. The first two plots are the distributions for the American location, while the last two plots, for the European one. The top plots are for the uni-directional (long-crested) seas, while the bottom plots are for the directional seas. Obviously the 10 000 year crest height for each case is just the 10^{-4} quantile.

As expected, the design wave based on the linear model (Rayleigh model) is much smaller than the predictions based on the second order sea model. Furthermore, for a Gaussian sea the tails of the crest height distribution when observed at a fixed location, do not depend on the shape of the spreading function. As a result the distribution for both uni-directional and directional sea is the same: compare the plots on the top with the plots on the bottom. The two solid lines are the tails of the distribution computed using the FORM and SORM methods. We shall consider the tails, computed using the SORM-method as the *true* tails. The dashed lines represent approximations of the tails computed using the Forristall and Dawson models.

Let us consider the two uni-directional spectra, i.e. the first and third plots of Fig. 2. The FORM-approximation gives slightly lower predictions that the SORM method. For example the predictions for the 10 000 year crest differ less than half meter, which of course is of no practical significance. The difference between Forristall's and Dawson's models is more profound, it gets up to a few meters on the third plot. (The bended dashed line is the approximation using Dawson's model, while the most conservative dashed tails are the Forristall's model.)

We turn now to the directional sea presented on the second and fourth plots. Since the parameters in Dawson's model do not depend on the spreading function, the resulting crest height distribution remains the same as for the uni-directional sea. The Dawson model overestimates the 10 000 year crest by more than two meters in the location close to the coast of America. Accidentally, Dawson's approximation gives a good 10 000 year crest height prediction for the European location, however the 100 year crest is overestimated by a few meters. Remarkably both the FORM and Forristall's tail approximations agree very well with the most complicated SORM method. We can also notice that the uni-directional seas produce higher crests than the directional ones, as has already been reported by several authors.

As a final remark, we would like to mention that for the computations in Forristall's and Dawson's models we have used the following definition of significant wave height, $h_s = 3.8\sqrt{\lambda_0}$, as proposed in Dawson and Wallendorf (2003). The standard definition of significant wave height, $h_s = 4\sqrt{\lambda_0}$ leads to an overestimation of the 10 000 year crest height by approximately 3 meters. Generally, it is not obvious how h_s should be defined: $h_s = 4\sqrt{\lambda_0}$ (four times standard deviation of the linear Gaussian part), or $4\sqrt{\mathsf{V}(X(0))}$ (four times the standard deviation of the full model)? Both the SORM and FORM methods are not affected by the significant wave height definition.

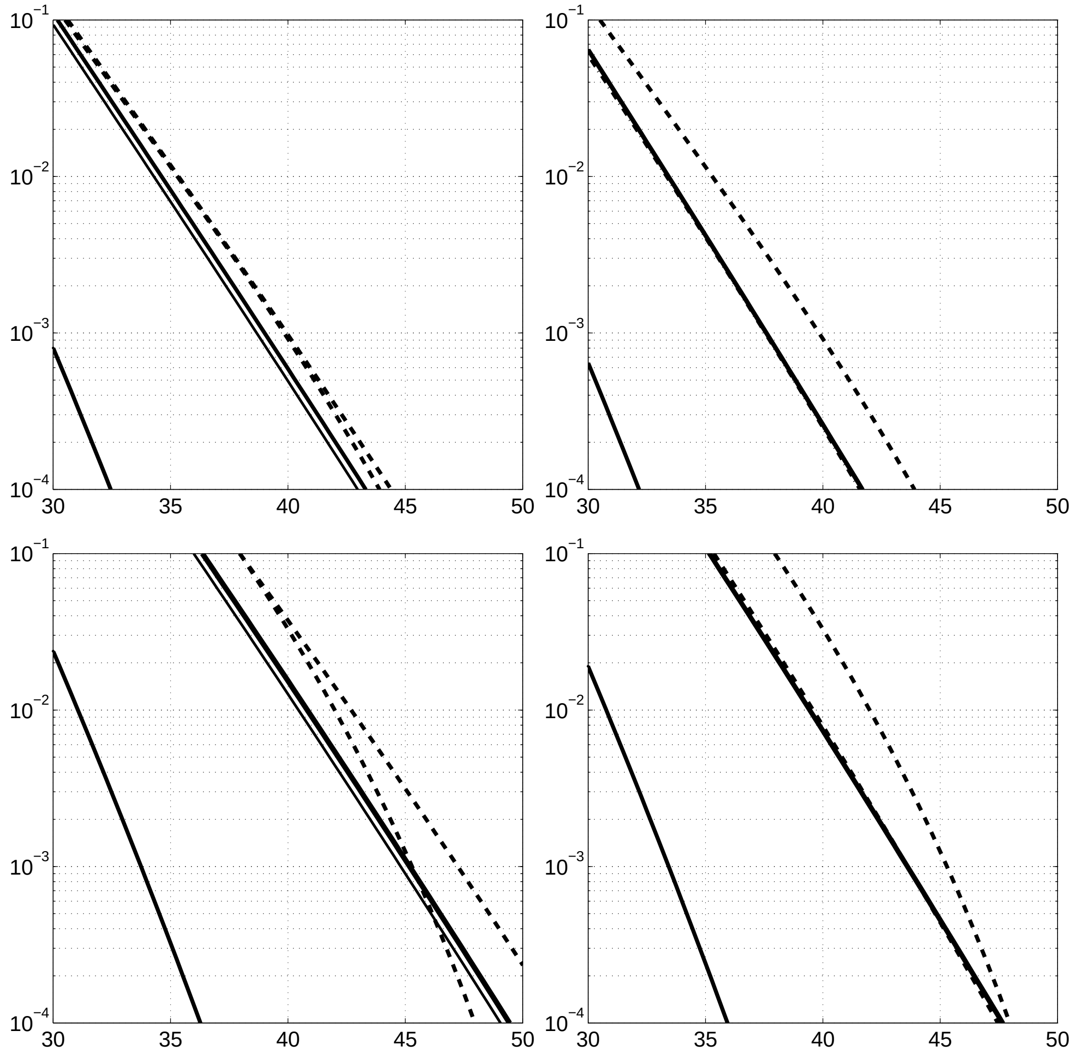

FIGURE 2. *The solid lines on the far left represent the design wave height distribution for a linear Gaussian sea. The two solid lines on the right are the tail distributions computed using the FORM and SORM methods. The two dashed lines are the approximations of the distribution tails using Forristall's and Dawson's method.*

4. Conclusions

We have presented an accurate method (SORM) for approximating the tails of the crest height distribution. This method was combined with the so called Rice method to give a methodology that was then used to find the design 100 year wave crest. This approach appears to give similar results to other methods presented in this paper. However, it is simpler and is derived without any assumptions of independence of crest heights, which makes it applicable in a wider range of situations.

References

[1] Breitung, K., (1988) Asymptotic crossing rates for stationary Gaussian vector processes. *Stochastic Processes and their Applications*, **29**, pp. 195-207.

[2] Anderson, C. W., Carter, D. J. T., Cotton, P. D., (2001) Wave Climate Variability and Impact on Offshore Design Extremes, *Report prepared for Shell International.*

[3] Mitsuyasu, H., Masuda A. and Kuo Y.-Y., (1979) The dispersion relation of random gravity waves Part 1. and Part 2., *J. Fluid Mech.*, **92**, pp. 717-730, 731-749.

[4] Baxevani, A. (2004) *Modelling Sea Surface Dynamics Using Crossing Distributions*, Lund University, PhD Thesis, Lund, Sweden

[5] Butler, R.W. and Machado, U. and Rychlik, I., (2003) Distribution of Wave Crest in non-Gaussian Sea, *Proceedings of the Thirteenth (2003) International Offshore and Polar Engineering Conference*, **3**, pp. 30-37.

[6] Dawson, T. H., (2002) Rayleigh Law and Stokes Correction for High Waves in Heavy Seas, *Proceedings of the fourth International Symposium on Ocean Wave Measurement and Analysis, San Francisco, ASCE*, **1**, pp. 484–493.

[7] Dawson, T.H. and L. A. Wallendorf, L.A., (2003) Effects on Wave Statistics of Non-Equilibrium High Frequency Waves in Random Seas, *Proceedings of The Thirteenth ISOPE Conference, Honolulu, Hawaii, ISOPE*, pp. 79–83.

[8] Ditlevsen, O. and Madsen, H.O., (1996) *Structural Reliability Methods*, J. Wiley and Sons, New York.

[9] Forristall, G. Z., (2000) Wave crest distributions: Observations and second order theory, *Journal of Physical Oceanography*, **30**, pp. 1931-1943.

[10] Forristall, G. Z., (2002) Wave Crest Sensor Intercomparison Study: An overview of WACSIS, *OMAE, Oslo.*

[11] Hagberg, O., (2004) The Rate of Crossings of a Quadratic Form of an n-Dimensional Stationary Gaussian Process, Technical Report, 2004:17, Lund University, Sweden.

[12] Hagberg, O., (2005) An Asymptotic expansion of the crossing rate of surface by a stationary Gaussian vector process, Technical Report, 2005:1, Lund University, Sweden.

[13] Hagberg, O., (2005) Saddle Point Approximation of Rice's Formula, Technical Report, 2005:2, Lund University, Sweden.

[14] Hasselmann, K., (1962) On the non-linear energy transfer in a gravity-wave spectrum, Part 1, General Theory, *Journal of Fluid Mechanics*, **12**, pp. 481-500.

[15] Haver, S., (2004) On the prediction of extreme wave crest heights, *Manuscript, Statoil, Stavanger, Norway.*

[16] Krogstad, H. and Barstow, S.F., (2004) Analysis and Applications of Second-Order Models for Maximum Crest Height, *Journal of Offshore Mechanics and Arctic Engineering*, **126**, pp. 66-71.

[17] Labeyrie, J., (1990) Time Scales and Statistical Uncertainties in the Prediction of Extreme Environmental Conditions, *Reliability Engineering and System Safety*, **32**, pp. 243-266.

[18] Leadbetter, M. R. and Lindgren, G. and Rootzén, H. (1983) *Extremes and related properties of random sequences and processes* , Springer-Verlag, New York.

[19] Machado, U.B. and Rychlik, I., (2004) Wave statistics in non-linear random sea, *Extremes*, **6**, pp. 125-146.

[20] Monbet, V. and Prevosto, M., (2000) Bivariate Simulation of Non-Stationary and Non-Gaussian Observed Processes: Application to Sea State Parameters, *Proceedings of The Tenth ISOPE Conference*, Seattle, USA, ISOPE, pp. 91-94.

[21] Naess, A., (1985) Statistical Analysis of Second-Order Response of Marine Structures, *Journal of Ship Research*, **29**, pp. 270-284.

[22] Naess, A., Machado, U.B., (2000) Response statistics of large compliant offshore structures *Proceedings of the 8th ASCE Specialty Conference on Probabilistic Mechanics and Structure Reliability, New York, USA, ASCE.*

[23] Podgórski, K. and Rychlik, I. and Machado, U. E. B., (2000) Exact Distributions for Apparent Waves in Irregular Seas, *Ocean Engineering* , **27**, pp. 979-1016.

[24] Prevosto, M. and Krogstad, H. E. and Robin, A., (2004) Probability Distributions for Maximum Wave and Crest Heights, *Coastal Engineering*, **40**, pp. 329-360.

[25] Prevosto, M., (2004) Statistics of Wave Crests from Second Order Irregular Wave 3D Models, *Rogue-Waves IFREMER.*

[26] Rice, S.O., (1944, 1945) Mathematical analysis of random noice, *Bell System Technical Journal*, **23, 24** pp. 282-332, 46-156.

[27] Rychlik I. and Leadbetter M. R., (2000) Analysis of ocean waves by crossing and oscillation intensities, *International Journal of Offshore and Polar Engineering*, **10**, pp. 282-289.

[28] Tromans, P. S., (2002) A spectral response surface method for calculating crest elevation statistics, *Proceedings of the 21st International Conference on Offshore Mechanics and Arctic Engineering, Oslo, Norway, OMAE.*

[29] Tromans, P.S. Vanderschuren L. (2004) A Spectral Response Surface Method for Calculating Crest Elevation Statistics, *Journal of Offshore Mechanics and Arctic Engineering*, **126**, pp. 51-53.

Appendix I

5. Modeling sea surface under stationary conditions

5.1. **Second order deterministic sea.** In this paper we model the sea surface elevation by a second order model, see [14] where corrections up to 5th order are given.

We begin with a linear sea model, which postulates that the sea surface is a sum of simple cosine waves. The linear part W_l, consisting of N cosine waves, is given by

$$W_l(t,\mathbf{p}) = \sum_{n=-N}^{N} \frac{A_n}{2}\, e^{i(\omega_n t - x\kappa_n\cos\theta_n - y\kappa_n \sin\theta_n)} := \sum_{n=-N}^{N} \frac{A_n}{2}\, w_n(t,\mathbf{p}), \tag{24}$$

where for each elementary wave, A_n denotes its complex amplitude ($A_{-n} = A_n^*$, the complex conjugate of A_n), with $A_0 = 0$. Since the field W_l needs to be real valued, the angular frequencies need to satisfy $\omega_{-n} = -\omega_n$. Moreover, $\kappa_{-n} = -\kappa_n$, where κ_n is the wave number corresponding to wave frequency ω_n through the *dispersion relation* ($\omega_n^2 = g\kappa_n \tanh(d\,\kappa_n)$, where d stands for the water depth and g for the gravity acceleration). We also assume that for $n > 0$, $\omega_n > 0$ and the direction of propagation θ_n satisfies $-\pi < \theta_n \le \pi$ with $\theta_{-n} = \theta_n + \pi$. Hence for $n > 0$ and $A_n = R_n e^{i\phi_n}$

$$\frac{A_n}{2}\, w_n(t,\mathbf{p}) + \frac{A_{-n}}{2}\, w_{-n}(t,\mathbf{p}) = R_n\, cos(\omega_n t - x\kappa_n \cos\theta_n - y\kappa_n \sin\theta_n + \phi_n),$$

is the cosine wave with amplitude R_n propagating along the direction θ_n.

The linear model is then corrected using *second order* or (*quadratic*) terms

$$W_q(t,\mathbf{p}) = \sum_{n=-N}^{N} \sum_{m=-N}^{N} \frac{A_n}{2}\frac{A_m}{2} E_{nm} w_n(t,\mathbf{p}) w_m(t,\mathbf{p}), \tag{25}$$

where the amplitude A, the angular frequency ω, the wave number κ and direction θ satisfy the same restrictions as in the linear model. The quadratic transfer function (QTF), $E_{nm} :=$

$E(\omega_n, \omega_m)$ is a real valued linear transformation that satisfies $E_{nm} = E_{mn}$, $E_{nm} = E_{-n-m}$ and $E_{-nn} = 0$, for all positive ω_n and ω_m.

Hence, the deterministic second order sea profile at time t and fixed point $\mathbf{p} = (x, y)$, can be writhen as

$$W(t, \mathbf{p}) = m + W_l(t, \mathbf{p}) + W_q(t, \mathbf{p}), \tag{26}$$

where m is the still water level, and the fields W_l and W_q are as in (24) and (25) respectively. In the following we assume that $m = 0$. Note also that since $m = 0$, $A_0 = 0$ and $E_{-nn} = 0$ hence $\frac{1}{T}\int_0^T W(t, \mathbf{p})\, dt \to 0$ as T increases to infinity, i.e. the average sea level is equal to still water level.

The exact form of the transfer function E_{nm} is given in Appendix I. A simplified form of the quadratic transfer function can be found for the special case of the uni-directional sea, i.e. the sea in which all waves travel along the same direction, $\theta_n = 0$ for $n > 0$, in deep water ($d = \infty$),

$$E_{nm} = \begin{cases} -\frac{1}{2g}|\omega_n^2 - \omega_m^2| & \text{if } n \cdot m < 0, \\ \frac{1}{2g}(\omega_n^2 + \omega_m^2) & \text{otherwise.} \end{cases}$$

Second order sea. Up to now we have considered a purely deterministic model for the sea surface. A random model is obtained by assuming that the complex valued amplitudes $A_n, n > 0$, are independent and normally distributed random variables, i.e.

$$A_n = \sigma_n(U_n - iV_n), \tag{27}$$

where U_n, V_n are independent standard normal variables, and σ_n^2 is the energy of waves with angular frequencies ω_n and ω_{-n}. From (27) it follows immediately that the linear part, defined in (24) is a Gaussian field. The resulting field $W(t, \mathbf{p})$ defined in (26) is non-Gaussian and usually referred to as second-order Gaussian sea.

It is often assumed that the Gaussian field W_l has a directional spectral density $S(\omega, \theta), \omega > 0$ that is traditionally factorized as

$$S(\omega, \theta) = S(\omega)G(\omega, \theta),$$

where $S(\omega)$ is a frequency spectrum, and the so called spreading function $G(\omega, \theta)$ satisfies $\int_{-\pi}^{\pi} G(\omega, \theta)\, d\theta = 1$. For more information on directional spectra, see [16]. A definition of the second order sea with continuous spectrum and some of its properties that are needed for the computation of the wave crest height distribution, are given in Appendix II.

In the rest of this paper we assume that the sea surface is observed at a fixed point $\mathbf{p}$ and denote the corresponding second order sea by $X(t) = W(t, \mathbf{p})$. It is easy to see that $\mathsf{E}[W_l(t, \mathbf{p})] = \mathsf{E}[W_q(t, \mathbf{p})] = 0$, and hence $\mathsf{E}[X(t)] = m$.

The following formulas can be found in [25]. The dispersion relation uniquely relates the wave number and the angular frequency, $\omega_n = g\kappa_n \tanh(d\kappa_n)$ where d is the water depth. Let us also denote $\omega_{nm} = \omega_n + \omega_m$ and $\boldsymbol{\kappa}_n = \kappa_n(\cos(\theta_n), \sin(\theta_n))$. We also make use of the following notation

$$\begin{aligned} C_{nm} &= ||\boldsymbol{\kappa}_n + \boldsymbol{\kappa}_m|| \tanh(d||\boldsymbol{\kappa}_n + \boldsymbol{\kappa}_m||), \\ k_{nm} &= \kappa_n \cos(\theta_n)\kappa_m \cos(\theta_m) + \kappa_n \sin(\theta_n)\kappa_m \sin(\theta_m), \\ D_{nm} &= \frac{2\omega_{nm}(g^2 k_{nm} - \omega_n^2\omega_m^2) + g^2(\kappa_n^2 2\omega_m + \kappa_m^2\omega_n) - \omega_n\omega_m(\omega_n^3 + \omega_m^3)}{2\omega_n\omega_m(\omega_{nm}^2 - gC_{nm})} \end{aligned}$$

then

$$E_{nm} = \frac{1}{2g}\left(\omega_{nm}^2 - \omega_n\omega_m - \frac{g^2 k_{nm}}{\omega_n\omega_n} + 2\omega_{nm}D_{nm}\right)$$

6. Appendix II

Suppose that the sea state is characterized by the directional spectrum $S(\omega,\theta) = S(\omega)G(\omega,\theta)$. In order to avoid non physical waves the frequency spectrum $S(\omega)$ is split into two parts; $S_0(\omega) = S(\omega)$, for $\omega < \omega^-$ or $\omega > \omega^+$, and $\tilde{S}(\omega) = S(\omega)$, for $\omega^- \le \omega \le \omega^+$. Denote by $\tilde{W}(t,\mathbf{p})$ the second order field defined by the truncated spectrum $\tilde{S}(\omega)G(\omega,\theta)$, as described in the following asymptotic procedure:

Let J, K be fixed integers. Consider $N = (J+1)(K+1)$ individual waves having angular frequencies $\omega_{jk} = \omega^- + j(\omega^+ - \omega^-)/J$, propagating along $\theta_{jk} = -\pi + 2k\pi/(K-1)$ carrying energy

$$\sigma_{jk}^2 = \tilde{S}(\omega_{jk})D(\omega_{jk},\theta_{jk})\frac{2\pi(\omega^+ - \omega^-)}{JK},$$

where $j = 0,\dots,J$, $k = 0,\dots,K$. For $n = 1 + j + k\cdot(J+1)$, let

$$\omega_n = \omega_{jk}, \quad \theta_n = \theta_{jk}, \quad \sigma_n^2 = \sigma_{jk}^2.$$

Let us now define a second order sea with N individual cosine waves by

$$W^N(t,\mathbf{p}) = W_l^N(t,\mathbf{p}) + W_q^N(t,\mathbf{p}),$$

where $W_l^N(t,\mathbf{p})$ and $W_q^N(t,\mathbf{p})$ are given in (24) and (25) respectively. The limiting process

$$\tilde{W}(t,\mathbf{p}) = \lim_{K,J\to\infty} W^N(t,\mathbf{p})$$

is obtained by letting K, J tend to infinity. Finally the second order sea is defined as the sum of two independent fields, the Gaussian field $W_0(t,\mathbf{p})$ with spectrum $S_0(\omega)G(\omega,\theta)$ and the field $\tilde{W}(t,\mathbf{p})$, i.e.

$$W(t,\mathbf{p}) = W_0(t,\mathbf{p}) + \tilde{W}(t,\mathbf{p}).$$

In the following we use $W^N(t,\mathbf{p})$ as an approximation of $\tilde{W}(t,\mathbf{p})$.

The field $W^N(t,\mathbf{p})$ can be written in a matrix form; define

$$[(U_1 - iV_1)w_1(t,\mathbf{p})\dots(U_N - iV_N)w_n(t,\mathbf{p})]^T = \mathbf{X}(t,\mathbf{p}) + i\mathbf{Y}(t,\mathbf{p}),$$

and

$$\begin{aligned} \mathbf{Q} &= [q_{mn}], \quad q_{mn} = (E_{m(-n)} + E_{mn})\sigma_m\sigma_n, \\ \mathbf{R} &= [r_{mn}], \quad r_{mn} = (E_{m(-n)} - E_{mn})\sigma_m\sigma_n, \\ \boldsymbol{\sigma} &= [\sigma_n], \quad \sigma_n = \sqrt{S(\omega_n)D(\omega_n,\theta_n)\frac{2\pi\omega_c}{N}}, \end{aligned} \tag{28}$$

with $m, n = 1, \dots, N$.

Using the relations in (28) we may write

$$\begin{aligned} W(t,\mathbf{p}) &= W_0(t,\mathbf{p}) + \boldsymbol{\sigma}^T \mathbf{X}(t,\mathbf{p}) + \frac{1}{2}\mathbf{X}(t,\mathbf{p})^T \mathbf{Q}\mathbf{X}(t,\mathbf{p}) \\ &+ \frac{1}{2}\mathbf{Y}(t,\mathbf{p})^T \mathbf{R}\mathbf{Y}(t,\mathbf{p}), \end{aligned}$$

with the field W_0 being independent of the fields $\mathbf{X}$ and $\mathbf{Y}$.

In this paper we are interested in measurements of the sea surface elevation at a fixed point $\mathbf{p} = (0, 0)$, hence we may write $X(t) = W(t, \mathbf{p})$. We may also denote by $X_0(t) = W_0(t, \mathbf{p})$ the Gaussian part of the process $X(t)$.

After some rather lengthy derivations including matrix diagonalization and some matrix algebra, for details see [19], we may write the process $X(t)$ in the equivalent form

$$X(t) = m + \sum_{j=0}^{2N} (\beta_j Z_j(t) + \gamma_j Z_j(t)^2), \tag{29}$$

where $\mathbf{Z}(t) = (Z_0(t), \dots, Z_{2N}(t))$ is a vector-valued stationary Gaussian process, such that for each t, $Z_j(t) \in N(0, 1)$ and the variables $Z_j(t), Z_k(t)$ are independent.

Let us also denote by $\dot{\mathbf{Z}}(t) = (\dot{Z}_0(t), \dots, \dot{Z}_{2N}(t))$, the derivative of vector $\mathbf{Z}(t)$. Then the joint density of the vectors $\mathbf{Z}(t)$ and $\dot{\mathbf{Z}}(t))$ is normal with $(\mathbf{Z}(t), \dot{\mathbf{Z}}(t)) \in N(\mathbf{0}, \Sigma)$, where

$$\Sigma = \begin{bmatrix} I & \Sigma_{12} \\ \Sigma_{12}^T & \Sigma_{22} \end{bmatrix} \tag{30}$$

while I is the identity matrix (note that the matrices Σ_{12}, Σ_{22} need not be identity matrices). Note that $Z_0(t) = X_0(t)/\beta_0$, where $\beta_0^2 = V(X_0(0))$, is independent of the processes $Z_j(t)$, $j = 1, \dots, N$. (Obviously $\gamma_0 = 0$ and $\sum \gamma_j = 0$.)

The linear part $X_l(t) = m + \sum \beta_j Z_j(t)$ is a Gaussian process with a spectrum $S(\omega)$ and variance

$$V(X_l(t)) = \sum_{j=0}^{2N} \beta_j^2 \approx \int_0^\infty S(\omega)\, d\omega = \lambda_0.$$

Equality for N going to infinity. The quadratic correction term $X_q(t) = \sum_{j=1}^{2N} \gamma_j Z_j(t)^2$ has mean zero and variance $V(X_q) = \sum_{j=1}^{2N} 2\gamma_j^2$. By the independence for fixed t of different $Z_j(t)$ and since $Z_j(t)$ and $Z_j(t)^2$ are uncorrelated, the variance of $X(t)$ is the sum of the variances of the terms in (29):

$$V(X(t)) = \sum_{j=0}^{2N} \beta_j^2 + \sum_{j=1}^{2N} 2\gamma_j^2.$$

Here, we should also mention that for efficiency in the computations the coefficients γ_j that are close to zero are omitted. In the examples where directional spectra are used, this means we have more than 100 non-zero components remaining from the more than 2000 we started with.

Finally, as we have mentioned before, the spectrum of the second order sea $X(t)$ is not equal to $S(\omega)$. The contribution of the quadratic component has to be removed. Consequently, one needs to solve the inverse problem to estimate the linear spectrum $S(\omega)$ and the spreading function $G(\omega, \theta)$, see [3].

7. Appendix III

The following generalization of Breitung's approximation can be found in [12].

Theorem 1. *Let $g : \mathbb{R}^n \to \mathbb{R}$ be a function such that the surface*

$$S = \{\mathbf{x} = (x_1, \dots, x_n); g(\mathbf{x}) = 0\}$$

has a point $\mathbf{x}_0$ such that $||\mathbf{x}_0|| = 1$ and $||\mathbf{x}|| > 1$ for all other $\mathbf{x} \in S$. By $\mathbf{x}$ we denote both the vector $(x_1, \dots, x_n)$ and the $n \times 1$ column matrix. Suppose $\mathbf{Z}(t)$ is an n-dimensional, stationary, differentiable, Gaussian vector process, and let $\dot{\mathbf{Z}}(t)$ denote its derivative. The correlation of the vector $(\mathbf{Z}(t),\dot{\mathbf{Z}}(t))$ is denoted by Σ,

$$\Sigma = \begin{bmatrix} I & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}. \tag{31}$$

For a family of processes $g(\mathbf{Z}(t)/\beta)$, $\beta > 0$, under some mild technical assumptions, the intensity of zero upcrossings is given by

$$\mu_\beta^+(0) = \frac{e^{-\beta^2/2}}{2\pi}(c + O(\beta^{-2})), \qquad c = \sqrt{\frac{\mathbf{x}_0^T\left(\Sigma_{22} - \Sigma_{21}G_0\Sigma_{12}\right)\mathbf{x}_0}{\det\left(I + P_0G_0P_0\right)}}, \tag{32}$$

as β tends to infinity, where $G_0 := \frac{1}{|\nabla g(\mathbf{x}_0)|}\left[\frac{\partial^2 g}{\partial x_i \partial x_j}(\mathbf{x}_0)\right]_{i,j=1,2,\dots,n}$ and $P_0 := I - \mathbf{x}_0\mathbf{x}_0^T$.

Since $\mathbf{Z}(t)$ is stationary, the same formula is valid for downcrossings.

Remark 2. *Formula (32) is in a sense finitely additive, i.e. if there is a finite number of points with minimal distance to the origin, the asymptotic formula for the upcrossing intensity of the process $g(\mathbf{Z}(t)/\beta)$ is the sum of the upcrossing intensities estimated using (32) for each point separately. Breitung's asymptotic approximation fails in the case of an infinite number of such points.*

Remark 3. *Theorem 1 is a generalizations of Breitung's result in two senses: In contrast to Theorem 1, Breitung demands the surface S be finite, and Theorem 1 contains the order of the error term.*

Remark 4. *Theorem 1 lends itself to a geometric interpretation. Note that $g(\mathbf{Z}(t)/\beta)$ crosses the zero level if and only if the vector process $\mathbf{Z}(t)$ crosses the surface βS. Hence, instead of saying that the formula is asymptotic "as β tends to infinity" we may say "as the surface S is inflated".*

We shall now demonstrate how the theorem can be used to approximate the upcrossing intensity $\mu^+(h)$ in the case of a second order sea. First we look at two special cases.

Small significant wave height h_s. Then the quadratic correction term in (29) may be ignored. That is, (29) simplifies to $X(t) = \mathbf{b}^T\mathbf{Z}(t)$. Define

$$g(\mathbf{x}) = 1 - \frac{1}{||\mathbf{b}||}\mathbf{b}^T\mathbf{x},$$

and note that, if $\beta = h/||\mathbf{b}|| := \beta_h$, the process $g(\mathbf{Z}(t)/\beta)$ downcrosses the zero level exactly when $X(t)$ upcrosses the level h. On the surface $g(\mathbf{x}) = 0$, the point closest to the origin is $\mathbf{x}_0 = \mathbf{b}/||\mathbf{b}||$, and since all second order derivatives of g are zero, Breitung's approximation gives

$$\mu_\beta^+(0) = \frac{1}{2\pi}\frac{\sqrt{\mathbf{b}^T\Sigma_{22}\mathbf{b}}}{||\mathbf{b}||}\exp\left(\frac{-\beta_h^2}{2}\right), \quad \beta_h = 4h/h_s, \tag{33}$$

which is clearly the exact form of Rice's formula for a Gaussian process with $h_s = 4||\mathbf{b}||$ and $t_z = 2\pi\frac{||\mathbf{b}||}{\sqrt{\mathbf{b}^T\Sigma_{22}\mathbf{b}}}$.

The significant wave height tending to infinity. The linear term in (29) is negligible. In this case we may write $X(t) = \mathbf{Z}(t)^T\Gamma\mathbf{Z}(t)$, where $\Gamma = \operatorname{diag}([\gamma_1, \ldots, \gamma_n])$ is the diagonal matrix with $\prod_{i=1}^n \gamma_i \neq 0$. Obviously we may also assume that $\gamma_1 \leq \gamma_2 \leq \cdots \leq \gamma_n$ and that $\gamma_n > 0$, since otherwise $X(t) \leq 0$ for all t. By defining

$$g(\mathbf{x}) = 1 - \frac{1}{\gamma_n}\mathbf{x}^T\Gamma\mathbf{x}, \tag{34}$$

it is readily seen that for $\beta = \sqrt{h/\gamma_n} := \beta_h$ the zero downcrossing intensity of the process $g(\mathbf{Z}(t)/\beta)$ equals the upcrossing intensity of the level h by the process $X(t)$. Consequently, $\mu_\beta^+(0) = \mu^+(h)$ and Theorem (1) may be used to compute $\mu_\beta^+(0)$. Note, however, that there are two points

$$\mathbf{x}_\pm := \pm\begin{bmatrix} 0 & 0 & \ldots & 1\end{bmatrix}^T$$

of minimal distance to the origin. Hence one has to compute formula (32) for each one of these points separately ($G_0 = -\frac{1}{\gamma_n}\Gamma$ for both points), and add the results.

In [12], explicit approximations of the error term are additionally provided. Consequently, the tails of the crest height distribution are exponentially distributed, i.e.

$$P(A_c > h) \approx \exp(-\frac{h}{2\gamma_n^2})(\frac{c}{\mu^+(0)} + O(h^{-1})),$$

where c is given in Theorem 1, while the zero upcrossing intensity $\mu^+(0)$ has to be estimated, for example by means of Monte Carlo methods.

7.1. **The general case of a second order sea.** Before we proceed any further, we should emphasize that Theorem 1 for the two cases of purely linear/quadratic sea model gave asymptotic formulas *as the level tends to infinity.* The general case is more complicated. This is not surprising since we have the mix of two different limiting cases. We have to construct somewhat artificial asymptotics. The idea is as follows. Fix the level h, assumed to be large, and let

$$p(\mathbf{x}) := \mathbf{b}^T\mathbf{x} + \mathbf{x}^T\Gamma\mathbf{x}.$$

Assume that there is only one point $\mathbf{x}_h$ on the surface

$$\{\mathbf{x} \in \mathbb{R}^n; p(\mathbf{x}) = h\}$$

of minimal distance to the origin, and define $\beta_h := ||\mathbf{x}_h||$. Let

$$g(\mathbf{x}) := 1 - \frac{1}{h}\left(\beta_h b^T\mathbf{x} + \beta_h^2\mathbf{x}^T\Gamma\mathbf{x}\right).$$

As before, the process $g(\mathbf{Z}(t)/\beta_h)$ crosses the level 0 when the process $p(\mathbf{Z}(t))$ crosses the level h; hence, using Breitung's method, with $\mathbf{x}_0 = \mathbf{x}_h/\beta_h$, for each level h separately, we have $\mu^+_{\beta_h}(0)$ equal to the u-upcrossing intensity for the process $p(\mathbf{Z}(t))$. Therefore, it is reasonable to believe that if β_h is large, then the term $O(\beta^{-2})$ for $\beta = \beta_h$ is small. Hence the approximation is good. However, the problem is that for each value h one defines a new function $g(\cdot)$, and hence we cannot use the theorem, which is valid for a fixed g, to motivate that the error term decreases to zero as h tends to infinity.

However, there are good reasons to use Breitung's approximation in this way instead of a formula which is truly asymptotic as the level h tends to infinity, especially if the linear terms dominate over the quadratic for the interesting levels of h. This is since, for h large enough, the h-crossings of the process will almost exclusively depend on the quadratic part. Our suggested use of Breitung's method will give a proper balance between the linear and the quadratic terms.

We turn now to the computation of Breitung,s approximation

$$\mu^+_{\beta_h}(0) = \exp(-\beta_h^2/2)c(\beta_h)/2\pi,$$

given in (1). Note that, opposed to the formulation in Theorem 1, we indicate c's dependence of β_h. Evaluating the terms gives

$$\begin{aligned}
\nabla g(\mathbf{x})^T|_{\mathbf{x}=\mathbf{x}_h/\beta_h} &= -\left(\frac{\beta_h}{h}\mathbf{b} + \frac{2\beta_h^2}{h}\Gamma\mathbf{x}\right)|_{\mathbf{x}=\mathbf{x}_h/\beta_h} = -\frac{\beta_h}{h}\left(\mathbf{b} + 2\Gamma\mathbf{x}_h\right), \\
G_0 &= \frac{-2\beta_h}{||\left(\mathbf{b} + 2\Gamma\mathbf{x}_h\right)||}\Gamma \quad P_0 = I - \frac{1}{\beta_h^2}\mathbf{x}_h\mathbf{x}_h^T.
\end{aligned}$$

Now we can use the following approximation

$$\mu(h) \approx \frac{\mathrm{e}^{-\beta_h^2/2}}{2\pi}c(\beta_h), \tag{35}$$

where

$$c(\beta_h) = \sqrt{\frac{\mathbf{x}_h^T\left(\Sigma_{22} - \Sigma_{21}G_0\Sigma_{21}\right)\mathbf{x}_h}{\det\left(I + P_0G_0P_0\right)}}$$

Remark 5. *The point of minimum norm, $\mathbf{x}_h$, can be found by standard optimization methods, see* [8].

Department of Mathematical Sciences, Gothenburg University, Chalmers University of Technology, SE-41296, Gothenburg, Sweden
E-mail address: anastass@chalmers.se

Dept. of Mathematical Statistics, Lund University, Lund, SE 22100, Sweden,
E-mail address: oskar@maths.lth.se

Department of Mathematical Sciences, Gothenburg University, Chalmers University of Technology, SE-41296, Gothenburg, Sweden
E-mail address: igor.rychlik@gmail.com