Maximum-entropy probability distribution of wind wave free-surface elevation

OCEANOLOGIA, 40 (3), 1998.
pp. 205–229.

KEYWORDS
Maximum-entropy principle
Maximum-entropy probability distribution
Random water waves
Wind waves
Free-surface elevations

Witold Cieślikiewicz
Institute of Hydro-Engineering,
Polish Academy of Sciences,
Kościerska 7, 80–953 Gdańsk, Poland;
e-mail: ciesl@ibwpan.gda.pl

Manuscript received June 16, 1998, in final form July 23, 1998.

Abstract

The probability density function of the surface elevation of a non-Gaussian random wave field is obtained. The derivation is based on the maximum entropy (information) principle with the first four statistical moments of the surface elevation used as constraints. The density function is found by the use of the Lagrangian multipliers method and it is shown that only two of four Lagrangian multipliers are independent. The applied method of numerical solution is described in detail and the useful nomograms that give the Lagrangian multipliers as functions of skewness and kurtosis are calculated and incorporated in the paper.

For slightly nonlinear waves the approximate maximum-entropy probability distribution is developed. The condition of the existence of this approximate distribution agrees with the empirical criterion for small deviations from the Gaussian distribution of random water waves.

The theoretical results compare well with field experiment data of Ochi and Wang (1984), even in the strongly non-Gaussian case.

1. Introduction

Wind waves are treated in general as a random process whose randomness follows from the nature of the generating forces and is a consequence of various instabilities in the wave evolution process. If we assume that the wind wave is a composition of denumerably many independent harmonic components, then by the central-limit theorem the free-surface elevation becomes Gaussian. In the case of weak nonlinear interactions a certain deviation from the Gaussian distribution may appear, as shown theoretically by Longuet-Higgins (1963). Among the experimental data, the probability distribution of which differs essentially from the Gaussian distribution, one can list the results of the 'Lubiatowo ′74' expedition (Bitner, 1980) and the data studied by Ochi and Wang (1984). The results of those experiments reveal the existence of strong deviations from the Gaussian distribution, particularly in the case of steep waves approaching the coastal zone. The laboratory investigations of Huang and Long (1980) show clearly the non-Gaussian character of a free-surface elevation.

The accurate determination of probability distribution is very important for the prediction of wave characteristics in engineering applications. With the increasing number of applications of remote sensing techniques based on optical as well as microwave systems (*e.g.* Synthetic Aperture Radar (SAR)), there is a need for the precise determination of probabilistic properties of a free-surface elevation.

There are three proposals for the non-Gaussian probability distributions of the surface elevation in the literature. The first one was given by Longuet-Higgins (1963) who used the Gram-Charlier series in Edgeworth's form. The second proposal is due to Tayfun (1980) and Huang *et al.* (1983), who devised a formula for the probability density based on the representation of free-surface elevation in the Stokes expansion of the second and third order.

Another non-Gaussian probability distribution of a free surface elevation based on *the maximum entropy (information) principle* was introduced by Cieślikiewicz (1988, 1990). This method is an alternative to the earlier approaches, and since no special representation of the random field of wind waves is used, one may expect that the method will find applications in a wide range of wave conditions.

In the present study the maximum-entropy approach is analysed in greater detail. The theoretical results are compared with the experimental data of Ochi and Wang (1984). These data, obtained during severe storms for a wide range of water depths, prove that the present approach is accurate even for steep waves. The calculated probability functions are highly consistent with the experimental values. The final part of this paper outlines the practical use of the non-Gaussian probability density determined by the maximum-entropy method in oceanography and engineering practice.

2. Maximum-entropy probability density for a random wind wave field

Jaynes (1957) introduced the maximum entropy principle as a criterion for selecting the probability distribution on the basis of partial knowledge (information) about the system. Essentially, the principle states that the probability distribution which best describes the available information, but is maximally non-committal with regard to the unavailable information, is the one that maximises the Shannon entropy subject to the given information as constraints. Jaynes (1968) also shows that 'the probability distribution which maximises the entropy is numerically identical to the frequency distribution which can be realised in the greatest number of ways'. The maximum entropy principle turns out to be fundamental for the application to real systems in physics.

In the case of random water waves we prescribe the constraints in the form of given values of the statistical moments of the free-surface elevation ξ. Such a choice emerges from the following two reasons: firstly, the Longuet-Higgins distribution is parametrised by the moments (cumulants, in fact), and we wish to secure the possibility of comparison; secondly, both in laboratory and full-scale measurements it is the statistical moments of a surface elevation that are often determined. Choosing them as constraints enables the theoretical results to be verified experimentally.

We assume that the mean value of the free-surface elevation ξ is zero. Further, let the variance μ_2 and the moments μ_3 and μ_4 of the third and fourth order of this random variable be given. Let $\Phi(\mu_2, \mu_3, \mu_4)$ denote the set of probability density functions $p(\xi)$ which reproduce these moments:

$$\Phi(\mu_2,\mu_3,\mu_4) = \Bigg\{ p(\xi): \quad p(\xi) \ge 0, \quad \int\limits_{-\infty}^{+\infty} p(\xi)\, d\xi = 1,$$

$$\int\limits_{-\infty}^{+\infty} \xi p(\xi)\, d\xi = 0, \quad \int\limits_{-\infty}^{+\infty} \xi^n p(\xi)\, d\xi = \mu_n \qquad \text{for} \quad n = 2,3,4 \Bigg\}. \tag{1}$$

We shall determine the so-called *representative distribution* for this set, *i.e.* the distribution which maximises the entropy on Φ. Owing to the maximum entropy principle, the representative distribution gives the fullest and most objective description of the 'statistical knowledge' contained in the moments μ_2, μ_3 and μ_4. The entropy S associated with the distribution p is defined as a functional of the form

$$S[p] = -\int\limits_{-\infty}^{+\infty} p(\xi)\, \ln p(\xi)\, d\xi. \tag{2}$$

We shall find the representative distribution p^* for the set Φ such that

$$S[p^*] = \max_{p\in\Phi} S[p] \tag{3}$$

by the use of the Lagrangian multipliers method.

Consider the functional

$$\tilde{S}[p] = -\int_{-\infty}^{+\infty} p(\xi)\left[\ln p(\xi) + \sum_{m=0}^{4} \alpha_m \xi^m\right] d\xi \tag{4}$$

in which α_m, $m = 0, 1, \ldots 4$, are the Lagrangian multipliers. From the condition for the functional $\tilde{S}$ to be extreme

$$\left.\frac{\delta\,\tilde{S}[p]}{\delta\,p(\xi)}\right|_{p=p^*} = 0\,, \tag{5}$$

where $\delta/\delta p(\xi)$ denotes a functional derivative, we obtain

$$p^*(\xi) = A^{-1}\exp\Big\{-\sum_{m=1}^{4} \alpha_m \xi^m\Big\} \tag{6}$$

in which $A^{-1} = \exp\{-1-\alpha_0\}$. To ensure the entropy (2) is finite we have to assume that

$$\alpha_4 > 0. \tag{7}$$

Using definition (1) of the set of probability density functions Φ, we can determine the unknown Lagrangian multipliers α_m and then the normalisation factor A. Thus, we have

$$\begin{cases} I_1 = 0 \\ I_n = \mu_n I_0 \qquad \text{for} \quad n = 2, 3, 4, \end{cases} \tag{8}$$

where

$$I_n = \int_{-\infty}^{+\infty} \xi^n \exp\Big\{-\sum_{m=1}^{4} \alpha_m \xi^m\Big\}\, d\xi. \tag{9}$$

The normalisation factor is given by

$$A = I_0. \tag{10}$$

System (8) consists of four nonlinear equations with respect to four unknown quantities $\alpha_1, \ldots, \alpha_4$. Using the obvious relations

$$\begin{aligned} &\int_{-\infty}^{+\infty} \frac{d}{d\xi}\,\Big\{\exp\Big[-\sum_{m=1}^{4} \alpha_m \xi^m\Big]\Big\}\, d\xi \equiv 0, \\ &\int_{-\infty}^{+\infty} \frac{d}{d\xi}\,\Big\{\xi\exp\Big[-\sum_{m=1}^{4} \alpha_m \xi^m\Big]\Big\}\, d\xi \equiv 0 \end{aligned} \tag{11}$$

one can readily verify the following identities (Cieślikiewicz, 1988, 1990)

$$\left\{\begin{array}{c} \alpha_1 I_0 + 2\alpha_2 I_1 + 3\alpha_3 I_2 + 4\alpha_4 I_3 \equiv 0 \\ I_0 - \alpha_1 I_1 - 2\alpha_2 I_2 - 3\alpha_3 I_3 - 4\alpha_4 I_4 \equiv 0 \end{array}\right. \tag{12}$$

which allow, by the use of (8), the expression, for example, of α_1 and α_2 in terms of α_3 and α_4

$$\left\{\begin{array}{l} \alpha_1 = -3\mu_2\alpha_3 - 4\mu_3\alpha_4 \\ \alpha_2 = \dfrac{1}{2\mu_2}(1 - 3\mu_3\alpha_3 - 4\mu_4\alpha_4). \end{array}\right. \tag{13}$$

Our task finally reduces to solving the system of four nonlinear equations (8) with unknown quantites α_3 and α_4. Note that the unknown Lagrangian multipliers α_3 and α_4 appear in (8) as parameters of integrals of the type (9). As we are unable to calculate these integrals analytically, numerical calculations are needed. It should be noted that the identities (12) play a key role in the numerical calculations as they provide a very efficient means of finding an initial guess good enough for conversion to a solution. The numerical algorithm will be presented in the next section. In section 4 an approximate analytical solution for small deviations of p^* from the Gaussian distribution will be presented.

Now let us consider two special cases of the problem. Firstly, we shall try to determine the representative distribution p^* for $\Phi(\mu_2, \mu_3)$ (when only the moments μ_2 and μ_3 are given). Then we have

$$\tilde{S}[p] = -\int\limits_{-\infty}^{+\infty} p(\xi)\left[\ln p(\xi) + \sum_{m=0}^{3} \gamma_m \xi^m\right] d\xi, \tag{14}$$

and

$$\frac{\delta \tilde{S}}{\delta p} = 0 \qquad \Longrightarrow \qquad p = B\exp(-\gamma_1\xi - \gamma_2\xi^2 - \gamma_3\xi^3). \tag{15}$$

If $\gamma_3 \neq 0$, then $S[p] = \infty$ and we thus see that a representative distribution of $\Phi(\mu_2, \mu_3)$ does not exist.

Finally, let us consider the problem of finding a representative distribution p^* of $\Phi(\mu_2)$ (when only μ_2 is given). In this case we obtain

$$p^*(\xi) = \frac{1}{\sqrt{2\pi\mu_2}} \exp\left(-\frac{\xi^2}{2\mu_2}\right) \tag{16}$$

which is the probability density function of the Gaussian distribution. It follows, therefore, that the representative distribution for $\Phi(\mu_2)$ is Gaussian.

3. Numerical evaluation of Lagrangian multipliers

For convenience, let us slightly modify our formulation in this section. It can be easily shown that the probability density function (6) may be expressed by the following formulae

$$\rho^*(\xi) = (\sigma B)^{-1} \exp\Big\{-\sum_{m=1}^{4} \beta_m (\xi/\sigma)^m\Big\}, \tag{17}$$

where

$$\beta_4 > 0, \tag{18}$$

and

$$\begin{cases} \beta_1 = -(3\beta_3 + 4\lambda_3\beta_4) \\ \beta_2 = \dfrac{1}{2}(1 - 3\lambda_3\beta_3 - 4\lambda_4\beta_4). \end{cases} \tag{19}$$

In these formulae the standard deviation σ, skewness λ_3 and kurtosis λ_4 are defined as

$$\sigma = \sqrt{\mu_2}, \qquad \lambda_3 = \frac{\mu_3}{\sigma^3}, \qquad \lambda_4 = \frac{\mu_4}{\sigma^4}. \tag{20}$$

The following system of equations determines the unknown multipliers β_3 and β_4

$$\begin{cases} J_1 = 0, & J_2 = J_0 \\ J_3 = \lambda_3 J_0, & J_4 = \lambda_4 J_0 \end{cases} \tag{21}$$

in which

$$J_n = \int_{-\infty}^{+\infty} \xi^n \exp\Big\{-\sum_{m=1}^{4} \beta_m \xi^m\Big\}\, d\xi. \tag{22}$$

The constant B in (17) is given by

$$B = J_0. \tag{23}$$

The representation of the probability density function in the form of eq. (17) is advantageous since the evaluation procedure for B, $\beta_1, \ldots, \beta_4$, determined by eqs. (19) to (21), does not depend on the variance σ^2. The parameters A, $\alpha_1, \ldots, \alpha_4$ can be retrieved by using the expressions:

$$\begin{cases} A = \sigma B \\ \alpha_m = \beta_m/\sigma^m & \text{for} \quad m = 1, \ldots, 4. \end{cases} \tag{24}$$

In order to determine the parameters of the probability function (17), we have to solve the system of equations (21) with respect to the unknowns β_3 and β_4. The quantities β_1 and β_2 are given by formulae (19). Because of the symmetry of the system (21), however, we solve a four-dimensional problem forgetting, as it were, the identities (19). This proves very effective on condition that the coordinates of the starting point satisfy identities (19). A general algorithm for numerical calculation by the use of the four-dimensional Newton method is described below.

Let

$$\beta = [\beta_1, \beta_2, \beta_3, \beta_4]^{\mathrm{T}}. \tag{25}$$

By means of (21) we can express our system in the form

$$F(\beta) = 0 \tag{26}$$

in which

$$F = [J_1,\ J_2 - J_0,\ J_3 - \lambda_3 J_0,\ J_4 - \lambda_4 J_0]^{\mathrm{T}}. \tag{27}$$

Then the iterative algorithm of the Newton method takes the form

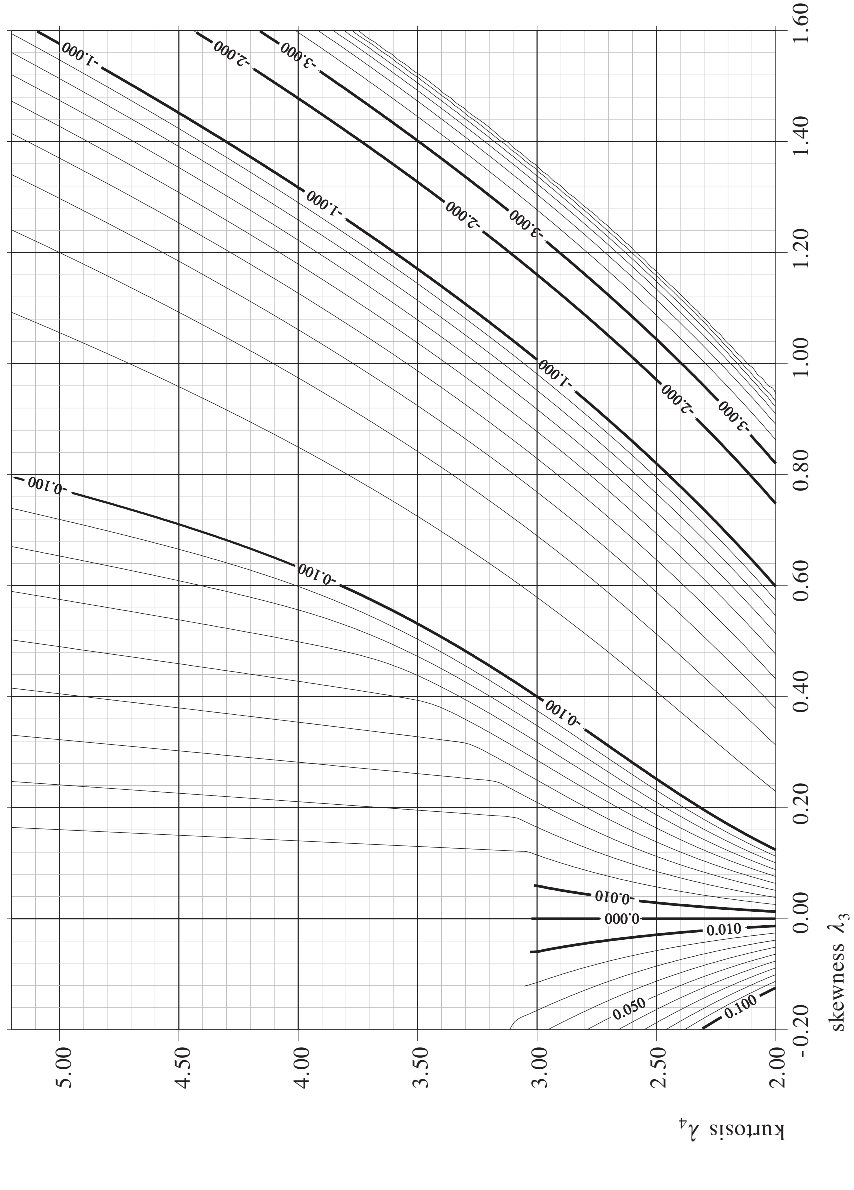
$$\mathcal{J}(\beta^{(k)})\,(\beta^{(k+1)} - \beta^{(k)}) = -F(\beta^{(k)}), \tag{28}$$

where the *Jacobian* matrix $\mathcal{J}(\beta)$ reads

$$\mathcal{J} = \begin{pmatrix} -J_2 & -J_3 & -J_4 & -J_5 \\ -J_3 + J_1 & -J_4 + J_2 & -J_5 + J_3 & -J_6 + J_4 \\ -J_4 + \lambda_3 J_1 & -J_5 + \lambda_3 J_2 & -J_6 + \lambda_3 J_3 & -J_7 + \lambda_3 J_4 \\ -J_5 + \lambda_4 J_1 & -J_6 + \lambda_4 J_2 & -J_7 + \lambda_4 J_3 & -J_8 + \lambda_4 J_4 \end{pmatrix}. \tag{29}$$

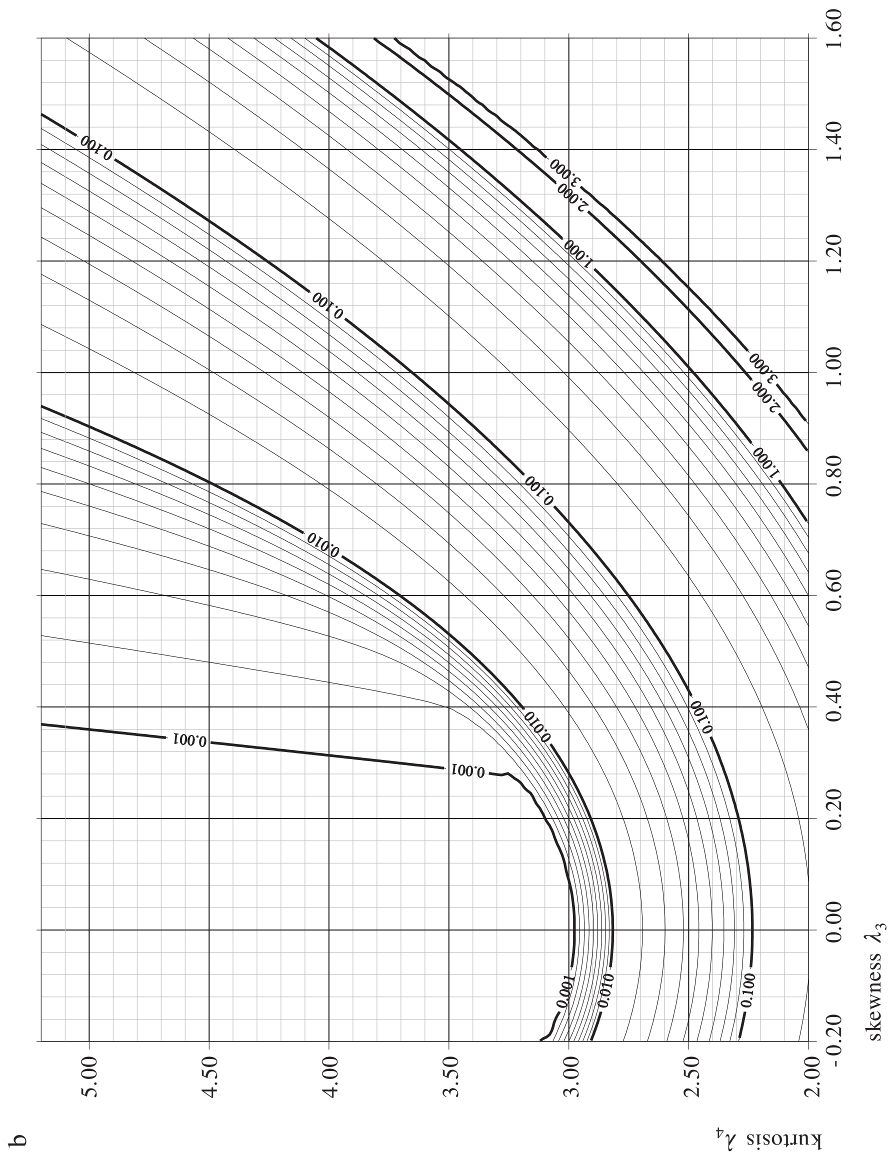
The numerical calculations are carried out on a grid covering a certain (λ_3, λ_4) region. This region is selected such that all points corresponding to the values of skewness λ_3 and kurtosis λ_4 from the experimental data of Ochi and Wang (1984) lie within. The system of equations (21) is solved at the discrete set of grid points with the grid spacing $\Delta\lambda_3 = 0.02$ and $\Delta\lambda_4 = 0.04$. The two-dimensional mesh covering the region of interest on plane (λ_3, λ_4) is fine enough to prepare nomograms for the sought – after Lagrangian multipliers. These nomograms for β_3 and β_4 are presented in Figs. 1a and 1b respectively. In Fig. 1c the nomogram for the constant B is shown. The two unfilled zones in each of the Figs. 1a, 1b and 1c in the upper-left and the lower-right corners correspond to regions (λ_3, λ_4) in which the convergence of the Newton method proved very poor.

The nomograms for the Lagrangian multipliers β_3, β_4 and constant B, which are the parameters of probability density function (17), are very useful in practical applications of the maximum-entropy probability distribution. One can avoid the complicated numerical calculations and, given the values of the skewness λ_3 and kurtosis λ_4, the parameters β_3, β_4 and B can be read off from Figs. 1a, 1b and 1c. Subsequently, the parameters β_1 and β_2



skewness λ_3
kurtosis λ_4
-0.20
0.00
0.20
0.40
0.60
0.80
1.00
1.20
1.40
1.60
5.00
4.50
4.00
3.50
3.00
2.50
2.00
-3.000
-2.000
-1.000
-0.100
-0.010
0.000
0.010
0.050
0.100

a


b
kurtosis λ_4
skewness λ_3
5.00
4.50
4.00
3.50
3.00
2.50
2.00
-0.20
0.00
0.20
0.40
0.60
0.80
1.00
1.20
1.40
1.60
0.001
0.010
0.100
1.000
2.000
3.000

c

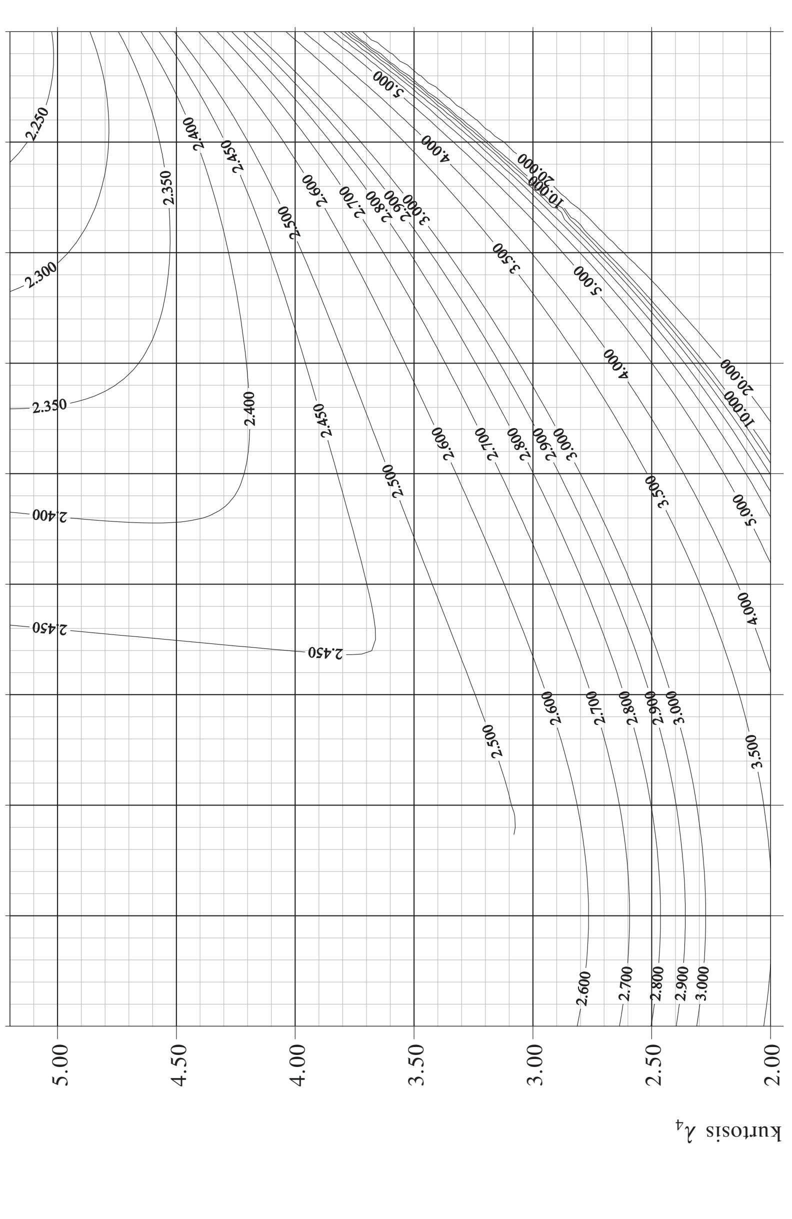


Fig. 1. Contour plots of parameters of maximum entropy distribution (17) as functions of skewness λ_3 and kurtosis λ_4: $\beta_3 = \beta_3(\lambda_3, \lambda_4)$ (a), $\beta_4 = \beta_4(\lambda_3, \lambda_4)$ (b), $B = B(\lambda_3, \lambda_4)$ (c)

can be obtained from eqs. (19). The maximum-entropy probability density function (17) for a given standard deviation σ, skewness λ_3 and kurtosis λ_4 can therefore be determined easily.

4. An approximate solution for small deviations from the Gaussian distribution

In this section, the formulation using αs and A will be used rather than that with βs and B. A characteristic function of the representative distribution (6) has the form

$$\theta^*(t) = A^{-1} \int_{-\infty}^{+\infty} \exp\Big\{ it\xi - \sum_{m=1}^{4} \alpha_m \xi^m \Big\}\, d\xi. \tag{30}$$

The addition to the normal density can be extracted by expanding the function $\exp\{-\alpha_1\xi - \alpha_3\xi^3 - \alpha_4\xi^4\}$ into a series

$$\begin{aligned}\theta^*(t) = A^{-1} \int_{-\infty}^{+\infty} d\xi\, e^{it\xi - \alpha_2\xi^2} \sum_{n=0}^{\infty} \sum_{k=0}^{n} \sum_{l=0}^{k} \frac{(-1)^n}{n!} \binom{n}{k} \binom{k}{l} \times \\ \times\ \alpha_1^{n-k} \alpha_3^{k-l} \alpha_4^{l} \xi^{n+2k+l}.\end{aligned} \tag{31}$$

The following equalities are readily verifiable:

$$\begin{aligned}\int_{-\infty}^{+\infty} \xi^n e^{it\xi - \alpha_2\xi^2}\, d\xi &= (-i)^n \frac{d^n}{dt^n} \int_{-\infty}^{+\infty} e^{it\xi - \alpha_2\xi^2}\, d\xi = \\ &= (-i)^n \sqrt{\frac{\pi}{\alpha_2}} \frac{d^n}{dt^n} \exp\left(-\frac{t^2}{4\alpha_2} \right) = \\ &= (i)^n \sqrt{\frac{\pi}{\alpha_2}} (2\alpha_2)^{-n/2} \exp\left(-\frac{t^2}{4\alpha_2} \right) H_n(t/\sqrt{2\alpha_2})\end{aligned} \tag{32}$$

in which $H_n(\cdot)$ denotes the n-th Hermite polynomial. Using (32) we can describe the characteristic function θ^* in the form of an infinite sum

$$\begin{aligned}\theta^*(t) = A^{-1} \sqrt{\frac{\pi}{\alpha_2}} \exp\left(-\frac{t^2}{4\alpha_2} \right) \sum_{n=0}^{\infty} \sum_{k=0}^{n} \sum_{l=0}^{k} \frac{(-1)^n}{n!} \binom{n}{k} \binom{k}{l} \times \\ \times\ \alpha_1^{n-k} \alpha_3^{k-l} \alpha_4^{l}\, i^{\,n+2k+l} (2\alpha_2)^{-(n+2k+l)/2} H_{n+2k+l}(t/\sqrt{2\alpha_2}).\end{aligned} \tag{33}$$

Calculating the successive derivatives of the function θ^* for $t = 0$ we obtain the moments of the probability distribution ξ

$$i^{\,n} \mu_n = \frac{d^n \theta^*(t)}{d\theta^n} \bigg|_{t=0}. \tag{34}$$

If the distribution (6) differs slightly from the Gaussian, then the exponent of the function $\exp\{-\alpha_1\xi - \alpha_3\xi^3 - \alpha_4\xi^4\}$ takes only small values for ξ such that $\exp(-\alpha_2\xi^2)$ is significantly greater than zero (say, for ξ in the region $\sqrt{\alpha_2}|\xi| < 3/\sqrt{2}$, which corresponds to $|\xi| < 3\sigma$ for the Gaussian distribution). For sufficiently small deviations, the first two terms of the outer sum in (33) will give reliable results. A large error for large values of ξ is non-essential, since the value of $\exp\{-\alpha_2\xi^2\}$ is then close to zero.

Taking the first two terms of the outer sum in (33) and using the following formulas for Hermite polynomials

$$H_{2n}(0) = (-1)^n(2n-1)!!, \qquad H_{2n+1}(0) = 0 \tag{35}$$

formula (34) yields the following system of equations with unknowns A, $\alpha_1, ..., \alpha_4$

$$\begin{cases} 1 - \dfrac{3\alpha_4}{(2\alpha_2)^2} = A\sqrt{\dfrac{\alpha_2}{\pi}} \\ \dfrac{\alpha_1}{2\alpha_2} + \dfrac{3\alpha_3}{(2\alpha_2)^2} = 0 \\ \dfrac{1}{2\alpha_2} - \dfrac{15\alpha_4}{(2\alpha_2)^3} = A\sqrt{\dfrac{\alpha_2}{\pi}}\,\mu_2 \\ -\dfrac{3\alpha_1}{(2\alpha_2)^2} - \dfrac{15\alpha_3}{(2\alpha_2)^3} = A\sqrt{\dfrac{\alpha_2}{\pi}}\,\mu_3 \\ \dfrac{3}{(2\alpha_2)^2} - \dfrac{105\alpha_4}{(2\alpha_2)^4} = A\sqrt{\dfrac{\alpha_2}{\pi}}\,\mu_4\,. \end{cases} \tag{36}$$

Solving this system we get

$$\begin{cases} A = \dfrac{4}{b}\sqrt{\dfrac{2\pi\mu_2}{a}} \\ \alpha_1 = \dfrac{2a^2\mu_3}{b\mu_2^2}, \qquad \alpha_2 = \dfrac{a}{2\mu_2} \\ \alpha_3 = -\dfrac{2a^3\mu_3}{3b\mu_2^3}, \qquad \alpha_4 = \dfrac{a^2(1-a)}{3b\mu_2^2}, \end{cases} \tag{37}$$

where

$$a = \mu_2(4\mu_2 - \sqrt{16\mu_2^2 - 5\mu_4})/\mu_4, \qquad b = 5 - a. \tag{38}$$

In order to guarantee the existence of this solution and to satisfy the condition (7) we have to assume that

$$\mu_4 < 3\mu_2^2 \quad \text{if} \quad \mu_3 \neq 0 \qquad \text{or} \qquad \mu_4 \leq 3\mu_2^2 \quad \text{if} \quad \mu_3 = 0. \tag{39}$$

Note that for $\mu_3 \to 0$ and $\mu_4 \to 3\mu_2^2$ the formulae (37) take the following form

$$A = \sqrt{2\pi\mu_2}, \qquad \alpha_2 = \frac{1}{2\mu_2}, \qquad \alpha_1 = \alpha_3 = \alpha_4 = 0 \tag{40}$$

leading to the Gaussian representative distribution. A numerical example for $\sigma = 0.5$, $\lambda_3 = 0.2$ and $\lambda_4 = 2.85$ is presented in Fig. 2. The density function (6) with parameters given by formulae (37) is drawn in this figure. For comparison, the Gaussian density function and a precise numerical solution obtained by the algorithm of the preceding section are also marked.

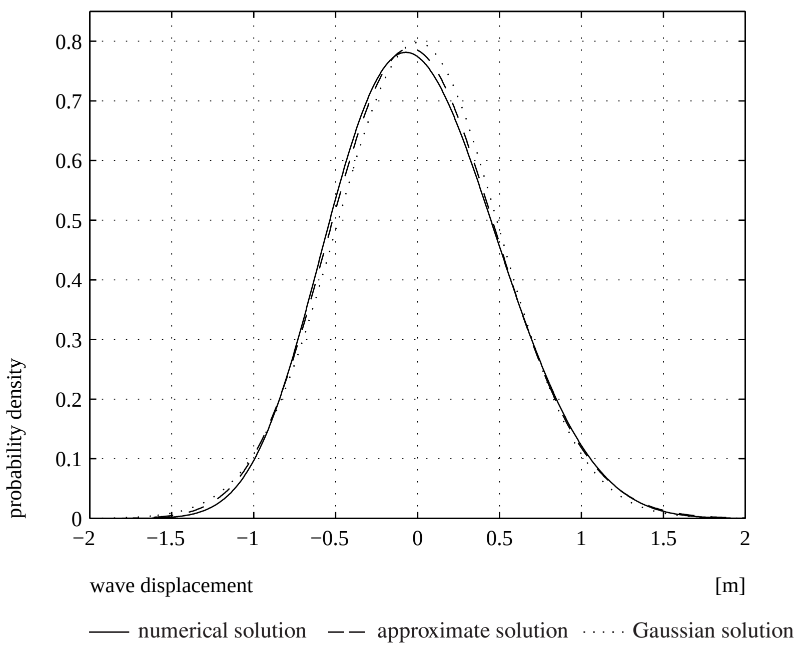


Fig. 2. Maximum-entropy probability distribution of surface elevation for $\sigma = 0.5$, $\lambda_3 = 0.2$ and $\lambda_4 = 2.85$

5. Comparison of maximum-entropy and observed probability distributions of free-surface elevation

In this section the maximum-entropy probability density function (17), with its parameters determined from experimental values of standard deviation σ, skewness λ_3 and kurtosis λ_4 of the free-surface elevation, is compared with observed frequency histograms. The parameters of theoretical distributions, *i.e.* βs and B in (17), are computed using the algorithm described in section 3. In that sense they are accurate, although their approximate values can be taken from the nomograms presented in Figs. 1a, 1b and 1c.

Below we refer to the results of the Atlantic Ocean Remote Sensing Land-Ocean Experiment (ARSLOE) undertaken in 1980 at the US Army Coastal Engineering Research Centre (CERC). The data obtained during the ARSLOE experiment are used to demonstrate that free-surface elevation follows the maximum-entropy probability distribution.

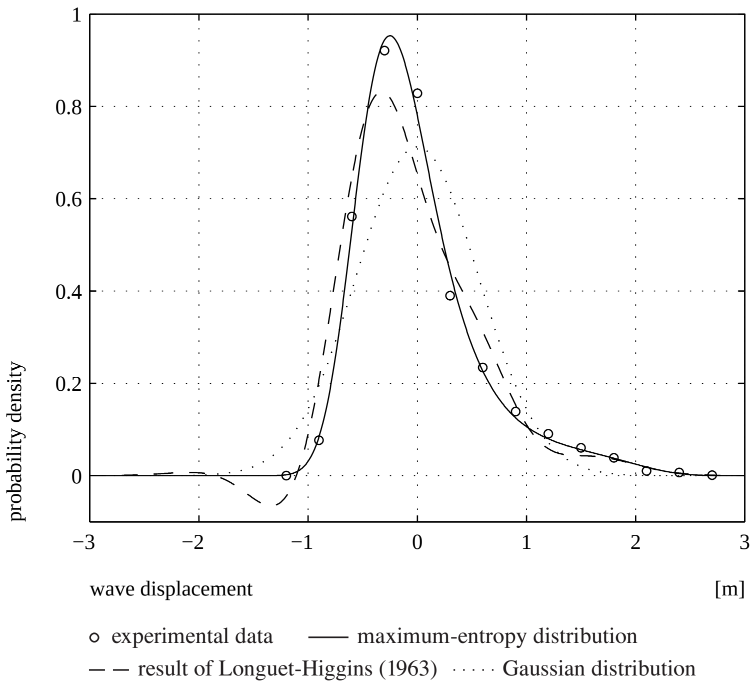


Fig. 3. Comparison of theoretical probability density functions and field data (Ochi and Wang, 1984) for $\sigma = 0.56$, $\lambda_3 = 1.26$, $\lambda_4 = 4.98$

In Fig. 3 the density function (17) is compared with the frequency histogram observed during a very severe storm of significant wave height 2.31 m in very shallow water (the average water depth during the storm was about 2 m). The observed histogram of free-surface elevations is taken from the paper of Ochi and Wang (1984). In that paper the skewness and kurtosis for the analysed record are also given as $\lambda_3 = 1.26$ and $\lambda_4 = 4.98$ [1]

[1] In the present paper we use the symbol λ_4 to denote kurtosis, while Ochi and Wang (1984) used the same symbol for the normalised cumulant of the 4th order, which is equal

respectively. These values, together with the standard deviation σ, enabled the parameters βs and B of the density (17) to be calculated. They are listed in Tab. 1. Both the observed and theoretical distributions and that of Longuet-Higgins (1963)

$$f(\xi) = \frac{1}{\sqrt{2\pi}\,\sigma}\,\mathrm{e}^{-\frac{1}{2}\zeta^2}\Big\{1 + \frac{1}{6}\lambda_3 H_3(\zeta) + \\ + \Big[\frac{1}{24}(\lambda_4 - 3)H_4(\zeta) + \frac{1}{72}\lambda_3^2 H_6(\zeta)\Big] + \ldots\Big\} \tag{41}$$

in which $\zeta = \xi/\sigma$, are also shown in this figure for comparison. The maximum entropy probability distribution shows a high degree of agreement with the observed distribution.

Table 1. Wind wave parameters of the ARSLOE experiment and the respective values of the normalisation factor and Lagrangian multipliers

Figure	h [m]	H_s [m]	σ [m]	λ_3	λ_4	B	β_1	β_2	β_3	β_4
3	2.0	2.31	0.56	1.26	4.98	2.2882	0.8125	0.6438	−0.3620	0.0542
4a	3.7	2.45	0.62	1.03	4.20	2.4023	0.7114	0.5485	−0.3067	0.0506
4b	8.8	3.55	0.89	0.85	3.87	2.4472	0.5470	0.5142	−0.2216	0.0347

There are no other examples in Ochi and Wang (1984) of observed frequency histograms of the free-surface elevation, with the skewness and kurtosis explicitly given. However, there are empirical formulae enabling λ_3 and λ_4 to be calculated in terms of the significant wave height H_s and the average water depth h given in their paper. The values of H_s and h are given in that paper for some other examples of the observed probability distribution, and two of them are used below for further verification of the maximum-entropy distribution of free-surface elevation. Let us first refer to the main conclusions of the investigations by Ochi and Wang (1984):

(i) The skewness of the free-surface elevation λ_3 depends only on the significant wave height H_s and the average water depth during the storm h.

(ii) The kurtosis of the free-surface elevation λ_4 is a function of λ_3 for $\lambda_3 > 0.2$.

(iii) For $\lambda_3 < 0.2$ there is some dispersion of values of λ_4 below three, but the probability distributions of the free-surface elevations for λ_3 and λ_4 in this region appear to be only slightly deviated from the Gaussian distribution.

---

to kurtosis – 3. The values of that normalised cumulant taken from Ochi and Wang (1984) henceforth will be recalculated to kurtosis.

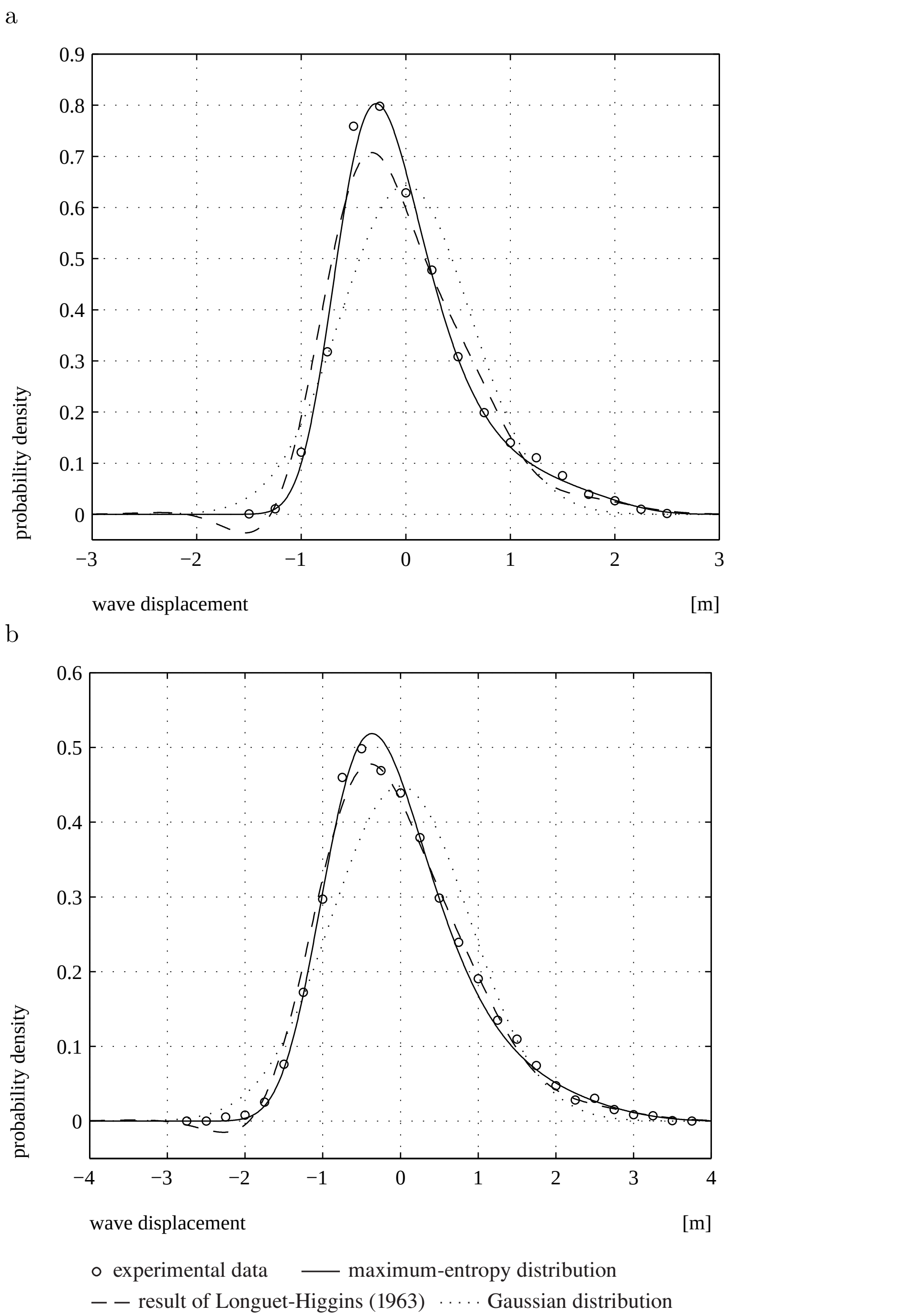


Fig. 4. Comparison of theoretical probability density functions and field data for $\sigma = 0.62$, $\lambda_3 = 1.03$, $\lambda_4 = 4.2$ (a) and $\sigma = 0.89$, $\lambda_3 = 0.85$, $\lambda_4 = 3.87$ (b)

Ochi and Wang (1984) proposed the empirical formulae $\lambda_3 = \lambda_3(H_s, h)$ and $\lambda_4 = \lambda_4(\lambda_3)$ corresponding to conclusions (i) and (ii) respectively. In Figs. 4a and 4b the maximum-entropy probability density function (17)

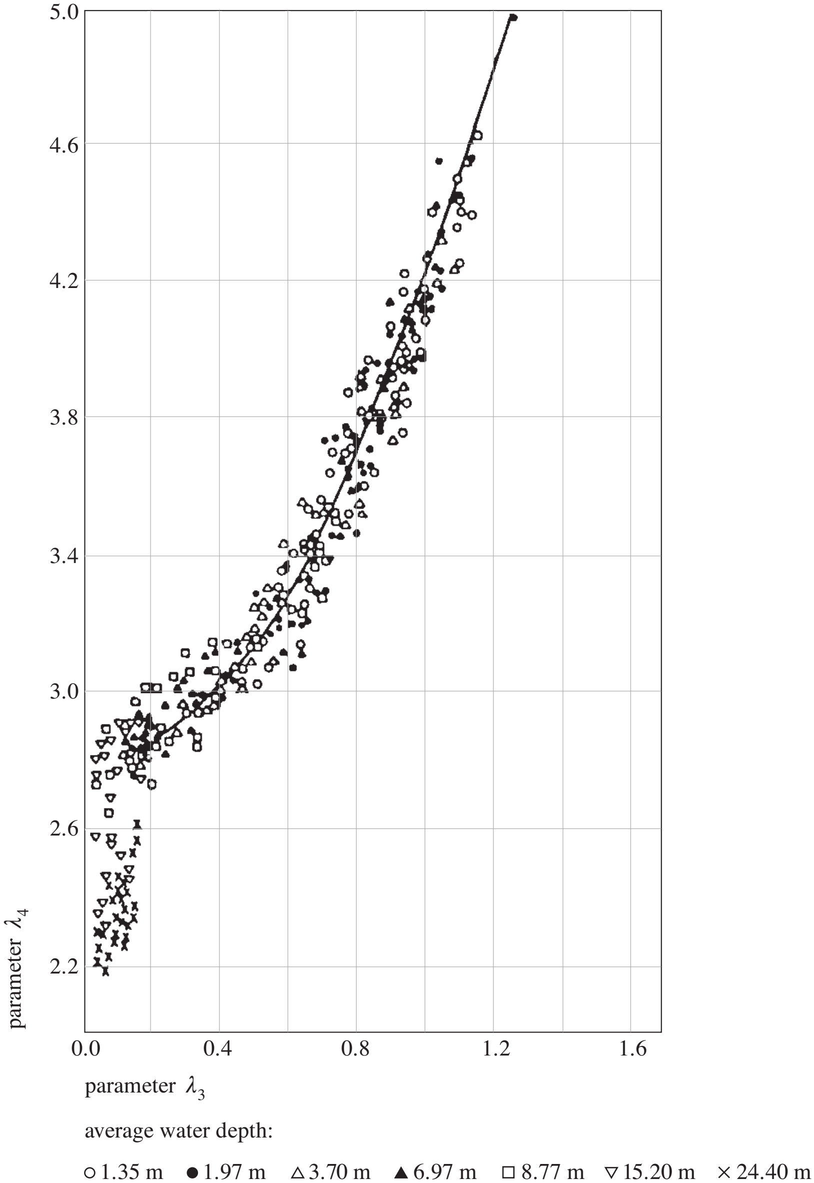
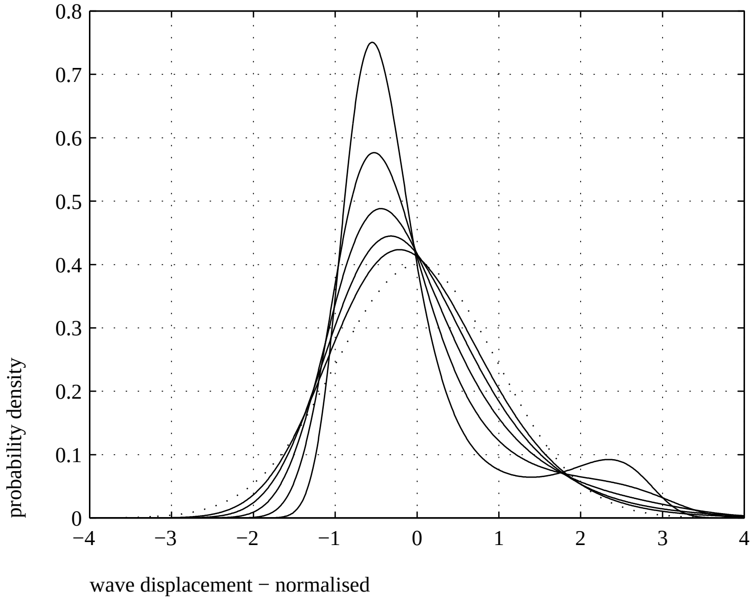


Fig. 5. Parameter λ_4 as a function of parameter λ_3 (Ochi and Wang (1984))

a



b

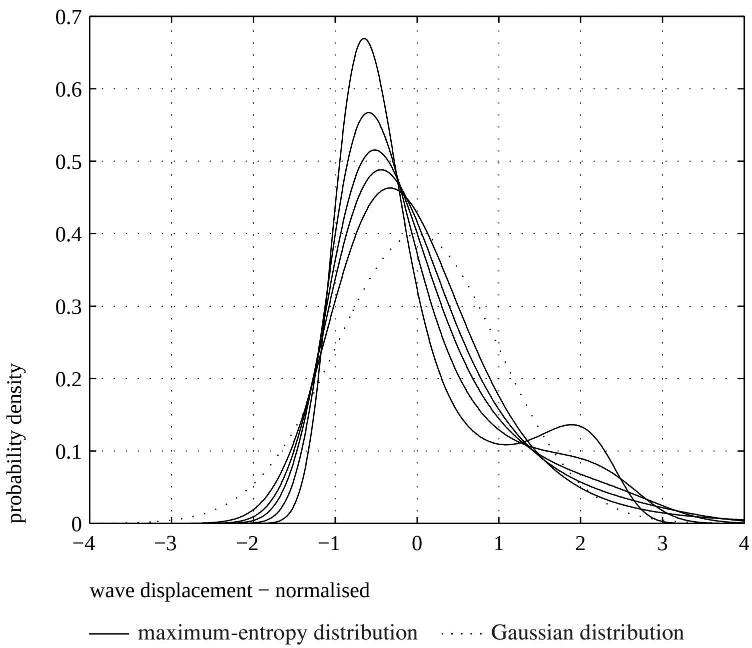


Fig. 6. Probability density functions for $\sigma = 1$, $\lambda_3 = 0.6$, 0.8, 1, 1.2, 1.4 and $\lambda_4 = 4.2$ (a) and $\sigma = 1$, $\lambda_3 = 1$ and $\lambda_4 = 3$, 3.4, 3.8, 4.2, 5 (b)

and the parameters βs and B calculated by using the skewness and kurtosis estimated from those empirical formulae are compared with the observed distributions of the free-surface elevations. This comparison is not direct in the sense that λ_3 and λ_4 are not exact and are influenced by the goodness of Ochi and Wang's empirical formulae. However, in this case the agreement is also very good. As before, the normal distribution and the distribution (41) are also marked in Figs. 4a and 4b. The numerical data for the examples presented are listed in Tab. 1.

Referring to conclusion (iii) above, it should be noted that this empirical criterion for small deviations from the Gaussian distribution $\lambda_3 < 0.2$ agrees well with the theoretical condition (39), which can be rewritten as $\lambda_4 < 3$. The consistency of these two conditions for the skewness and kurtosis respectively can be clearly seen from the diagram presented in Fig. 5 (Fig. 7 in Ochi and Wang (1984)), in which the observed values of the kurtosis λ_4 are plotted against the skewness λ_3.

Figs. 6a and 6b show the numerical examples demonstrating to some extent the dependence of the shape of density function (17) on the skewness and kurtosis respectively. The standard deviation is chosen as $\sigma = 1$ for all the examples. The parameters βs and B for the densities plotted in Fig. 6a

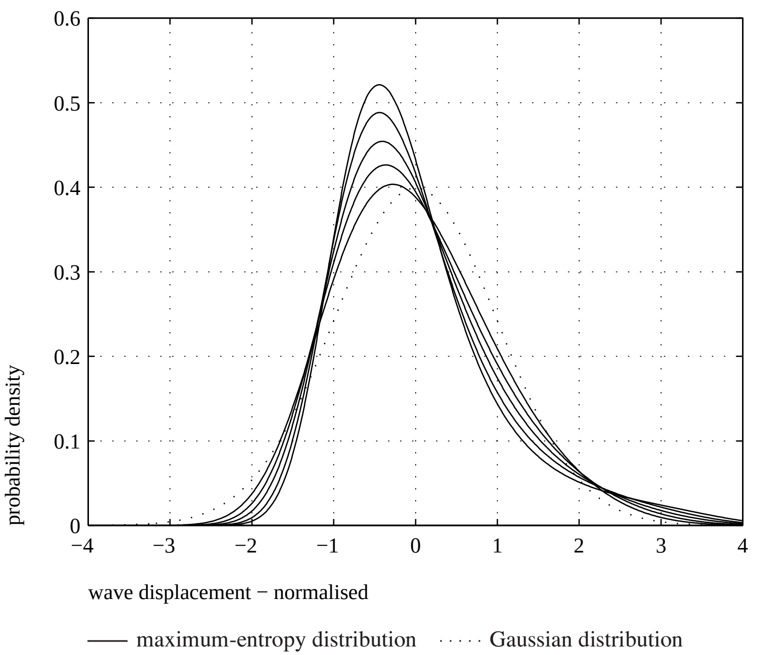


Fig. 7. Probability density functions for $\sigma = 1$, $(\lambda_3, \lambda_4) = (0.4, 3)$, $(0.6, 3.3)$, $(0.8, 3.74)$, $(1, 4.2)$, $(1.2, 4.82)$

are calculated for $\lambda_3 = 0.6, 0.8, 1, 1.2, 1.4$ and the kurtosis $\lambda_4 = 4.2$, while for the family of curves presented in Fig. 6b, the following data were chosen: $\lambda_3 = 1$ and $\lambda_4 = 3,\ 3.4,\ 3.8,\ 4.2,\ 5$. The dependence of the normalisation constant B on the skewness λ_3 is weak, as all the curves in Fig. 6a cross the vertical axis $\xi = 0$ at approximately the same point (it is obvious from (17) that $B = [\sigma\rho^*(0)]^{-1}$). Note that this result is reflected in the formulae (37).

Let us now come back to the above conclusion (ii), whose principal message is that in the case of the free-surface of wind waves the choice of pairs (λ_3, λ_4) is not arbitrary. In other words, only some combinations of the constraints (μ_3, μ_4) of the set of probability density functions $\Phi(\mu_2, \mu_3, \mu_4)$ defined in (1) are admissible in nature. Fig. 7 shows the maximum-entropy probability density functions of the free-surface elevation (17) for pairs (λ_3, λ_4) lying on the empirical curve $\lambda_4 = \lambda_4(\lambda_3)$ of Ochi and Wang (1984).

6. Discussion and conclusions

Although the maximum-entropy probability density function in its general form $p(\xi) = \exp\left(-\sum_{m=0}^{M} \alpha_m x^m\right)$, where α_m, $m = 0, 1, \ldots, M$ are Lagrange multipliers, has been known for a long time, it does not seem to be widely used in practice, despite its attractive physical interpretation it does not seem to be widely used in practice. Of course, the maximum-entropy distribution on $(-\infty, +\infty)$ in the case when only the first two moments are given as constraints is the normal distribution that is so important in physics. However, the simplest interesting case of the maximum-entropy probability distribution on $(-\infty, +\infty)$ occurs when the first four moments are given. This case leads to the difficult problem of solving a system of four nonlinear equations. There are no general methods for solving systems of more than one nonlinear equation. For problems in more than two dimensions the root finding becomes virtually impossible without additional insight. To the author's knowledge, the maximum-entropy probability distribution in the case of known first four statistical moments was used for the first time in Cieślikiewicz (1988, 1990). The parameters of that distribution, *i.e.* the Lagrangian multipliers, were found thanks to the simple identities (12) which transformed the four-dimensional problem into a two-dimensional one. In the present paper the method of solution is described in greater detail, and is extended by incorporating the useful nomograms that yield parameters of the maximum-entropy probability distribution (the Lagrangian multipliers) when the skewness and kurtosis coefficients are given.

It is not known *a priori* if the system of nonlinear equations (8) (or (21)) has a solution $\alpha_1, \ldots, \alpha_4$ (or $\beta_1, \ldots, \beta_4$). The general problem of the existence of a solution to the above-mentioned system is very difficult for systematic investigation. In this paper such a solution is found by numerical calculations over the defined region in the (λ_3, λ_4) plane. We can hope that the numerical solutions presented in this paper are approximations to the 'real' solutions. We then gain a deeper insight into the problem of the existence of a maximum-entropy probability distribution for the considered case when the first four moments are given. Note that if such a distribution exists, it is unique, since the integrand of the entropy functional S defined in (3) is convex and the constraints are linear (Dowson and Wragg, 1973).

It is shown in this paper that the maximum-entropy probability distribution exists for pairs of skewness and kurtosis (λ_3, λ_4) that are 'typical' of free-surface elevations. This does not, however, mean that the maximum-entropy distribution agrees with the 'real' probability distribution followed by the free-surface elevation. There is always a certain minimal level of given information below which the maximum-entropy probability distribution becomes inadequate. For example, in the case of linear random waves in deep water, it is enough to know the second moment of the free-surface elevation (we assume the mean value to be equal to zero). In this case the maximum-entropy distribution, which is in fact the normal distribution, describes the stochastic characteristic of the wave displacement very well. However, it is clearly not true for the highly nonlinear waves occurring in shallow water during a severe storm. Hence, the important question arises whether the amount of additional information contained in two higher statistical moments is sufficient to make the maximum-entropy distribution satisfactorily close to the underlying true probability distribution. The answer is one of the main aims of the present study and appears to be in the affirmative. The hypothesis that a wind wave random field of free-surface elevations follows the maximum-entropy probability distribution consistent with the constraints given by the first four moments has been confirmed by comparison with the field data.

In the case of small deviations from the Gaussian distribution, *i.e.* for slightly nonlinear waves, the approximate maximum-entropy probability distribution is developed. The condition of the existence of that approximation agrees with the empirical criterion for small deviations from the Gaussian distribution.

Now, let us discuss the relationship between the method of determining the probability distribution using the maximum-entropy principle with the moment constraints, and methods of fitting the probability distribution to the data based on standard methods of parameter estimation. The most important thing, of course, is that the maximum-entropy principle provides the formula for the probability density function, while the standard methods need the distribution function to be assumed and only its unknown parameters are estimated. However, it is quite obvious that the evaluation procedure for the parameters of the maximum-entropy density function (*i.e.* the Lagrangian multipliers), with the moment constraints in the background, is equivalent to the parameter estimation based on *the method of moments* applied to the density function (6) (or (17)). It is shown in the Appendix that this procedure is also equivalent to *the method of maximum likelihood* estimation.

The probability distribution selected by means of the maximum-entropy principle has a very clear and attractive interpretation. In fact, in making statistical inferences on the basis of partial information contained in the measured data, we must use that probability which has maximum entropy (2) subject to whatever is known – for instance the first statistical moments, as in our case. This is the only unbiased assignment we can make; using any other would introduce additional information which by assumption we do not have.

The argumentation given above is a poorly probabilistic one. However, it is very important to note that when the nonlinearities of the wave motion can be neglected, the maximum-entropy probability distribution of free-surface elevation becomes Gaussian. Thus, for linear random waves the maximum-entropy distribution, as well as being Gaussian, also possesses a very strong 'mechanical' background. This comes from the central-limit theorem applied to a sea surface represented by the sum of infinite number of simple harmonic waves.

The probability distribution (17) can be easily utilised in practical applications, when the mean value, variance, skewness and kurtosis are given, by using the nomograms $\beta_3 = \beta_3(\lambda_3, \lambda_4)$, $\beta_4 = \beta_4(\lambda_3, \lambda_4)$ and $B = B(\lambda_3, \lambda_4)$ in Figs. 1a, 1b and 1c. Of course, these nomograms are determined for the density function (17) regardless of the application, and hence may be used in cases other than free-surface elevations of water waves. For wind waves, however, if it is possible to forecast the first four statistical moments starting from some basic properties of the wind wave random field (*e.g.* wind velocity, fetch, water depth), the *a priori* probability distribution in the form of (17) can be predicted. Such a possibility arises, firstly, from the empirical findings of Ochi and Wang (1984), which show a strong dependence of kurtosis λ_4 on skewness λ_3 (*i.e.* the function $\lambda_4 = \lambda_4(\lambda_3)$ can be empirically determined). Secondly, it follows from the theoretical results of Longuet-Higgins (1963), who evaluated the skewness λ_3 explicitly in terms of the spectral density of free-surface elevation of waves in deep water. The more general formulae, valid in intermediate water depths, can be found in Cieślikiewicz and Gudmestad (1993).

References

Bitner E. M., 1980, *Nonlinear effects of the statistical model by shallow water wind waves*, Appl. Ocean Res., 2, 63–73.

Cieślikiewicz W., 1988, *Determination of the probability distribution of a sea surface elevation via the maximum-entropy method*, Rep. Inst. Hydro-Eng. Polish Acad. Sci., Gdańsk, 26 pp., (in Polish).

Cieślikiewicz W., 1990, *Determination of the surface elevation probability distribution of wind waves using the maximum-entropy principle*, [in:] *Water waves kinematics*, A. Tørum and O. T. Gudmestad (eds.), NATO ASI, Ser. E: Appl. Sci., 178, Kluwer Acad. Publ., Dordrecht–Boston–London, 345–348.

Cieślikiewicz W., Gudmestad O. T., 1993, *Stochastic characteristics of orbital velocities of random water waves*, J. Fluid Mech., 255, 275–299.

Dowson D. C., Wragg A., 1973, *Maximum-entropy distributions having prescribed first and second moments*, IEEE Trans. Inform. Theory, 19, 689–693.

Huang N. E., Long S. R., 1980, *An experimental study of the surface elevation probability distribution and statistics of wind-generated waves*, J. Fluid Mech., 101, 179–200.

Huang N. E., Long S. R., Tung C. C., Yuan Y., Bliven L. F., 1983, *A non-Gaussian statistical model for surface elevation of nonlinear random wave fields*, J. Geophys. Res., 88, 7597–7606.

Jaynes E. T., 1957, *Information theory and statistical mechanics*, Phys. Rev., 106, 620–630.

Jaynes E. T., 1968, *Prior probabilities*, IEEE Trans. Sys. Sci. Cybern., SEC–4, 227–241.

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

Ochi M. K., Wang W. C., 1984, *Non-Gaussian characteristics of coastal waves*, Proc. 19th Coast. Eng. Conf., Houston, Texas, ASCE, New York, I, 516–531.

Tayfun M. A., 1980, *Narrow-band nonlinear sea waves*, J. Geophys. Res., 85, 1548–1552.

Appendix

Let $p(\xi;\alpha)$ denote a distribution function of form (6) where α represents a family of unknown parameters $\alpha_1, \ldots, \alpha_4$. Then the likelihood function L for a random sample of size N is defined as

$$L(\xi_1, \xi_2, \ldots, \xi_N; \alpha) = \prod_{n=1}^{N} p(\xi_n; \alpha). \tag{A1}$$

According to *the method of maximum likelihood* we choose as an estimate of α the set of values which maximises L for the given values of $\xi_1, \ldots, \xi_N$. As it is more convenient to work with $\ln L$ than with L itself, the required values of α can be found by solving the following *likelihood equations*

$$\frac{\partial \ln L}{\partial \alpha_m} = \sum_{n=1}^{N} \frac{\partial \ln p(\xi_n; \alpha)}{\partial \alpha_m} = 0 \qquad \text{for} \quad m = 1, \ldots, 4 \tag{A2}$$

with respect to α_m, $m = 1, \ldots, 4$.

According to (10) $A = I_0(\alpha_1, \ldots, \alpha_4)$, and thus $p(\xi;\alpha)$ of form (6) may be written

$$p(\xi;\alpha) = \exp\Big\{-\ln I_0 - \sum_{m=1}^{4} \alpha_m \xi^m\Big\}. \tag{A3}$$

Therefore, we have for $m = 1, \ldots, 4$

$$\frac{\partial \ln p(\xi;\alpha)}{\partial \alpha_m} = -\frac{1}{I_0}\frac{\partial I_0}{\partial \alpha_m} - \xi^m. \tag{A4}$$

Using the above formula in (A2) gives the likelihood equations in the form

$$\sum_{n=1}^{N} \left(\frac{1}{I_0}\frac{\partial I_0}{\partial \alpha_m} + \xi_n^m\right) = 0 \qquad \text{for} \quad m = 1, \ldots, 4. \tag{A5}$$

From the definition of I_m (9) it follows that

$$\frac{\partial I_0}{\partial \alpha_m} = I_m \qquad \text{for} \quad m = 1, \ldots, 4 \tag{A6}$$

so (A5) can be written as

$$I_m = I_0 M_m \qquad \text{for} \quad m = 1, \ldots, 4, \tag{A7}$$

where

$$M_m = \frac{1}{N} \sum_{n=1}^{N} \xi_n^m \tag{A8}$$

is the mth sample moment of the random variable ξ.

Without any loss of generality we can select the origin of the random variable ξ such that

$$M_1 = 0. \tag{A9}$$

Then (A7) reads

$$\begin{cases} I_1 = 0 \\ I_m = M_m I_0 \qquad \text{for} \quad m = 2, 3, 4 \end{cases} \tag{A10}$$

which corresponds nicely to (8). The only difference is that the population moments μ_n in (8) are replaced in (A10) by the sample moments M_n.