Statistical modeling of the sea surface in the presence of abnormal waves

A. Zapevalov[1*] and *A*. Knyazkov[1]

[1]Marine Hydrophysical Institute, RAS, Kapitanskaya str., 2, 299011, Sevastopol, Russia

Abstract. A wide range of fundamental and applied problems requires a detailed description of the statistics of abnormal sea waves (freak waves or rouge waves). These waves are characterized not only by a change in energy, but also by strong nonlinearity, leading to extreme values of the higher cumulants. The possibilities and limitations of modeling the probability density function (PDF) of sea surface elevations by a two-component Gaussian mixture at extreme values of skewness and excess kurtosis are analyzed. The parameters of a two-component Gaussian mixture are calculated from known values of statistical moments. Model PDFs in the form of a two-component Gaussian mixture are compared with PDFs based on direct wave measurement data, and also compared with the known Gram-Charlier distribution. It is shown that with positive values of the excess kurtosis, the PDF in the form of a two-component Gaussian mixture can be constructed at the limit values of the skewness and excess kurtosis obtained in different regions of the World Ocean. With large negative values of the kurtosis, the shape of the probability density function is strongly distorted, which indicates the limit of applicability of a two-component Gaussian mixture to the description of such situations. Keywords: Two-Component Gaussian Mixture, Sea Surface, Elevation Distribution.

1 Introduction

One of the most important characteristics of the sea surface is the elevation probability density function (PDF), which is used in solving a wide range of applied problems related to the impact of waves on ships, offshore platforms, and coastal structures. It is also used in tasks related to remote sensing of the ocean. Three main types of models are used for statistical description of sea waves. The first type is dynamic models developed on the basis of hydrodynamic equations [1-3]. The second type is numerical spectral wave models [4, 5]. The third type is models developed within the framework of the theory of random processes [6-9].

Here we will limit ourselves to analyzing models of the third type. Unknown parameters of the third type of models are calculated based on known statistical moments. These models describe the sea surface well when deviations from the Gaussian distribution are sufficiently small [10]. With large deviations, distortions of the PDF of surface elevations occur [11].

* Corresponding author: sevzepter@mail.ru

The relevance of constructing PDF elevations of the sea surface at large values of excess kurtosis is primarily related to the analysis of the probability of the appearance of abnormally high waves (freak waves or rouge waves). According to wave measurements carried out in different areas of the World Ocean, large values of kurtosis indicate a high probability of occurrence of freak waves [12-14]. PDF studies with significant deviations of skewness from zero are of interest for remote sensing problems. These deviations lead to a displacement of the leading edge of the reflected radio pulse of the altimeter and an error in determining the level of the sea surface [15-17]. This error caused by the deviation of the distribution of sea surface elevations from the Gauss distribution is called skewness bias.

Wave measurements in the coastal zone at relatively shallow depths, where the approximation of “deep water” is not performed, have shown the existence of freak waves whose trough is larger than the crest [18]. According to wave measurements carried out in the coastal zone of the Baltic Sea, in situations where freak waves were recorded, waves were observed in 17.5%, in which the depression exceeded the crest by one and a half times [19]. In the coastal zone of the Baltic Sea, in situations where freak waves were recorded, waves were observed in which the trough exceeded the crest by one and a half times [19]. Measurements in the Western Black Sea region of Turkey (depth 12.5 m) showed for freak wave that skewness changing in the range −0.01 and −0.4 and the excess kurtosis in the range of 0.1–1.2 [20].

The purpose of this work is to assess the possibility and limitations of using a two-component Gaussian mixture at extreme values of skewness and excess kurtosis of sea surface elevations.

2 Simulation and numerical analysis

2.1 Two-component Gaussian mixture

Gaussian mixtures are great features that are widely used for modeling in various fields of science and practice [9, 21, 22]. PDF in the form of a two-component Gaussian mixture has the form [23]

$$P_S(\xi)=\frac{\alpha_1}{\sqrt{2\pi}\sigma_1}\exp\left(-\frac{(\xi-m_1)^2}{2\sigma_1^{\ 2}}\right)+\frac{\alpha_2}{\sqrt{2\pi}\sigma_2}\exp\left(-\frac{(\xi-m_2)^2}{2\sigma_2^{\ 2}}\right) \tag{1}$$

Where α_i is the weight of the i–th component ($i=1,2$), m_i is the mathematical expectation, $\sigma_i^{\ 2}$ is the variance. The weight coefficients satisfy two conditions

$$\alpha_i\in(0,1),\ \alpha_1+\alpha_2=1 \tag{2}$$

The procedure for calculating the parameters of a function $P_S(\xi)$ is based on a relation linking its statistical moments with the statistical moments of its components

$$\mu_j=\alpha_1\mu_{i,1}+\alpha_2\mu_{i,2} \tag{3}$$

$$\mu_{j,i}=\int\xi^j\frac{\alpha_i}{\sqrt{2\pi}\sigma_i}\exp\left(-\frac{(\xi-m_i)^2}{2\sigma_i^{\ 2}}\right)d\xi \tag{4}$$

Where μ_3 and μ_4 are third and fourth statistical points. The statistical moment of order n is defined as

$$\mu_n = \left\langle \xi^n \right\rangle \tag{5}$$

Where the symbol $\langle \ldots \rangle$ means averaging. Here and elsewhere we will assume that $\mu_2 = 1$, then $\mu_3 = A$ is an skewness, $\mu_4 - 3 = E$ is an excess kurtosis.

The procedure for calculating the parameters of a two-component Gaussian mixture is described in detail in [10].

2.2 Gram-Charlier distribution

The approximation of PDF by Gram-Charlier series is based on the well-known decomposition into a series by derivatives of the function [24]

$$N(\xi) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\xi^2\right) \tag{6}$$

Which are defined by the expression

$$\frac{d^n}{dx^n} N(\xi) = (-1)^n H_n(\xi) \cdot PN(\xi) \tag{7}$$

Where $H_n(\xi)$ are orthogonal Chebyshev-Hermite polynomials of order n. In general, a PDF of a random variable with a unit variance can be represented in the form

$$P_{GC}(\xi) = \sum_{i=0}^{\infty} C_n H_n(\xi) \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\xi^2\right) \tag{8}$$

Where C_n are the coefficients of the series.

Coefficients C_n are calculated based on statistical moments. For the elevation of the sea surface, as a rule, the statistical moments of the elevations of the sea surface are known only up to the fourth order inclusive, therefore, truncated series are used

$$P_{GC}(\xi) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\xi^2\right) \left\{1 + \frac{\mu_3}{6} H_3(\xi) + \frac{(\mu_4 - 3)}{24} H_4(\xi)\right\} \tag{9}$$

Polynomials $H_3(\xi)$ and $H_4(\xi)$ are given by expressions

$$H_3(\xi) = \xi^3 - 3\xi \tag{10}$$

$$H_4(\xi) = \xi^4 - 6\xi^2 + 3 \tag{11}$$

It should be emphasized that the use of truncated series leads to distortions at the tails of the distribution, which manifest themselves as negative PDF values or the appearance of local extremes [11, 24]. These distortions impose a limit on the range of acceptable values ξ in which the function can be constructed $P_{GC}(\xi)$.

2.3 Numerical Simulation

The functions calculated for several values of skewness and excess kurtosis are shown in Fig. 1. The values of statistical moments μ_3 equal to -0.2 and 0.3, as well as $\mu_4 - 3$ equal to -0.4 and 1 approximately correspond to their limit values observed in the coastal zone of the Black Sea [25, 26].

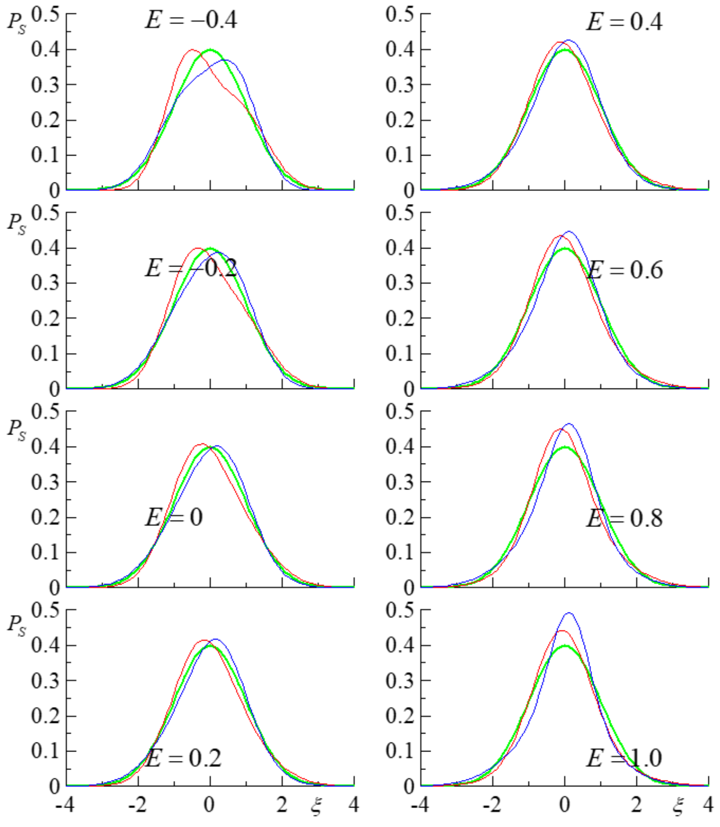

Fig. 1. PDF model in the form of a two-component Gaussian mixture. The red and blue curves are a two-component Gaussian mixture calculated at $A = 0.3$ and $A = -0.2$; green curve is Gaussian distribution.

The PDF of sea surface elevations must be non-negative and satisfy the requirement of single-modality. In a linear wave field, which is a superposition of sinusoidal waves, the distribution of surface elevations is described by the Gaussian distribution [27]. Sea waves are a weakly nonlinear process, so deviations from the Gaussian distribution are relatively small. This imposes another requirement on the PDF, the number of inflection points must be equal to two.

To control the single-modality and convexity-concavity of the PDF, the first and second derivatives of the function $P_S(\xi)$ were also analyzed, since for some values of skewness and excess kurtosis, it is not possible to construct a single-modal distribution of a two-component Gaussian mixture [10]. At $E = -0.4$, although there is no inflection of the function, but its second derivative has additional local maxima (see Fig. 2).

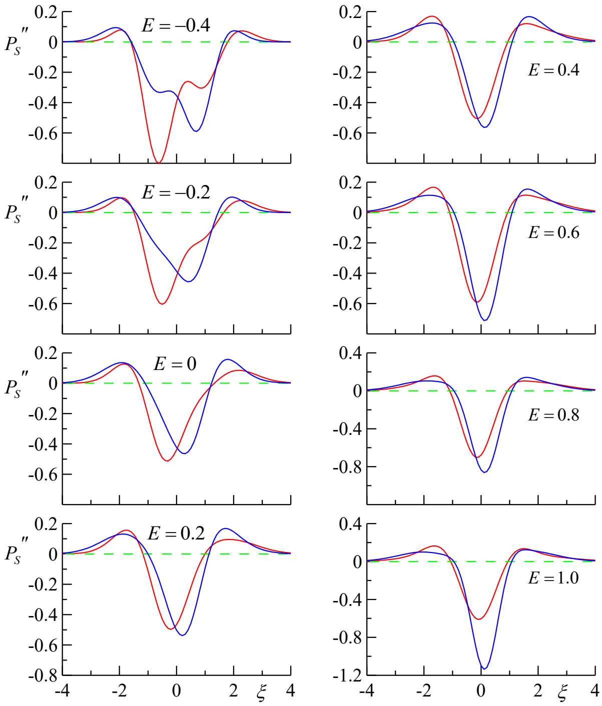

Fig. 2. The second derivative of PDF in the form of a two-component Gaussian mixture $P_S''(\xi)$. The red curve corresponds $A = 0.3$; the blue curve corresponds $A = -0.2$.

3 Discussion

3.1 Comparison with observational data

Let's compare the model and calculated from the measurement data of the probability density functions of sea surface elevations. To do this, we will use the data of direct wave measurements obtained on the stationary oceanographic platform of the Marine Hydrophysical Institute of the RAS [25]. The data of continuous measurements were divided into sessions lasting 20 minutes, for each of which the values of skewness and excess kurtosis were calculated, according to which the model was built $P_S(\xi)$, as well as the experimental PDF $P_E(\xi)$. To construct the function $P_E(\xi)$, the waveforms of each measurement session were normalized in such a way as to obtain a single variance. Next, a histogram of the distribution of elevations was constructed with a width $\Delta\xi = 0.25$. The function was obtained from the histogram by normalizing by the total number of points in the waveform and by the width $\Delta\xi$.

Situations in which the kurtosis exceeded the level of 0.4 are relatively rare. During December 2018, when the wave data analyzed in our work were obtained, 15 similar situations were recorded with a total of 2232 measurement sessions. The maximum value of excess kurtosis was 0.80. The functions $P_S(\xi)$ and $P_E(\xi)$ for the same values of skewness and excess kurtosis are shown in Figs. 3.

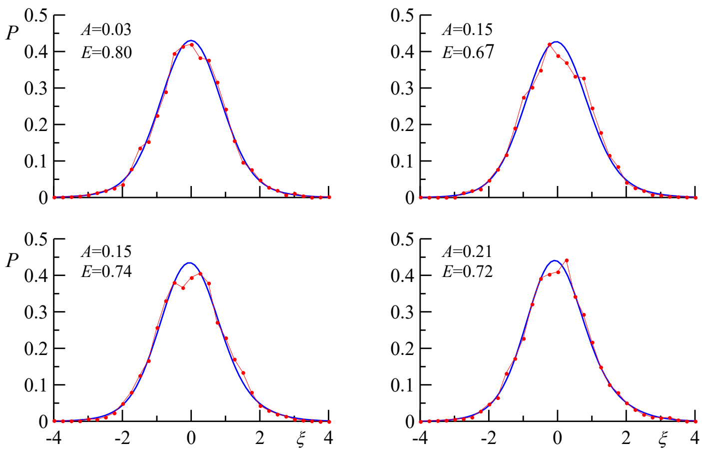

Fig. 3. Comparison of PDF in the form of a two-component Gaussian mixture (blue curve) and empirical PDF (red curve).

3.2 Comparison with the Gram-Charlier distribution

In the wave measurement data used here, the values of skewness and excess kurtosis are in the ranges $-0.2 < A < 0.3$, $-0.4 < E \leq 0.8$. In some situations, in particular, when the approximation of deep water is not performed, when abnormal waves are present, the ranges of variation of the higher statistical moments are much wider [20, 26]. To test the function $P_S(\xi)$ beyond the specified ranges, we will use a distribution based on a truncated Gram-Charlier series $P_{GC}(\xi)$ [28]. Taking into account the distortions on the tails of the distribution that occur when using truncated Gram-Charlier series [11, 24], the comparison will be carried out in the region $|\xi| < 2.5$. Previously, PDF in the form $P_{GC}(\xi)$ were verified on the data of direct wave measurements [26]. It has been shown that the relative error is small in the region $|\xi| < 2.5$.

It should be noted that for the problems of the impact of waves on ships and coastal structures, large crest and large trough are of the greatest interest, which make the main contribution to the distribution tails. In altimetric determination of sea level from spacecraft, it is important to know the deviation of the median distribution of elevations relative to the level of the nonperturbed surface (relative to $\xi = 0$) [16, 17]. Thus, the comparison $P_S(\xi)$ and $P_{GC}(\xi)$ is primarily of interest for the tasks of studying the sea by means of remote sensing.

The approximations of the PDF constructed within the framework of two models for extreme values of skewness and excess kurtosis are shown in Fig. 4.

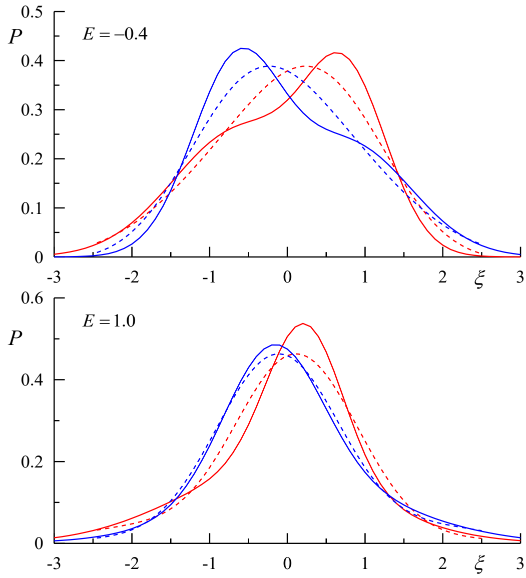

Fig. 4. Comparison of the functions $P_S(\xi)$ (solid curve) and $P_{GC}(\xi)$ (dashed lines). Red curves correspond $A=-0.4$; blue curves correspond $A=0.4$.

According to [20] in shallow water (depth 12.5 m) in the presence of abnormal waves, the skewness can approach -0.4. In experiments on a stationary oceanographic platform (depth about 30 m), the skewness varies from -0.23 to 0.37 [26].

To control PDF distortion, Fig. 5 shows the second derivatives of the PDF.

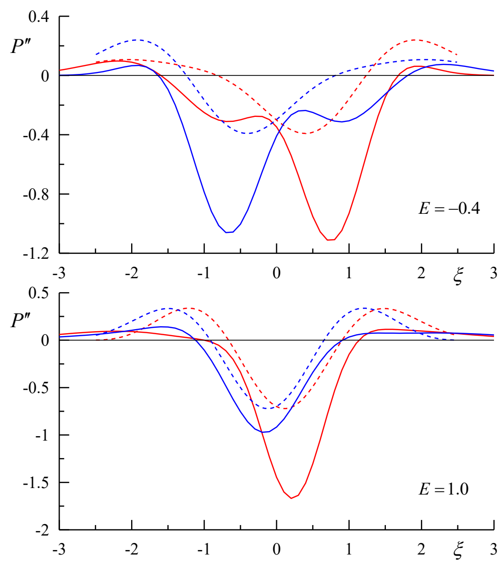

Fig. 5. Comparison of the second derivatives of the PDF in the form of a Gaussian mixture (solid curve) and the Gram-Charlier distribution (dashed lines). Red curves correspond $A=-0.4$; blue curves correspond $A=0.4$.

4 Conclusion

Currently, the main model describing the statistical characteristics of sea surface elevations remains the model constructed on the basis of truncated Gram-Charlier series. The known limitations associated with its application stimulated the search for new approaches. One of such approaches is the use of the PDF of sea surface elevations in the form of a two-component Gaussian mixture.

The two-component Gaussian mixture describes well the elevations of the sea surface with relatively small values of skewness and excess kurtosis. With positive values of excess kurtosis, the PDF in the form of a two-component Gaussian mixture can be constructed at the limit values of skewness and excess kurtosis obtained in different regions of the World Ocean. With large negative values of the excess kurtosis, the shape of the probability density function is strongly distorted, which probably indicates the limit of applicability of the two-component Gaussian mixture model to the description of such situations. Perhaps we are at the boundary of compatibility of this model and the unimodality condition, which is a natural physical condition in the description of surface waves.

Acknowledgment

The work was performed at the Marine Hydrophysical Institute of the Russian Academy of Sciences as part of the assignment on the topic: FNNN-2021-0004 “Fundamental studies of oceanological processes which determine the state and evolution of the marine environment influenced by natural and anthropogenic factors, based on observation and modeling methods”.

References

1. F. Fedele, F. Arena, APS **17(1)**, 026601 (2005)
2. Y. Hou, G. Song, X. Zhao, J. Song, Q. Zheng, Chinese Journal Oceanology and Limnology **24(1)**, 1-5 (2006) DOI: 10.1007/BF02842767
3. M.A. Tayfun, M.A. Alkhalidi, Offshore Technology Conference Asia (2016) DOI: 10.4043/26436-ms
4. S.Y. Annenkov, V.I. Shrira, J. Physical Oceanography **44(6)**, 1582–1594 (2014)
5. S.Yu. Mikhailichenko, A.V. Garmashov, V.V. Fomin, Ecological Safety of Coastal and Shelf Zones of Sea **2**, 52-57 (2016)
6. F.M. Bréon, N. Henriot, J. Geoph. Res.: Oceans **111**, C06005 (2006) DOI: 10.1029/2005JC003343.
7. A.S. Zapevalov, Yu.B. Ratner, Physical Oceanography **13(1)**, 1-13 (2003) DOI: 10.1023/A:1022444703787
8. V.I. Tatarskii, J. Atmospheric and Oceanic Technology **20**, 1697-1705 (2003) DOI: 10.1175/1520-0426(2003)020<1697:MROTCD>2.0.CO;2
9. Z. Gao, Z. Sun, S. Liang, Ocean Engineering **213**, 107815 (2020) DOI: https://doi.org/10.1016/j.oceaneng.2020.107815
10. A.S. Zapevalov, A.S. Knyazkov, Physical Oceanography **29(4)**, 395-403 (2022) DOI: 10.22449/1573-160X-2022-4-395-403.
11. O.K. Kwon, Communications in Statistics - Theory and Methods (2020) DOI: 10.1080/03610926.2020.1833219

12. C. Guedes Soares, Z. Cherneva, E.M. Antão, Ocean Engineering **31**, 1147–1167 (2004) DOI: 10.1016/J.OCEANENG.2003.10.014
13. V.A. Ivanov, V.A. Dulov, S.Yu. Kuznetsov, S.F. Dotsenko, et all., Geography, Environment, Sustainability **05(1)**, 84-111 (2012)
14. P. Stansell, Applied Ocean Research **26(1–2)**, 35-48 (2004)
15. A.S. Zapevalov, Izvestiya, Atmospheric and Oceanic Physics **48(2)**, 200–206 (2012) DOI: 10.1134/S0001433812020120.
16. S.I. Badulin, V.G. Grigorieva, P.A. Shabanov, et all., Advances in Space Research **68(2)**, 978-988 (2021) DOI: 10.1016/j.asr.2019.11.040
17. N. Pires, M. Fernandes, C. Gommenginger, R. Scharroo, Remote Sensing **8(7)**, 576 (2016) DOI: 10.3390/rs8070576.
18. J. Glejin, V. Sanil Kumar, et all., Indian Journal of Geo-Marine Sciences **43(7)**, 1339-1342 (2014)
19. I. Didenkulova, C. Anderson, Natural Hazards and Earth System Science **10(9)**, 2021-2029 (2010)
20. E. Bilyay, B.O. Ozbahceci, A.C. Yalciner, Black Sea. Nat. Hazards Earth Syst. Sci. **11(3)**, 659–666 (2011) DOI: 10.5194/nhess-11-659-2011.
21. Yu.I. Zhuravlev, I.V. Isaev, Zhurnal Vychislitel'noi Matematiki i Matematicheskoi Fiziki **19(3)**, 726–738 (1979)
22. M.A. Carreira-Perpiñán, IEEE Trans. on Pattern Analys. and Mach. Intell **22(11)**, 1318–1323 (2000)
23. N.N. Aprausheva, S.V. Sorokin, Pattern Recognition and Image Analysis **23(3)**, 341-347 (2013) DOI: 10.1134/S1054661813030024.
24. W. Lin, , J.E. Zhang, J. of Computational and Applied Mathematics **407(2)**, 113945 (2021) DOI: 10.1016/j.cam.2021.113945
25. A.S. Zapevalov, A.V. Garmashov, Physical Oceanography **28(4)**, 414-425 (2021) DOI: 10.22449/1573-160X-2021-4-414-425.
26. A.S. Zapevalov, A.N. Bol'shakov, V.E. Smolov, Oceanology **51(3)**, 406–413 (2011) DOI: 10.1134/S0001437011030222.
27. M.S. Longuet-Higgins, Mathematical, Physical and Engineering Sciences **249(966)**, 321–387 (1957) DOI:10.1098/rsta.1957.0002
28. M.J. Kendall, A. Stewart, *The advanced theory of statistics*. Vol. I. Distribution theory (Butler & Tanner Ltd, London, 1958)