Statistical Moment and Cumulant of the Fifth Order of the Sea Surface Elevation

Alexander Zapevalov[a],
Marine Hydrophysical Institute, RAS. 2, Kapitanskaya St., Sevastopol, Russian Federation
sevzepter@mail.ru

Keywords: Sea surface, Skewness, Asymmetry, Fifth statistical moment, Fifth cumulant.

Abstract: For the field of sea waves, the relationship between skewness λ_3, the fifth statistical moment μ_5 and the fifth cumulant λ_5 of surface elevations is analyzed. The analysis is carried out on the basis of direct wave measurements obtained on a stationary oceanographic platform of the Marine Hydrophysical Institute. The oceanographic platform is installed on the Black Sea. For the deep water condition, the limits of variation of the parameters μ_5 and λ_5 are determined. It is shown that the correlation coefficient between and is equal to 0.92, which makes it possible to describe the dependence $\mu_5 = \mu_5(\lambda_3)$ by linear regression. The obtained regression allows us to close the system of equations for calculating the parameters of the probability density function of sea surface elevations in the form of a two-component Gaussian mixture. The statistical relationship between λ_3 and λ_5, as well as between μ_5 and λ_5 is weak, the correlation coefficients are respectively equal to -0.18 and 0.21.

1 INTRODUCTION

Statistical description of the sea surface is one of the most urgent tasks of modern oceanology.

Existing mathematical models based on the Stokes decomposition or on the basis of spectral analysis do not yet allow calculating statistical moments older than the fourth order [Hou et al., 2006; Annenkov and Shrira, 2013]. The skewness and excess kurtosis obtained in these models are always positive, which does not correspond to the results of measurements in the field, where negative values of the third and fourth cumulants are observed [Guedes Soares, 2004; Zapevalov and Garmashov, 2021].

Sea waves are a weakly nonlinear process, so the distribution of sea surface elevations differs from the Gauss distribution [Babanin and Polnikov, 1994; Zapevalov and Garmashov, 2021]. To describe the probability density function of sea surface elevations, truncated Gram-Charlier and Edgeworth series are usually used [Burdyugov, 2022], as well as approximations in the form of a two-component Gaussian mixture [Gao et al., 2020].

For the correct use of both approximations, information is needed about the fifth statistical moment, which, as a rule, is not determined in experiments. The parameters of distributions constructed on the basis of Gram-Charlier and Edgeworth series are calculated according to the specified statistical moments [Kendall and Stewart, 1958]. These series represent the desired function in the form of a decomposition over orthogonal Chebyshev-Hermite polynomials. The low order of series truncation caused by the lack of information about statistical moments older than the fourth order leads to distortions at the tails of the distribution [Blinnikov and Moessner, 1998; Kwon, 2020]. When calculating the parameters of a two-component Gaussian mixture, information about the first five statistical moments is required [Zapevalov and Ratner, 2003]. If this information is missing, there is an ambiguity in their definition [Zapevalov and Knyazkov, 2022].

One of the possible ways to solve this problem is to determine the relationship between statistical moments (cumulants) of different orders. In this paper, based on the data of direct wave measurements, the relationship of the third statistical

[a] https://orcid.org/ 0000-0001-9942-2796

moment with the fifth statistical moment and the fifth cumulant is analyzed.

2 MATERIALS AND METHODS

2.1 Statistical Moment

Statistical moments of a random variable ξ can be defined as

$$\mu_n = \langle \xi^n \rangle , \quad (1)$$

where *n* is the order of the statistical moment, the symbol denotes averaging. For the Gaussian distribution central statistical moments are

$$\left.\begin{aligned} \mu_{2n} &= \frac{(2n)!}{2^n n!}\mu_2^n , \\ \mu_{2n+1} &= 0, \quad n \ge 1, \end{aligned}\right\} . \quad (2)$$

Here, instead of statistical moments, it is more convenient to use their polynomial combinations, called cumulants. Cumulants are denoted as λ_n. Moments and cumulants contain the same information and can be uniquely expressed through each other. If the mean value of the random variable ξ is zero, then

$$\left.\begin{aligned} \lambda_1 &= 0 \\ \lambda_2 &= \mu_2 \\ \lambda_3 &= \mu_3 \\ \lambda_4 &= \mu_4 - 3\mu_2^2 \\ \lambda_5 &= \mu_5 - 10\mu_3\mu_2 \end{aligned}\right\} \quad (3)$$

Here and further we will assume that $\mu_2 = 1$. In this case, the cumulants λ_3 and λ_4 have their own names, λ_3 is a skewness, λ_4 is an excess kurtosis. Since the third cumulant and the third statistical moment coincide, then in the analysis and in the equations we will use we will use both λ_3 and μ_3.

2.2 Wave Measurements

To analyze the statistical characteristics of sea surface waves, data obtained on the stationary oceanographic platform of the Marine Hydrophysical Institute were used. The depth at the location of the stationary oceanographic platform is about 30 m. The features of the meteorological regime in the measurement area are described in the works [Shokurova et al., 2016; Solov'ev and Ivanov, 2007; Mikhailichenko et al., 2016]. Measurements were carried out in the summer and autumn of 2005 and 2006, as well as in the winter of 2018.

For measurements, a string wave recorder was used, the sensor of which is an uninsulated wire. To eliminate the interference created by the platform supports, the wave recorder moved away from the nearest support at a distance of more than 6 m. For statistical analysis, wave records were divided into fragments lasting 20 min, for which statistical characteristics were calculated.

3 RESULTS AND DISCUSSION

3.1 Parameterization of the probability density function

If the type of probability density function is unknown, then it can be constructed from known statistical moments. Let's start with a situation where the approximation of the probability density function is carried out using a two-component Gaussian mixture. It has the form [Aprausheva and Sorokin, 2013]

$$P_S(\xi) = \frac{\alpha_1}{\sqrt{2\pi}\sigma_1}\exp\left(-\frac{(\xi - m_1)^2}{2\sigma_1^2}\right) + \\ + \frac{\alpha_2}{\sqrt{2\pi}\sigma_2}\exp\left(-\frac{(\xi - m_2)^2}{2\sigma_2^2}\right), \quad (4)$$

where α_i is the weighting coefficient of the *i*–th component, m_i is the mathematical expectation, σ_i^2 is the variance. The weight coefficients satisfy the condition

$$\alpha_1 + \alpha_2 = 1 . \quad (5)$$

The function has six unknown parameters. Taking into account condition (5), their number can be reduced to five. These five parameters are calculated from the system of equations [Gao et al., 2020]

$$\left.\begin{aligned} \alpha_1\mu_{1,1}+\alpha_2\mu_{1,2}&=0\\ \alpha_1\mu_{2,1}+\alpha_2\mu_{2,2}&=1\\ \alpha_1\mu_{3,1}+\alpha_2\mu_{3,2}&=\mu_3\\ \alpha_1\mu_{4,1}+\alpha_2\mu_{4,2}&=\mu_4\\ \alpha_1\mu_{5,1}+\alpha_2\mu_{5,2}&=\mu_5 \end{aligned}\right\}, \qquad (6)$$

where

$$\mu_{j,i}=\int \xi^j \frac{1}{\sqrt{2\pi}\sigma_i}\exp\left(-\frac{(\xi-m_i)^2}{2\sigma_i^2}\right)d\xi. \qquad (7)$$

Here, the first index indicates the order of the statistical moment, the second index indicates the number of the component of the Gaussian mixture.

When the statistical moment μ_5 is unknown, one of the parameters (the weight coefficient) remains free and must be determined from additional conditions [Zapevalov and Ratner 2003; Zapevalov and Knyazkov 2022]. As a result, there is ambiguity in the construction of the approximation (4).

Consider the situation when the approximation is carried out by expanding the desired function in terms of the derivatives of the standardized normal distribution

$$PN(\xi)=\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}\xi^2\right). \qquad (8)$$

Derived functions $PN(\xi)$ are defined by the expression.

$$\frac{d^n}{d\xi^n}PN(\xi)=(-1)^n H_n(\xi)\cdot PN(\xi), \qquad (9)$$

where $H_n(\xi)$ are Chebyshev-Hermite polynomials of order n that have the orthogonal property. The Edgeworth distribution can be written as

$$P(\xi)=PN(\xi)-\frac{\mu_3}{3!}PN^{\mathrm{III}}(\xi)+\frac{(\mu_4-3)}{4!}PN^{\mathrm{IV}}(\xi)+ +\frac{\mu_3^2}{6!}PN^{\mathrm{VI}}(\xi)-\frac{(\mu_5-10\mu_3)}{5!}PN^{\mathrm{V}}(\xi)+\ldots \qquad (10)$$

The expressions in parentheses are the cumulants of the fourth and fifth orders, respectively.

Taking into account that $\mu_3=\lambda_3$, we can draw the following conclusion. For problems of approximating the probability density function of sea surface elevations, it is necessary to know both the relationship between the parameters λ_3 and μ_5 and the relationship between the parameters λ_3 and λ_5.

3.2 Fifth order statistical moment

The dependence of the fifth statistical moment on the skewness is shown in Figure 1 When λ_3 changing in the range from -0.3 to 0.4, the values μ_5 mostly lie in the range from -3 to 5. There is a high correlation between the parameters λ_3 and μ_5, the correlation coefficient $r(\lambda_3,\mu_5)=0.92$. The dependence $\mu_5=\mu_5(\lambda_3)$ can be described by linear regression

$$\mu_5 = 9.28\lambda_3 + 0.11. \qquad (11)$$

The standard deviation is 0.56.

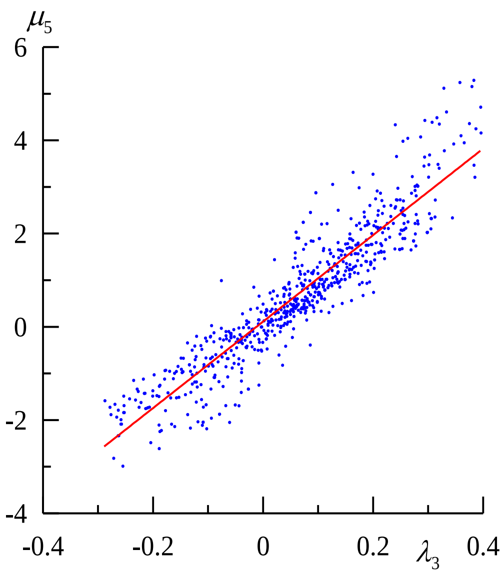

Figure 1: Dependence of the fifth statistical moment μ_5 on the skewness λ_3. Points are measurement data, a straight line is regression (11).

The parameters of a two-component Gaussian mixture are calculated from the specified values of statistical moments. The high correlation between λ_3 and μ_5 makes it possible to use regression (1) to calculate these parameters when μ_5 is unknown. The system of equations (3) becomes closed after substituting μ_5 calculated using (1) into the right part of the fifth equation.

3.3 Fifth order cumulant

The fifth order cumulant is a linear function of two arguments

$$\lambda_5 = \mu_5 - 10\,\mu_3\,. \tag{12}$$

Cumulant values vary from -1.4 to 2.0 with an average value of 0.07.

Dependencies $\lambda_5 = \lambda_5(\lambda_3)$ and $\lambda_5 = \lambda_5(\mu_5)$ are shown in Figure 2. The statistical relationship between λ_3 and λ_5, as well as between μ_5 and λ_5 is weak, the correlation coefficients are respectively equal to -0.18 and 0.21.

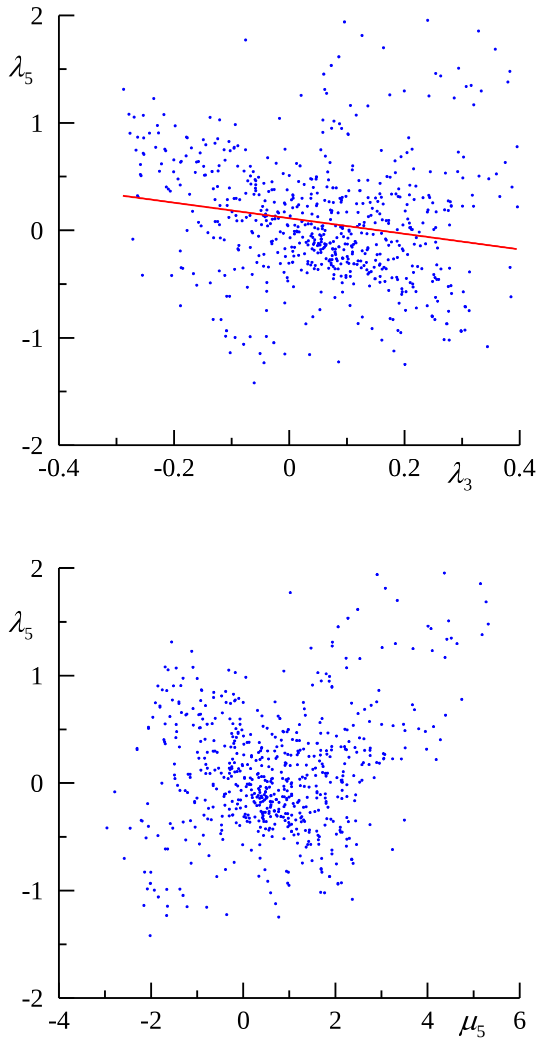

Figure 2: Functional connection λ_3 , μ_5 and λ_5 .

The red line on the upper fragment of the Figure 2 shows the dependence

$$\lambda_5^{(R)} = \mu_5^{(R)} - 10\,\lambda_3\,, \tag{13}$$

where the fifth statistical moment $\mu_5^{(R)}$ is calculated according to the regression (11). The standard deviation $\lambda_5^{(R)}$ from λ_5 is 0.54.

4 CONCLUSIONS

The analysis of wave measurements carried out in the field showed the following. The fifth statistical moment of sea surface elevation varies from -3 to 5. It can be calculated using linear regression (11) for the third statistical moment with a standard error of 0.56. The correlation coefficient between the statistical moments of the third and fifth orders is 0.92. This result, in particular, allows us to close a system of equations for calculating the parameters of a two-component Gaussian mixture approximating the probability density function of surface elevations.

The fifth cumulant of surface elevations varies from -1.41 to 1.96. It is weakly correlated with the third statistical moment. The correlation coefficient is -0.18. According to the known values of the third statistical moment, the fifth cumulant can be restored using regression (13) with a standard error of 0.54. The possibility of using regression (13) to change the order of truncation of the Edgeworth series when modeling the probability density function needs additional analysis.

ACKNOWLEDGEMENTS

The investigation was carried out within the framework of the state task on theme FNNN-0555-2021-0004 "Fundamental studies of the oceanological processes which determine the state and evolution of the marine environment influenced by natural and anthropogenic factors, based on the observation and modeling methods".

REFERENCES

Annenkov, S.Y., Shrira, V.I., 2013. Large-time evolution of statistical moments of wind–wave fields. *J. Fluid Mech.* 726, pp. 517–546.

Aprausheva, N.N., Sorokin, S.V., 2013. Exact equation of the boundary of unimodal and bimodal domains of a two-component Gaussian mixture. *Pattern Recognition and Image Analysis*, Vol. 23, No. 3. pp. 341–347.

Babanin, A.V., Polnikov, V.G., 1994. On non-Gaussian wind waves. *Physical Oceanography*. No. 3. pp. 241–245.

Blinnikov, S., Moessner, R., 1998. Expansions for nearly Gaussian distributions. *Astron. Astrophys. Suppl. Ser.*, Vol. 130, pp. 193–205.

Burdyugov, V.M., 2022. Modeling the shape of the reflected radio pulse of an altimeter using a truncated distribution of sea surface elevations. *Processes in GeoMedia*. № 3 (33). pp. 1724-1729.

Gao, Z., Sun, Z., Liang, S., 2020. Probability density function for wave elevation based on Gaussian mixture models. *Ocean Engineering*, 213, 107815.

Guedes Soares C., Cherneva Z., Antão E.M., 2004. Steepness and asymmetry of the largest waves in storm sea states. *Ocean Engineering* Vol. 31. pp. 1147–1167.

Hou, Y., Song, G., Zhao, X., Song, J., Zheng, Q., 2006. Statistical distribution of nonlinear random water wave surface elevation. *Chinese Journal Oceanology and Limnology*. Vol. 24, No. 1. pp. 1-5.

Kendall, M.J., Stewart, A., 1958. *The Advanced Theory of Statistics. Vol. I. Distribution theory*. London: Butler & Tanner Ltd.

Kwon, O.K., 2020. Analytic expressions for the positive definite and unimodal regions of Gram-Charlier series. *Communications in Statistics - Theory and Methods*. doi: 10.1080/03610926.2020.1833219

Mikhailichenko, Yu., Garmashov, A.V., Fomin, V.V., 2016. Verification of the SWAN wind waves model by observations on the stationary oceanographic platform of the Black Sea hydrophysical polygon of RAS. *Ecological Safety of Coastal and Shelf Zones of Sea.* No.2, pp. 52-57. (in Russian)

Shokurova, I.G., Garmashov, A.V., Toloknov, Yu.N., Korovushkin, A.I., 2016. Analysis of the meteorological regime in the north-western part of the Black Sea from observational data on fixed offshore platform. *Ecological Safety of Coastal and Shelf Zones of Sea*. No. 2. pp. 41-51. (in Russian)

Solov'ev, Yu.P., Ivanov, V.A., 2007. Preliminary results of measurements of atmospheric turbulence over the sea. *Physical Oceanography*. Iss. 3. pp. 154 - 172.

Zapevalov, A.S., Garmashov, A.V., 2021. Skewness and kurtosis of the surface wave in the coastal zone of the Black Sea. *Physical Oceanography*. Vol. 28, No. 4, pp. 414-425.

Zapevalov, A.S., Knyazkov, A.S., 2022 Statistical description of the sea surface by two-component Gaussian mixture. *Physical Oceanography*. Vol. 29, No. 4. pp. 395-403.

Zapevalov, A.S., Ratner, Yu.B., 2003. Analytic model of the probability density of slopes of the sea surface. *Physical Oceanography*. Vol. 13, No. 1. pp. 1-13.