CHAPTER 7

Method of quantile function (BOLIVAR)

The basis of the POT method is independence of wave heights in different storms. In the previous chapters we showed, however, that a storm sequence may be regarded as a Markov's chain. In order to exclude the limitations of the POT method and to take into account the asymptotic characteristics of AMS, let us consider n samples, consisting of heights h^+_{ij} of the largest waves in the *k* strongest storms in year number *i,(i=1,..,n; j=1,…,k).*

In accordance with definition (6.1), each of these samples is comprised of wave heights observed during different storms (no more than one height is taken from each storm). Let us sort each of the samples in decreasing order. This will give us the following ranked series:

$$\begin{array}{l} \left(h^+_{11} \geq h^+_{12} \geq \ldots \geq h^+_{1k}\right)\ 1^{st}\ year \\ \left(h^+_{21} \geq h^+_{22} \geq \ldots \geq h^+_{2k}\right)\ 2^{nd}\ year \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \left(h^+_{n1} \geq h^+_{n2} \geq \ldots \geq h^+_{nk}\right)\ n^{th}\ year \end{array} \qquad (7.1)$$

The number k of elements in each sample may vary but $k \geq 1$. For *k =1* (i.e. one storm in a year) this constitutes a sample of the highest waves.

$$\left(h^+_{11}, \ldots, h^+_{n1}\right) \qquad (7.2)$$

for all *n* years considered. Their distribution was discussed in Chapter 2 of this review.

The maximum wave height h_{max} which is possible once in n years, is the extreme (i.e. last term) element of the samples, both (7.2) and (7.1).

The order statistics h^+_{ij} are estimates of quantiles x_p. Their probabilistic properties are represented by the joint probability distribution function

$$G(x_1, \ldots, x_k) = P\left\{h^+_{i1} < x_1, \ldots, h^+_{ik} < x_k\right\} \qquad (7.3)$$

This is called the quantile function. A method of estimation of highest waves, which is based on equations (7.1)-(7.3), is called BOLIVAR [Boukhanovsky, Lopatoukhin, Rozhkov, 1997; Boukhanovsky, Lopatoukhin, Rozhkov, 1998 (a,b); Rozhkov et al., 1999; Boukhanovsky, Lavrenov, Lopatoukhin, Rozhkov, 1999]. Its name is derived from characters of the author names.

Heights h^+ of the highest waves in a sequence of storms in a single year can be considered correlated. Fig. 7.1 shows point diagrams of the highest waves h^+_{i1} and h^+_{i2} in the two strongest storms during an individual year in the Black, Baltic, and Barents Seas.

It can be seen from Fig. 7.1 that there is a relationship between h^+_{i1} and h^+_{i2}. This is related to the fact that h^+ in the second strongest, the third and other corresponding storms in a single year must be lower than the maximum one by definition, which generates a correlated sequence of maximum wave heights in storms.

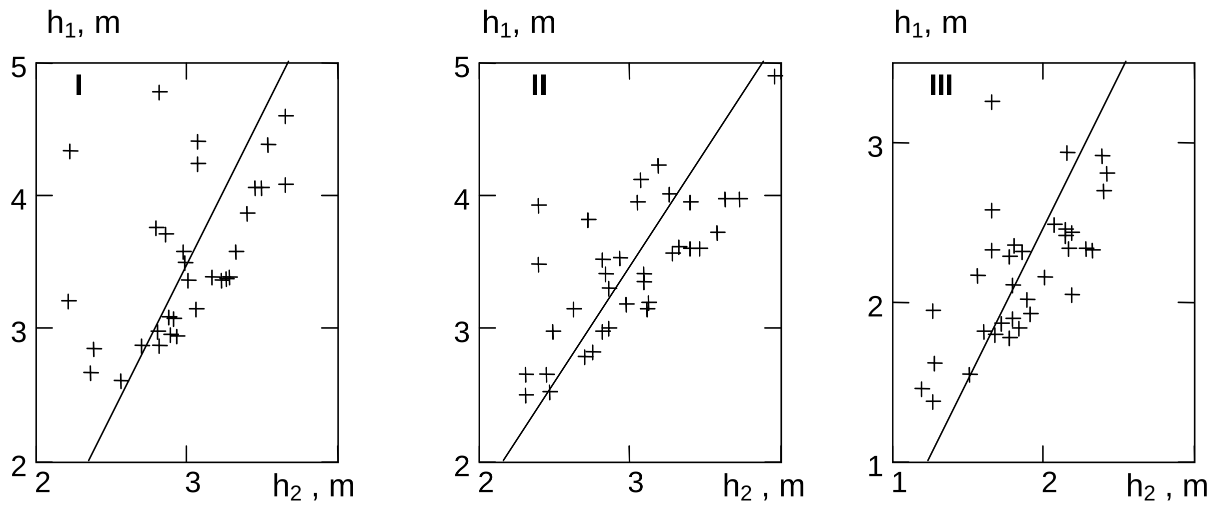

Figure. 7.1: Point diagram of first and second annual maxima (+) of mean wave heights in the Black (I), Baltic (II) and Barents (III) Seas. The solid line represents linear regression.

It is known from statistics that ranking of a sample with the distribution density *f(x)*, even if initially not correlated, leads to correlation between its statistic of i[th] and j[th] order [David, 1969]:

$$cov\left(x_i^*, x_j^*\right) \approx \frac{i}{n}\left(1-\frac{j}{n}\right)\frac{1}{nf(x_{i/n})f(x_{i/n})}$$

In particular, for $n \approx 30–40$ the correlation between the first two elements in the ranked sample reaches 0.6 – 0.7. Assuming that not only the extreme element of the sample h_{i1}^+ but also the following element h_{i2}^+ has the same asymptotic distribution , we get:

$$G(x_1, x_2) = G(x_1)G(x_2|x_1) \qquad (7.4)$$

where *G(•)* is Gumbel distribution (2.2).

Then the parameters of the conditional distribution $G(x_2|x_1)$ should depend on parameters of $G(x_1)$. Fig. 7.2 shows a comparison of empirical distributions of the first (annual highest), second and third highest wave heights with the theoretical distribution (2.2).

Fig. 7.2 confirms the hypothesis that the empirical data is distributed according to (2.2). Fig. 7.3 shows quantile diagrams (the median, upper and lower 10% bounds of the distribution) of the first eight maxima of wave height. It can be seen that distributions of ranked wave heights in a single year are similar. Thus, with some degree of approximation, it is possible to produce certain generalized relations between medians $Me(h_k^+)/Me(h_1^+)$ and r.m.s deviations $\sigma_{h_k^+}/\sigma_{h_1^+}$ of the first and other maxima. These are represented in fig. 7.4.

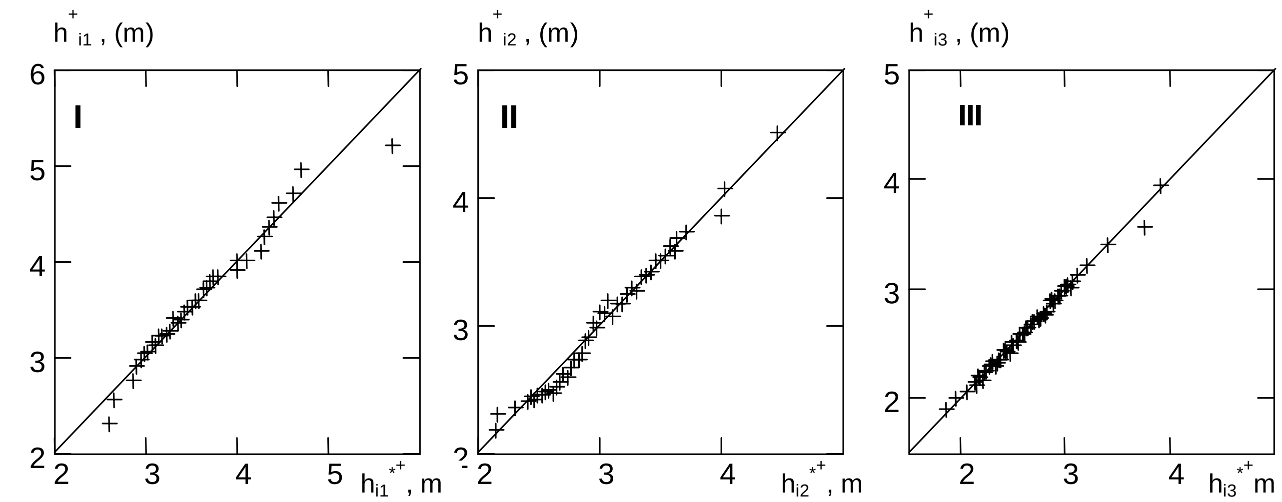

Figure.7.2: Bi-plots of annual maxima distribution.
The Black sea. I, II, III: the first, second, and the third annual maxima, respectively.

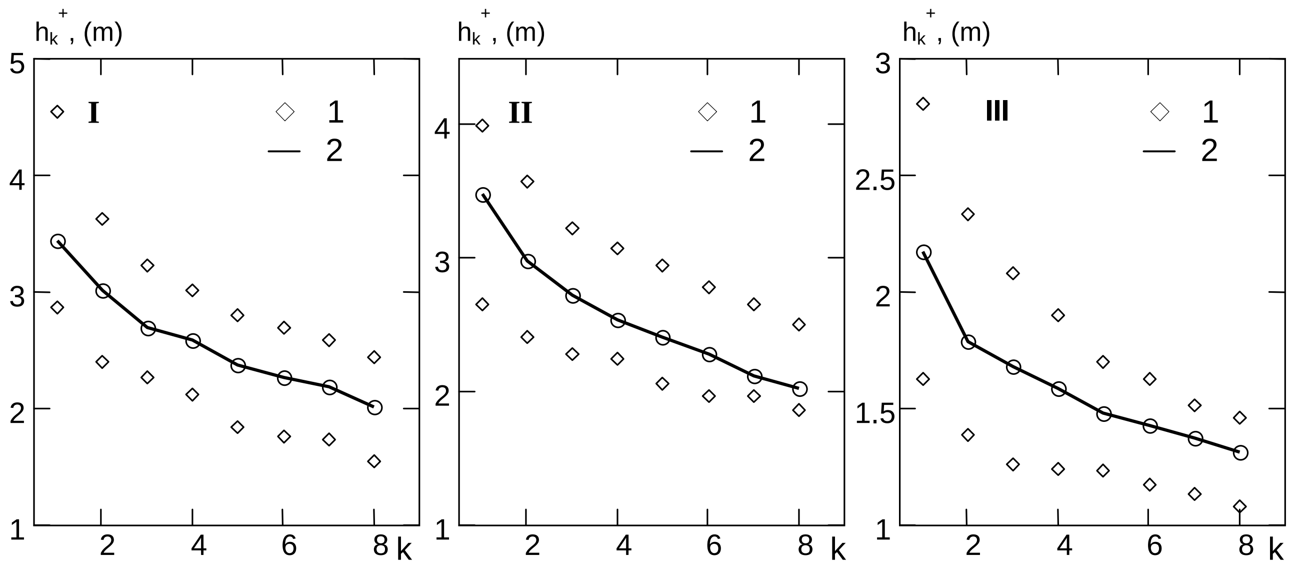

Figure. 7.3: Quantile diagrams for consequential annual maxima h_k in the Black (I), Baltic (II) and Barents (III) Seas. 1: upper and lower 10% bounds of distribution, 2: median.

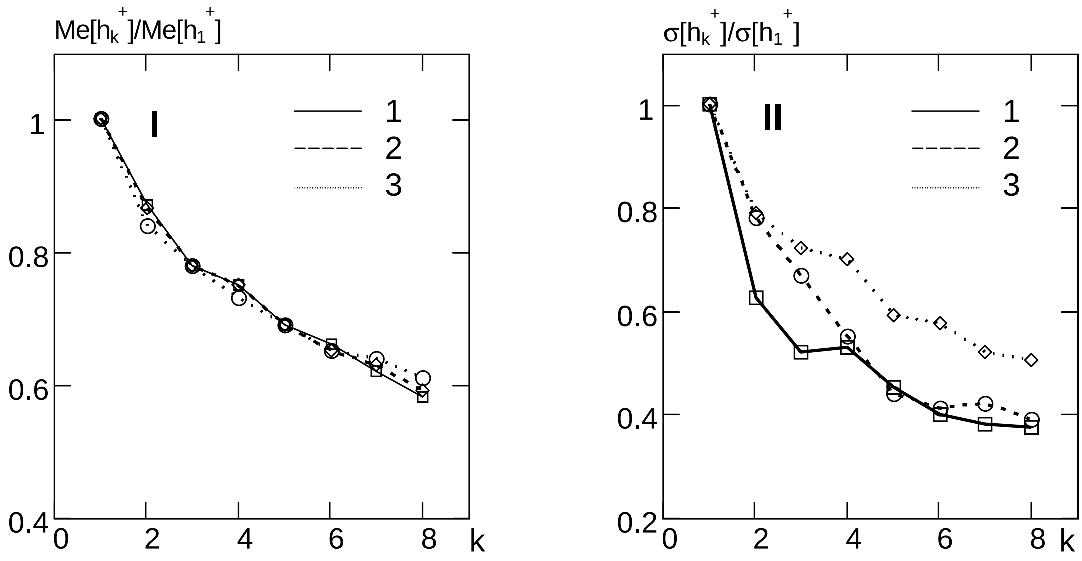

Figure. 7.4. Normalised median (I) and r.m.s deviation (II) of consequential annual maxima h_k.
1: Black Sea, 2: Baltic Sea, 3: Barents Sea.

Stochastic simulation based on the storm and weather window model (6.9) is the main technique for computing the distribution (7.3). It takes into account both intra-annual and year-to-year cycles. Fig. 7.5 shows quantile diagrams (the median, upper and lower 10% distribution bounds) computed using a 35-year (see table 6.1) and a 100-year simulated series, from (6.9). It follows from Fig. 7.5 that the agreement between observed and simulated data is satisfactory, which, in turn, confirms the correct choice of probabilistic model.

Let us consider distributions (7.3) and (5.1). Let p_k denote the probability that during the year number i the number of storms of certain intensity was k. Then the multi-dimensional distribution of probability of maximum wave heights in a sequence of storms that exceed a specified threshold will be

$$F(x_1,\dots x_m)=\sum_{k=1}^{m} p_k(x_1,\dots,x_k),\ m=1,2,3\dots$$

Distribution (7.5) generalizes (5.1) because it does not require an assumption that wave heights in the storms are independent. Distribution (7.3) is valid for waves from different storms. According to the way the distribution (7.1) was constructed, it makes it possible to distinguish wave heights that are to be expected once in 100 or 50 years, and also to estimate wave heights h_1^+ and h_2^+, that can take place during a single year during an interval of n years. For example, the height of the first annual maximum in the southern part of the Baltic Sea, at a return period of 100 years, is $h_1^{(100)}$ =5.1 m.

This coincides with the prediction by the AMS method. The second annual maximum, at 100 year return period, is $h_2^{(100)}$ =4.1 m.

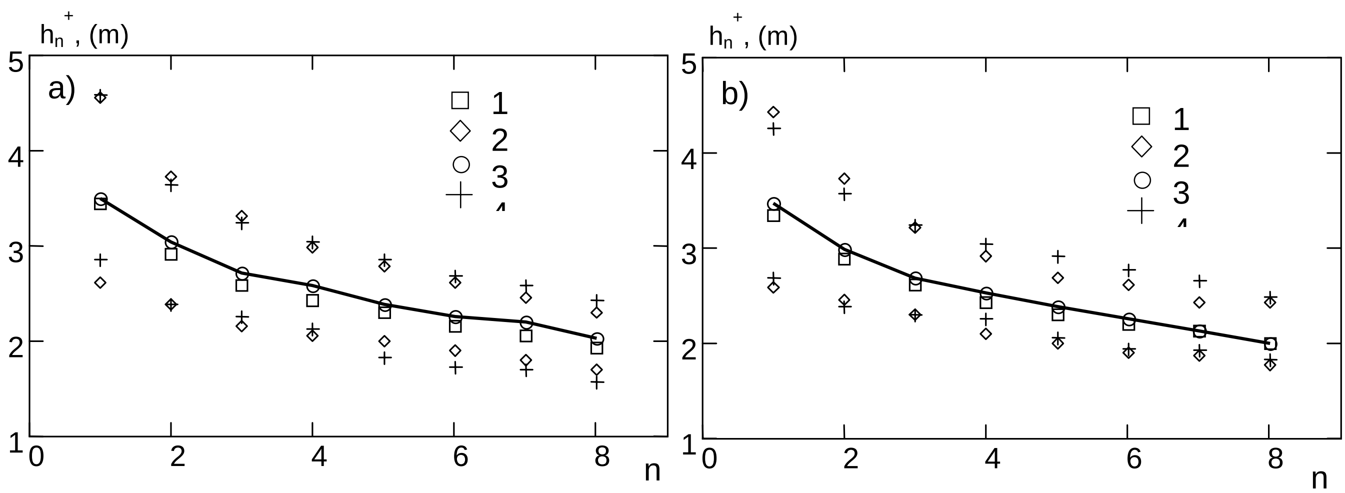

Figure 7.5: Medians (1,3) and deciles (2,4) of wave heights in the eight strongest storms in (a) the Black and (b) the Baltic Seas. (1,2): data from Table 6.1. (3,4): simulations using the probabilistic model (6.9).

This differs from the wave height possible once in 50 years, which is $h_1^{(50)}$=5.0 m. For a point in the Barents Sea we obtained the following estimates: $h_1^{(100)}$= 6.4m, $h_2^{(100)}$= 4.6m, $h_1^{(50)}$= 6.0m, $h_2^{(50)}$=4.4 m.

Fig. 7.6 shows distributions of the first, second, and third annual maxima in the southern part of the Baltic Sea on the Gumbel probability plot (see (2.5)). They were computed using (7.3). In addition, the figure contains distributions *G(h)* of all maxima exceeding *Z* =3.9 m and *Z*=3.2 m.

It can be seen from fig. 7.6 that for wave heights exceeding a sufficiently high value (i.e. 3.9 m), the contribution of the second maxima to distribution *G*(*h*) is insignificant. Distribution *G(h)* corresponds closely to that of all first annual maxima (up to the accuracy of assigning a certain probability $p_n=1/\lambda T$ to its quantiles).

At the same time, the question of whether the distribution can be reliably estimated using data on its several upper quantiles remains open. It can be seen that for a somewhat lower threshold level *Z= 3.2 m*, data (6) denoted by circles represent a mixture of first, second, and third annual maxima. This gives more weight to the “left” tail of the distribution and, in turn, leads to lower estimates for long return period wave heights.

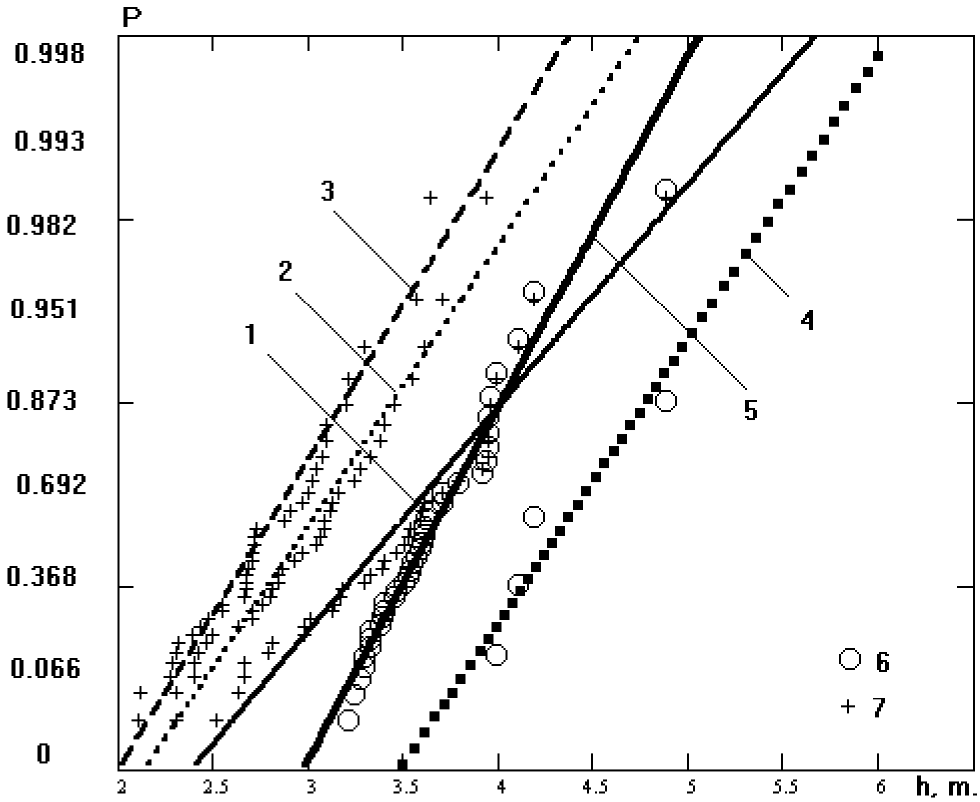

Figure 7.6: Distribution of annual wave height maxima.
1,2,3: The first, second, and third maxima simulated using (7.3). 4: Distribution of all maxima exceeding Z=3.9(m). 5: Distribution of all maxima exceeding Z=3.2 (m). 6: A mixture of all strong storms.
7: Samples made of observed first, second, and third maxima.

---oooOooo---