



# Article Research on Threshold Selection Method in Wave Extreme Value Analysis

Huashuai Liu <sup>1,2,3,\*</sup>, Fan Yang <sup>1,2,3</sup> and Hongchuan Wang <sup>1,2,3</sup>

- <sup>1</sup> Nanjing Hydraulic Research Institute, Nanjing 210029, China; fyang@nhri.cn (F.Y.); hcwang@nhri.cn (H.W.)
- <sup>2</sup> Key Laboratory of Port, Waterway and Sedimentation Engineering of MOT, Nanjing 210024, China
- <sup>3</sup> State Key Laboratory of Hydrology Water Resources and Hydraulic Engineering, Nanjing 210029, China
- \* Correspondence: lhuashuai@126.com

Abstract: Climate change poses higher requirements on ocean engineering design, and reasonable estimation of design wave heights plays a crucial role in coastal protection and offshore engineering. Extreme value analysis is widely used in frequency calculations of wave parameters, among which the peak over threshold method based on the generalized Pareto distribution is proven to be an effective method, and the different selection of extreme wave samples in this method has a great influence on the calculation results. In this study, long-term significant wave height series were utilized to investigate the long-range correlation of significant wave heights, and thresholds were determined based on the changes of long-range correlations. This approach assumes that extreme events are generally caused by different physical processes, where extreme events result from massive disturbances leading to abnormal states, and long-range correlations are not affected or minimally affected by extreme events. Thus, thresholds can be determined based on changes of long-range correlations by removing extreme events. Comparing this method to graphical diagnostic techniques, we demonstrated its rationality in determining extreme wave height thresholds. Moreover, the automatic threshold selection offered by this method helps to mitigate errors associated with subjective judgments in traditional approaches.

**Keywords:** design wave height; extreme value analysis; threshold; multifractal detrended fluctuation analysis

# 1. Introduction

The determination of extreme significant wave height and its return period plays a crucial role in coastal protection and the design criteria of marine structures [1–4]. To achieve this estimation, extensive research has been conducted on extreme value theory, primarily focusing on sampling methods and probability distribution models [5–9].

Typical sampling methods include the block maxima method (BM) [5] and the peak over threshold (POT) method [7,10]. The basic idea behind these methods is to first identify extreme value samples, then fit these samples using extreme value distribution models to construct long-term distributions, and finally extrapolate the required return values. The challenge with using the block maxima method lies in the lack of long-term observation data and the potential omission of informative data [11]. For instance, if some sub-maxima within an interval are larger than the maxima in another interval, it indicates that this method not only wastes useful information but also increases the inaccuracy of parameter estimation. To address this issue, the POT method [12] is employed to identify and select wave height values that exceed a given threshold from the original wave data [13]. This increases the number of samples and reduces the error in parameter estimation. The POT method has been considered more reliable in many studies due to its ability to gather more sample information [14,15].

In practice, when using the POT method, the first challenge is to select an appropriate threshold. The threshold must be high enough to ensure convergence to the generalized



Citation: Liu, H.; Yang, F.; Wang, H. Research on Threshold Selection Method in Wave Extreme Value Analysis. *Water* 2023, *15*, 3648. https://doi.org/10.3390/ w15203648

Academic Editors: Ruey-Syan Shih and Chia-Ming Fan

Received: 1 September 2023 Revised: 4 October 2023 Accepted: 15 October 2023 Published: 18 October 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). Pareto distribution (GPD), but not too high to allow enough values in the POT sample for reliable estimation. Several scholars have proposed various threshold selection methods, such as graphical diagnostic methods [16,17], probability-based methods [18], and hybrid methods [19–21]. Among these proposed methods, the graphical diagnostic method [16] has been widely accepted and used. Researchers [22-24] have employed the graphical diagnostic method to determine the threshold. The advantage of this method is that it allows the understanding of data characteristics through graphical analysis. However, the drawback is that it may produce multiple thresholds, making it challenging to determine the precise value. It requires expert judgment, and, therefore, the sensitivity of the threshold to the return period is often analyzed to ensure the rationality of threshold selection. To address these drawbacks, Thompson et al. [25] proposed an automatic threshold selection method based on the characteristic of the modified scale parameter (ATSMP). This method presents a pragmatic automated, simple, and computationally inexpensive threshold selection method based on the distribution of the difference of parameter estimates when the threshold is changed. However, the threshold selected by this method is usually not unique. To determine a suitable threshold within a stable range, Liang et al. [26] introduced the automatic threshold selection method based on the characteristic of extrapolated significant wave heights (ATSME), which can obtain a unique threshold value. Compared with the GPD parameter plot, the ATSME can select a suitable threshold without subjective effects and additional work. Compared with the ATSMP, the ATSME simplifies the process of threshold selection, and additional work is not needed.

While there are many threshold selection methods available, it is challenging to determine which method is optimal. Due to the importance of threshold selection in return period calculations, further research on threshold selection methods is necessary [27]. Shao et al. proposed an automated sampling method for extracting independent and identically distributed samples from long-term wave processes in the Yellow Sea. They studied the differences in extreme wave heights between tropical cyclones and non-tropical cyclones and evaluated the impact of considering tropical cyclone extreme wave heights on return periods. However, when applying this method to regions with complex weather conditions, further analysis is needed. Scarrott et al. [28] pointed out that extreme and non-extreme events are caused by different physical processes. Ozger [29] used detrended fluctuation analysis to study the long-range correlation of significant wave height series. For extreme wave events, their occurrences are influenced by external extreme factors and are very rare, leading to minimal long-term dependence on the normal wave sequence, even negligible. For instance, if the original significant wave height series undergoes data transformation and extreme wave height data are removed, it has almost no effect on the original trend. In contrast, removing non-extreme wave height data would result in significant differences from the original trend. Therefore, to obtain the threshold for extreme significant wave heights, the variation of long-range correlation exponent of significant wave height series can be analyzed. The multifractal detrended fluctuation analysis (MF-DFA) method proposed by Kantelhardt et al. [30] can be used to study the long-range correlation of time series. The MF-DFA method considers the long-range correlation characteristics of the data themselves when determining the threshold for extreme events, making it an objective approach. This method has been successfully applied in various fields. Du et al. [31] used the percentile method, fixed threshold method, and MF-DFA method separately to determine the threshold for extreme precipitation. Among these three methods, the MF-DFA method was considered the most objective and scientific method. Liu et al. [32] analyzed the applicability of non-parametric methods, parametric methods, and the DFA method in determining thresholds for extreme precipitation. The research findings indicated that the DFA method was proven to be the most suitable approach, capable of providing a set of suitable thresholds for extreme precipitation in large regions characterized by spatiotemporal heterogeneity. Zhang et al. [33] employed the universal multifractal analysis method to analyze extreme precipitation data, demonstrating that this method takes into consideration both the physical processes and probability distribution of

precipitation. It provides a formal framework for the spatiotemporal assessment of extreme precipitation at the regional scale. Du et al. [34] determined the threshold for extreme temperatures using the MF-DFA method. In the paper, it is argued that the DFA exponent obtained through the MF-DFA method can identify the long-range correlations in the system's evolution over a certain period. Furthermore, extreme events do not significantly affect the long-term correlations of the entire sequence, or even if there is an effect, it is minimal. Therefore, the DFA exponent can effectively serve as the threshold for extreme events. However, the use of the MF-DFA method for identifying extreme values was predominantly concentrated in meteorology, with no apparent application in the field of ocean engineering. Furthermore, there was no literature discussing its suitability for threshold selection in extreme value analysis. Therefore, this paper extends its application to the domain of ocean engineering. The paper employs this method to determine extreme wave heights and applies them to extreme value analysis, evaluating their feasibility in this context.

The present study utilizes the MF-DFA method to determine the threshold for extreme significant wave heights by analyzing the variation of the long-range correlation index in the significant wave height series. Firstly, the long-range correlation in the significant wave height series is detected. Then, the threshold for extreme significant wave height events is determined by the property that extreme values have minimal impact on long-range correlation is utilized. The threshold obtained through this method is compared with thresholds obtained through other methods, and the GPD is used for return period analysis to evaluate the rationality of this approach. The structure of this paper is as follows:

In Section 2, the theoretical methods as well as the dataset are introduced. The Section 3 presents the validation of the threshold determination using the MF-DFA method and compares it with the graphical diagnostic approach. The Section 4 discusses the rationale of MF-DFA threshold selection method. Finally, Section 5 provides a summary and conclusion of the study.

#### 2. Materials and Methods

# 2.1. POT Method

The POT method extracts a series of independent peak significant wave heights that are above a specified threshold as samples. Let the distribution function of the sequence be F(x), and define  $F_u(y)$  as the conditional distribution function of the random variable X that exceeds the threshold  $\mu$ :

$$F_{u}(y) = P(X - \mu \le x | X > \mu) = \frac{F(\mu + y) - F(\mu)}{1 - F(\mu)} = \frac{F(x) - F(\mu)}{1 - F(\mu)}$$
  
$$\Rightarrow F(x) = F_{u}(y)(1 - F(\mu)) + F(\mu)$$
(1)

When the threshold is sufficiently high, the conditional excess distribution function  $F_u(y)$  converges to the generalized Pareto distribution [7], with its cumulative distribution function given by:

$$F_{u}(y) \approx G(x; \mu, \sigma, \xi) = \begin{cases} 1 - \left[1 + \xi\left(\frac{x-\mu}{\sigma}\right)\right]^{-1/\xi}, \xi \neq 0\\ 1 - \exp\left(-\frac{x-\mu}{\sigma}\right), \xi = 0 \end{cases}$$
(2)

where  $\mu$ ,  $\sigma$ , and  $\xi$  represent the location, scale, and shape parameters, respectively. When  $\xi = 0$ , the GPD corresponds to the exponential distribution.

The quantile of the GPD model for a T-year return period is given by:

$$x_T = \begin{cases} \mu + \frac{\sigma}{\xi} [(\lambda T)^{\xi} - 1], \ \xi \neq 0\\ \mu + \sigma \ln(\lambda T), \ \xi = 0 \end{cases}$$
(3)

where  $\lambda = \frac{N_u}{N}$ ,  $N_u$  is the total number of exceedances of the threshold  $\mu$ , and N is the total number of years of records.

The reasonable determination of the threshold is a crucial prerequisite for correctly estimating the parameters of the GPD model. If the selected threshold is too large, it may result in too few exceedance data, leading to larger parameter variances in the estimation. On the other hand, if the selected threshold is too low, it may not guarantee the convergence of the distribution, causing significant estimation bias. In this study, three different methods, namely multifractal detrended fluctuation analysis, mean residual life plot method, and threshold stability plot method, are utilized to determine the threshold. Each method is used to fit the exceedance samples to the GPD, and the rationality of the threshold selection methods is compared.

## 2.2. Threshold Determination Using the Detrended Fluctuation Analysis Method

The specific algorithm steps of the MF-DFA method are as follows:

Step 1: For a significant wave height time series x(t), t = 1, 2, ..., N, determine the "profile"

$$Y(i) = \sum_{t=1}^{i} (x(t) - \overline{x}), i = 1, 2, \cdots, N$$
(4)

where  $\overline{x}$  is the mean of x(t).

Step 2: Divide the profile Y(i) into  $N_s = int (N/s)$  nonoverlapping segments of equal length s. Since the length N of the series is often not a multiple of the considered time scale s, a short part at the end of the profile may remain. In order not to disregard this part of the series, the same procedure is repeated starting from the opposite end. Thereby,  $2N_s$  segments are obtained altogether.

Step 3: Calculate the local trend for each of the  $2N_s$  segments by a least-square fit of the series. Then determine the variance

$$F^{2}(v,s) = \frac{1}{s} \sum_{i=1}^{s} \left\{ Y[(v-1)s+i] - y_{v}(i) \right\}^{2}$$
(5)

for each segment  $v, v = 1, 2, ..., N_s$  and

$$F^{2}(v,s) = \frac{1}{s} \sum_{i=1}^{s} \left\{ Y[N - (v - N_{s})s + i] - y_{v}(i) \right\}^{2}.$$
 (6)

for  $v = N_s + 1$ ,  $N_s + 2$ , ...,  $2N_s$ . Here,  $y_v(i)$  is the fitting polynomial in segment v. In this paper, we set the polynomial as linear and  $s_{\min} = 6$ ,  $s_{\max} = 3000$ .

Step 4: Calculate the *q*th order fluctuation function

$$F_q(s) = \left\{ \frac{1}{2N_s} \sum_{v=1}^{2N_s} \left[ F^2(s,v) \right]^{q/2} \right\}^{1/q}$$
(7)

for  $q \neq 0$  and

$$F_0(s) = \exp\left\{\frac{1}{4N_s} \sum_{v=1}^{2N_s} \ln\left[F^2(s, v)\right]\right\}$$
(8)

for q = 0. For q = 2, the standard DFA procedure is retrieved.

Step 5: Determine the scaling behavior of the fluctuation functions by analyzing log–log plots  $F_q(s)$  versus *s* for each value of *q*.

$$F_a(s) \propto C s^{h(q)} \tag{9}$$

The slope of log  $F_q(s)$  and log *s* is the generalized Hurst exponent h(q). In this paper, the standard DFA method is used for analysis, i.e., the case of q = 2. The scaling exponent

H = 0.5 means that the time series are uncorrelated; 0 < H < 0.5 implies the series presents an anti-persistent behavior and 0.5 < H < 1 implies that the original series has long-rang correlation. In this paper, the physical interpretation of the DFA exponent is not considered. Only the impact of extreme values on the DFA exponent is taken into account to determine the threshold for extreme events.

The process of determining the threshold for extreme events using the MF-DFA method for a given time series, denoted as {x(t), t = 1, 2, ..., N}, is as follows:

- (1) Determine the maximum  $x_{\text{max}}$  and minimum values  $x_{\text{min}}$  of the time series  $\{x(t)\}$ .
- (2) Determine the central point *R* of the sequence  $\{x(t)\}$ , you can either take the average of all data points or choose a median value that lies between the maximum and minimum values.
- (3) Starting from the maximum value of  $\{x(t)\}$ , sequentially discard data points within intervals  $\{x(t), x(t) \ge x_{\max} d \times k\}$  until reaching the central point *R*. In this process, a series of new sequence  $Y_J$ ,  $J = x_{\max} d \times k$  is obtained, where *d* is the interval size.
- (4) Calculate the fractal exponent  $D_J$  for each new sequence  $Y_J$  and observe how it changes with the discarded interval size *J*.
- (5) When the change in D<sub>I</sub> starts to become smooth and converges to the original DFA exponent of the data {x(t)}, take the corresponding J value as the threshold for extreme events in the sequence {x(t)}. The degree of convergence to the original value is not unique and may fluctuate slightly around the original exponent. Therefore, to determine the convergence point, the variance var<sub>j</sub> of the sequence of exponents D<sub>I</sub> can be calculated. Variance can be defined as follows:

$$\operatorname{var}_{j}^{2} = \frac{1}{N-1} \sum_{j=1}^{N} \left( DFA_{j} - E \right)^{2}$$
(10)

$$E = \frac{1}{N} \sum_{j=1}^{N} DFA_j \tag{11}$$

In this paper, we take d = 0.01, R = 2. In the above formula,  $DFA_j$  represents the DFA exponent of the j-th sequence, N represents the total number of sequences, and E represents the average of all exponents  $DFA_j$ . The convergence point of  $DFA_j$  can be determined by identifying the converging interval var<sub>j</sub>. The paper employs a BG algorithm for detecting the turning points in the sequence, which corresponds to the threshold of significant wave heights that cause a turning in long-range correlation. The chi-square test is used to detect the convergence point. When there are multiple turning points, the critical points must adhere to the following three criteria for the chi-square test: (1) the critical turning points should pass the significance level test; (2) the points before the critical turning points should not exhibit significant differences.

#### 2.3. Method Validation

Randomly generate two non-stationary time series  $\{x_1\}$  and  $\{x_2\}$ , each with a length of 10,000 and different degrees of long-range correlation. Randomly select any points from them and alter their values such that the selected values are greater than the maximum value in the original sequence and smaller than the minimum value in the original sequence, defining them as extreme value points. Obtain new sequences  $\{x_{11}\}$  and  $\{x_{22}\}$ , and use the method described above to calculate how the DFA exponent changes as extreme values are reduced in the series.

From Figure 1, it can be observed that after excluding the extreme values from the sequence, the DFA exponent of the resulting new series exhibits little to no change compared to the DFA exponent of the original series. As more extreme value points are removed, up to the point of the original sequence's maximum value, the DFA exponent of the new series

1.00 0.90 0.80 0.70 0.60 0.50 0.40 0.30 0.20 -25.0 25.0 -30.0 -20.0 -15.0 -10.0 -5.0 0.0 5.0 10.0 15.0 20.0 30.0 Cert-53-65-8-64 .... 4 1.45 -10 -30 -25 -20 -15 0 20 25 30 -5 5 10 15 x11 0.90 0.80 0.70 0.60 0.50 0.40 0.30 0.20 0.10 -30.0 -25.0 -20.0 -15.0 -10.0 -5.0 5.0 10.0 15.0 20.0 25.0 30.0 0.0 8. 19 - 2007 VOGNARZ AND AND A COMPANY -25 -20 20 25 30 -30 -15 -10 -5 0 10 15 x22

remains nearly equal to that of the original series. This suggests that the discarded points have minimal impact on the calculation of the DFA exponent for the original series and can be considered as extreme values in the current series.

**Figure 1.** The change in DFA exponent after excluding data from different intervals in a series with extreme values.

To validate the effectiveness of the method, the same procedure was applied to the original sequences  $\{x_1\}$  and  $\{x_2\}$ . If the method is effective, meaning extreme values have no impact on the DFA exponent of the series, then the extreme thresholds for series  $\{x_{11}\}$  and  $\{x_1\}$  should be the same, and the DFA exponents should also be the same. The same parameters were used during the calculation process, and the results are shown in Figure 2. From Figure 2, it can be observed that when the values are taken as 18.19 and -18.45, the DFA exponent of the new series precisely starts to converge to the exponent of the original series, the value of 18.19 and -18.45 can be considered as the thresholds for extreme values in the series  $\{x_1\}$ . The DFA exponent for series  $\{x_1\}$  is 0.8556, and for sequence  $\{x_{11}\}$ , it is 0.8541, with only minor differences in numerical values, indicating the method's effectiveness.



**Figure 2.** The change in DFA exponent after excluding data from different intervals in the original series.

# 2.4. The Study Area and Data

In the study, ERA-5 reanalysis data spanning from 1979 to 2020 were utilized. The daily maximum significant wave height data from multiple points in the South China Sea were extracted as the research samples. ERA-5 reanalysis data are the latest generation of reanalysis data created by the Copernicus Climate Change Service operated by ECMWF. It has higher resolution compared to its predecessor ERA-Interim, with a spatial resolution of 80 km and a temporal resolution of 1 h.

The reliability of ERA-5 reanalysis data for long-term significant wave height series has been validated in previous literature [35–37]. Table 1 and Figure 3 provide the location information and spatial distribution of the study points.

Location	Lat. (° N)	Lon. (° E)	Water Depth (m)	Maximum Wave Height (m)
P1	14	110	1274	9.75
P2	11	116	2671	5.74
P3	20	119	3032	12.65
P4	17	114	3053	9.55
P5	20	108	49	9.97
P6	21	116	128	10.54

Table 1. Location information of the study sites.



Figure 3. Distribution map of study locations.

#### 3. Results

#### 3.1. The Long-Range Correlation of the Significant Wave Height Series

Due to extreme events being caused by external disturbances leading to abnormal states and formed by different physical processes than non-extreme events, the DFA exponent obtained from the MF-DFA method measures the long-range correlation within a certain time scale, while the overall long-range correlation is not significantly affected by extreme events or is minimally influenced. Based on this idea, the threshold for extreme events can be determined by analyzing the variation of the DFA exponent. Before using this method, the long-range correlation in the significant wave height series should be examined. If the series exhibits long-range correlation, the threshold for extreme wave heights can be determined based on the variation of the DFA exponent. By applying the MF-DFA method described in Section 2 to analyze the significant wave height series, the values of the fractal exponent h(2) can be calculated, and it can be determined whether the significant wave height series at each location exhibits long-range correlation.

Based on the analysis of the significant wave height series using the MF-DFA method, the following results were obtained:

According to the Figure 4, the slopes of the fluctuation functions for points P1 to P6 were calculated, resulting in fractal exponents of the significant wave height series at each location: 0.81, 0.77, 0.81, 0.80, 0.59, and 0.79, all of which are greater than 0.5. This indicates that all six points exhibit positive long-range correlation. As a result, the MF-DFA method can be used to determine the threshold. Given the characteristic that long-range correlation is not significantly affected by extreme values or is minimally influenced by them, the extreme values are sequentially removed from the significant wave height series. The

variation in the fractal exponent is then calculated to determine the threshold for extreme significant wave heights.



Figure 4. The log-log plots of fluctuation function.

### 3.2. Threshold Determination

In Section 3.1, the long-range correlation in the significant wave height series for the selected points was examined, and the results showed that all the selected points exhibit long-range correlation. Using the MF-DFA method described in Section 2.2, the threshold for extreme significant wave heights was determined, and in the calculation of the variation of the DFA exponent, were set as d = 0.01, R = 2.

Figure 5 presents the DFA index variation sequence and variance sequence for the significant wave height at point P1, refer to Figure S1 in the supplementary materials for all sites. From the Figure 5, it is evident that there is a clear trend in the sequence. The BG algorithm was then applied to calculate the turning points, and the chi-square test was used to determine the convergence point. The calculation results are presented in Table 2 (using P1 as an example), refer to Table S1 in the supplementary materials for the calculation results of all sites.



Figure 5. Cont.



Figure 5. Determination of extreme wave height threshold using the MF–DFA method.

Table 2. Significance test of the turning point.

Turning Point	n	$\chi^2$	$\chi^2_{(a/2)}$	$\chi^2_{(1-a/2)}$	Significance Test
5.19	52	50.03	64.3	38.56	No
5.12	56	71.14	68.8	42.06	Yes
4.87	74	90.6	88.85	58.01	Yes
4.76	83	140.27	98.78	66.08	Yes
4.44	112	121.82	130.47	92.38	No
4.35	121	113.72	140.23	100.62	No

Figure 5 provides the variation of the DFA index with the removal of extreme significant wave height and the variance of the DFA index for point P1. Table 2 presents the significance test results for each site. According to the criteria for critical turning points, the threshold for extremely significant wave height at point P1 is chosen as 5.12 m. The thresholds for extremely significant wave height at the other points are determined to be 3.96, 6.06, 5.59, 4.45, and 5.15 m for points P2, P3, P4, P5, and P6, respectively.

# 3.3. Return Period

In addition to threshold selection, another important consideration when using the POT method is the choice of the minimum time span ( $\Delta$ t), which can impact the estimation of return periods [38]. The selection of the time span should be sufficiently long to ensure independence among samples, and different studies have provided varying results for the optimal range of  $\Delta$ t [39]. In this section, to determine the most suitable time span for the data used in this study, we investigate the influence of different time intervals ( $\Delta$ t) on the return period. The results, as shown in Table 3 (using point P1 as an example), indicate that a time interval of 5 days is selected as the optimal time span.

Table 3.	Variation	of return	periods	at different i	ntervals.

Return Periods	3 Days	4 Days	5 Days	6 Days	7 Days
50 year (m)	8.53	8.53	8.53	8.53	8.53
100 year (m)	9.15	9.15	9.13	9.13	9.13
150 year (m)	9.51	9.51	9.47	9.47	9.47
200 year (m)	9.77	9.77	9.71	9.71	9.71

Based on the threshold determined using the MF-DFA method and selecting a time interval of 5 days, independent samples above the threshold are obtained. These samples are then fitted to the GPD. Parameter estimation is performed for the GPD, and a test is conducted to validate the reasonableness of selecting this distribution.

From the probability density plot and quantile–quantile plot in Figure 6, it can be observed that the fitting of extreme samples to the distribution function is quite satisfactory,

indicating that the chosen threshold is appropriate. Using the same method, Table 4 provides the return periods of 50-year, 100-year, 150-year, and 200-year extreme wave heights for each location in the study under different threshold values.



Figure 6. Goodness of fit test at a threshold of 5.12 (PP plot and QQ plot).

Locactions\Return Periods	50 Year (m)	100 Year (m)	150 Year (m)	200 Year (m)
P1	8.53	9.13	9.47	9.71
P2	6.17	6.24	6.28	6.31
P3	11.38	12.32	12.87	13.26
P4	9.84	10.42	10.75	10.99
P5	8.73	9.38	9.74	10.01
P6	10.67	11.42	11.85	12.14

Table 4. Return periods of different locations.

# 3.4. Comparison with Other Threshold Selection Methods

To illustrate the importance of threshold selection in estimating return periods, we conducted a sensitivity analysis of return periods with respect to threshold variations and identified a stable range of thresholds that yield reliable return period estimates. Taking P1 point as an example, we considered a threshold range from 2.0 m to 6.5 m with an interval of 0.05 m, and a sample time interval of 5 days. The results of return periods varying with thresholds are shown in Figure 7. From Figure 7a, it can be observed that as the threshold increases within the range of 2 m to 4.6 m, the return values of significant wave heights gradually increase. This is because as the threshold increases, small significant wave height samples are gradually excluded from the analysis, and since these excluded samples are not extreme values, the remaining samples better represent the significant wave heights. Hence, the return values change continuously within this threshold range, making it inappropriate for threshold selection. Within the range of 4.6 m to 5.55 m, the return values of significant wave heights stabilize, indicating that the exclusion of samples due to an increase in the threshold no longer significantly affects the return values. At this point, the selected samples can represent extreme events while maintaining a sufficient sample size for reasonable return period estimation. Therefore, this range is generally considered suitable for threshold selection. For thresholds greater than 5.55 m, the confidence intervals of return values for significant wave heights exhibit considerable fluctuations, indicating substantial uncertainty in return period estimation within this threshold range. As the threshold increases, reasonable extreme wave height values are gradually excluded, and the sample size becomes insufficient, leading to significant errors in model estimation. Consequently, thresholds in this range are also not recommended for selection.



Figure 7. Stable threshold range.

To further demonstrate the rationality of the stable threshold range, Table 5 provides the variation of return values for different return periods within this range. From the table, it can be seen that the changes in return values for 50-year, 100-year, 150-year, and 200-year return periods are relatively small within the stable threshold range, indicating that threshold variations have little effect on the return periods within this range. Hence, selecting thresholds within this range is reasonable. Table 6 calculates the differences in return values of significant wave heights for neighboring thresholds for the 50-year, 100-year, 150-year, and 200-year return periods at the P1 point. The differences are very small within the stable threshold range, which further confirms the rationality of selecting the stable threshold range.

Table 5. The stable threshold range and corresponding return period of significant wave heights.

Return Periods	Stable Threshold Range (m)	Return Period Significant Wave Heights (m)	The Average of Wave Height (m)
50 year	(4.60, 5.65)	(8.50, 8.55)	8.53
100 year	(4.60, 5.55)	(9.09, 9.22)	9.16
150 year	(4.60, 5.55)	(9.43, 9.61)	9.52
200 year	(4.60, 5.55)	(9.65, 9.90)	9.78

Table 6. Differences in return periods within stable threshold range.

<b>Return Periods</b>	Stable Threshold Range (m)	Range of Differences (m)	Width of Differences (m)
50 year	(4.60, 5.65)	(-0.01, 0.09)	0.10
100 year	(4.60, 5.55)	(-0.05, 0.06)	0.11
150 year	(4.60, 5.55)	(-0.08, 0.10)	0.18e
200 year	(4.60, 5.55)	(-0.10, 0.14)	0.24

Table 7 provides the stable threshold ranges for all stations. Based on the sensitivity analysis of significant wave heights to thresholds, each threshold within the stable range can be selected as the theoretically appropriate threshold. When choosing the lowest threshold as the appropriate one, the threshold for different return periods can be uniquely

determined, and more samples can be used to extrapolate the significant wave heights for return periods with weaker estimation uncertainties. When choosing the median threshold as the appropriate one, the return significant wave heights can be more robust. On the other hand, when choosing the highest threshold as the appropriate one, the reliability of the return significant wave heights can be higher.

Table 7. Stable threshold ranges at each location.

Locations	P1	P2	P3	P4	P5	P6
Stable threshold range (m)	(4.65, 5.55)	(3.75, 4.95)	(5.90, 6.35)	(4.70, 5.70)	(4.10, 5.15)	(4.80, 5.55)

The above threshold sensitivity analysis highlights the importance of threshold selection for the estimation of return period results. To validate the effectiveness of the MF-DFA method proposed in this study for threshold determination, several commonly used threshold determination methods were chosen for comparison, namely, the mean residual life plot method and the parameter stability plot method [14].

#### 3.4.1. The Mean Residual Life Plot

Let  $X_1, X_2, \dots, X_n$  be a set of independent and identically distributed random variables with a common distribution F(x). Now, we choose a threshold  $u_0$ , and if  $X_1 > \mu, X_1 - \mu$  is the excess above the threshold, for any  $\mu > \mu_0$ , the Mean Excess Function (MEF)  $e(\mu)$  is defined as:

$$e(\mu) = E(X - \mu | X > \mu) = \frac{\sigma_{\mu 0}}{1 - \xi} + \frac{\xi}{1 - \xi}\mu$$
(12)

You can plot a scatter plot of the threshold  $\mu$  and the arithmetic mean of the excess values using the given equation. When the shape parameter  $\xi$  is a constant, this plot will approximate a straight line. In this plot, the threshold  $\mu$  is taken as the *x*-axis, and the arithmetic mean of the excess values is taken as the *y*-axis. The resulting line will have a slope of  $\frac{\zeta}{1-\zeta}$  and an intercept of  $\frac{\sigma_{\mu 0}}{1-\zeta}$ . Exactly, if after a certain threshold  $\mu_0$ , the slope of the line remains constant, then the point  $\mu_0$  where the slope stabilizes can be considered as the threshold value. In other words, after this point, the relationship between the threshold and the arithmetic mean of the excess values becomes more stable, indicating that the excess values are well-behaved and can be reliably modeled. Therefore,  $\mu_0$  is a suitable threshold value for the Mean Excess Function method.

From the Figure 8, and combined with the number of extreme value samples, we select the threshold as 4.80 m. To verify the rationality of this threshold selection, we provide the probability plot and quantile plot for the GPD fit. From Figure 9, it is evident that when the threshold is set to 4.80 m, there is little difference between the empirical values and the fitted values obtained from the GPD model. This result indicates a good fit for the chosen threshold. Refer to Figure S2 in the supplementary materials for the mean residual life plot of the remaining sites.

#### Mean Residual Life Plot



**Figure 8.** Mean residual life plot. Lines represent empirical mean residual life plot and confidence intervals.



Figure 9. Goodness of fit test at a threshold of 4.80 (PP plot and QQ plot).

# 3.4.2. Parameter Stability Plot

The method of selecting the threshold based on the mean excess plot has a certain subjectivity. In the mean excess plot, different interpretations of approximate linearity may lead to different choices of the threshold value  $\mu_0$ . Coles proposed using the stability of parameter estimates for determining the optimal threshold. If the generalized Pareto distribution is a reasonable fit for the excesses over the threshold, then the scale parameter  $\sigma_{\mu}$  can be estimated as:

$$\sigma_{\mu} = \sigma_{\mu_0} + \xi(\mu - \mu_0) \tag{13}$$

The above equation  $\sigma_{\mu}$  will vary with the threshold  $\mu$ , and to account for this variation, the corrected scale parameter is introduced as:

$$\sigma^* = \sigma_\mu - \xi\mu \tag{14}$$

In the above equation,  $\sigma^*$  is held constant with respect to the threshold  $\mu$ , so the corrected scale parameter and shape parameter should remain constant for values greater than the threshold  $\mu_0$ . By plotting the confidence intervals of  $\sigma^*$  and  $\xi$  against the threshold  $\mu$ , the appearance of constant segments can be used to determine the optimal threshold position.

Figure 10 shows the variation of the shape parameter and corrected scale parameter with the threshold, using point P1 as an example, for parameter stability plots of the

remaining points, please refer to Figure S3 in the supplementary materials. From Figure 10, it can be observed that when the threshold is in the range of (4.85 m, 5.40 m), both the shape parameter and the corrected scale parameter approach constant values. In order to obtain more extreme value samples, we choose the threshold to be as small as possible, and in this case, we select 4.85 m as the threshold. To validate the reasonability of the threshold selection, we provide the probability plot and quantile plot for GPD fitting, and the results in Figure 11 confirm that the threshold selection is reasonable, and the estimated uncertainty is acceptable.



Figure 10. Parameter stability plot.



Figure 11. Goodness of fit test at a threshold of 4.85 (PP plot and QQ plot).

### 4. Discussion

To illustrate the importance of proper threshold selection in the estimation of return periods, we analyzed the impact of threshold variations on the return periods. Figure 5 shows the variations of the 50-year, 100-year, 150-year, and 200-year return periods, highlighting the significance of selecting an appropriate threshold. Using the MF-DFA method, we determined a unique threshold based on the long-range correlation of the sequences

and compared it with traditional threshold selection methods to demonstrate the rationality of the MF-DFA method in threshold determination. Table 7 provides the stable threshold range for each station, indicating that the threshold variation within this range has a minor effect on the return periods. The reasonable threshold is highly likely to fall within this interval. In Table 8, we present the threshold selection results obtained from different methods, where the MF-DFA method provides a unique value, while the mean excess plot method and parameter stability plot method give possible ranges. The specific threshold value still requires the user's expertise for selection. Combining the stable threshold range provided in Table 7, we observe that the thresholds determined by the MF-DFA method are all within this stable range. Additionally, they also fall within the threshold selection ranges of the mean excess plot method and parameter stability plot method, further validating the rationality of using the MF-DFA method for threshold selection. Table 9 displays the 50-year, 100-year, 150-year, and 200-year return periods at P1 point obtained using different thresholds. The results show that the return periods obtained using these methods have relatively small differences, further confirming the reliability of the MF-DFA method.

Locations	MF-DFA (m)	Mean Excess Function (m)	Parameter Stability Plot (m)
P1	5.12	(4.80, 5.40)	(4.85, 5.40)
P2	3.96	(3.90, 4.70)	(3.95, 4.40)
P3	6.06	(4.90, 6.50)	(5.80, 6.10)
P4	5.59	(4.70, 6.00)	(4.80, 5.60)
P5	4.45	(3.60, 4.50)	(4.10, 4.50)
P6	5.15	(4.98, 5.35)	(4.90, 5.20)

Table 8. Thresholds determined by different threshold selection methods.

Return Periods	MF-DFA (m)	Mean Excess Function (m)	Parameter Stability Plot (m)
50 year	8.53	8.54	8.55
100 year	9.13	9.19	9.22
150 year	9.47	9.58	9.61
200 year	9.71	9.85	9.89

**Table 9.** Return periods estimated by different thresholds.

From the above process, we can observe that when using the mean excess plot and parameter stability plot methods for threshold selection, the stable parameters may not be distinct enough, requiring substantial expertise for interpretation. Often, multiple thresholds may satisfy the selection criteria, necessitating sensitivity analysis of the return period of significant wave height to determine the appropriate threshold. On the other hand, the MF-DFA method can provide a unique threshold. The threshold determined using the MF-DFA method for estimating return periods exhibits minimal differences compared to the traditional methods, highlighting the reliability of the MF-DFA method in determining extreme wave height thresholds.

#### 5. Conclusions

This study utilized long-term significant wave height data to investigate the importance of threshold selection in the POT method for estimating return periods in extreme value analysis. The study proposed using the MF-DFA method to determine the threshold for extreme wave heights and compared it with graphical diagnostics, highlighting the advantages and feasibility of the MF-DFA method in threshold determination. The following conclusions were drawn:

The MF-DFA method has demonstrated the existence of long-range correlations in the significant wave height series, serving as the foundation for threshold determination using the MF-DFA method. Extreme events and non-extreme events are typically driven by different physical processes, and long-range correlations are minimally affected or have a minor influence from extreme events. Based on this idea, extreme significant wave height thresholds were determined by observing changes in long-range correlations after removing extreme values from the significant wave height series. When the removal of values significantly affects long-range correlations, that value is considered the threshold for extreme effective wave heights.

This study calculated fractal indices for different extreme significant wave height series, identified points of abrupt change in the fractal indices, and determined that point as the threshold for extreme significant wave heights. The threshold determined by the MF-DFA method was compared with the mean residual life plot method and parameter stability plot method. The study analyzed the return period results under different threshold methods and, in conjunction with the stable threshold range, highlighted the rationality of the MF-DFA method in threshold selection.

In contrast to traditional graphical diagnostic techniques, the MF-DFA method provides a deterministic threshold. Furthermore, it considers the inherent long-range correlation characteristics of the data during the determination of extreme event thresholds, offering objectivity that is less influenced by the subjectivity of researchers. Thus, the MF-DFA method can be considered as a threshold selection method in wave extreme value analysis.

In summary, the MF-DFA method offers a reliable and objective approach to threshold selection for extreme wave height analysis. It considers the data's long-range correlation properties and provides a definite threshold, making it a valuable tool for wave extreme value analysis.

**Supplementary Materials:** The following supporting information can be downloaded at: https: //www.mdpi.com/article/10.3390/w15203648/s1, Figure S1: Determination of extreme wave height threshold using MF-DFA method; Figure S2: Mean residual life plots; Figure S3: Parameter stability plots; Table S1: Significance test of turning point.

**Author Contributions:** The concept of the study was developed by H.L. H.L.: performed numerical simulations, analysis, and manuscript writing; F.Y.: data curation, validation, review and editing; H.W.: methodology, funding acquisition. All authors have read and agreed to the published version of the manuscript.

**Funding:** The authors would like to thank the support of the National Key Research and Development Program of China (Grant No. 2021YFB2600701).

**Data Availability Statement:** The significant wave height which are presented in this article are provided in open access (http://climate.copernicus.eu/climate-reanalysis, accessed on 1 January 2021).

Conflicts of Interest: The authors declare no conflict of interest.

#### References

- 1. Cai, Y.; Reeve, D.E. Extreme value prediction via a quantile function model. Coast. Eng. 2013, 77, 91–98. [CrossRef]
- Sartini, L.; Besio, G.; Cassola, F. Spatio-temporal modelling of extreme wave heights in the Mediterranean Sea. Ocean Model. 2017, 117, 52–69. [CrossRef]
- 3. Gao, J.; Shi, H.; Zang, J.; Liu, Y. Mechanism analysis on the mitigation of harbor resonance by periodic undulating topography. Ocean Eng. 2023, 281, 114923. [CrossRef]
- Gao, J.-L.; Chen, H.-Z.; Mei, L.-L.; Liu, Z.; Liu, Q. Statistical analyses of wave height distribution for multidirectional irregular waves over a sloping bottom. *China Ocean Eng.* 2021, 35, 504–517. [CrossRef]
- Fisher, R.A.; Tippett, L.H.C. Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Math. Proc. Camb. Philos. Soc.* 1928, 24, 180–190. [CrossRef]
- Muraleedharan, G.; Lucas, C.; Soares, C.G. Regression quantile models for estimating trends in extreme significant wave heights. Ocean Eng. 2016, 118, 204–215. [CrossRef]
- 7. Pickands, J., III. Statistical inference using extreme order statistics. *Ann. Stat.* **1975**, *3*, 119–131. Available online: https://www.jstor.org/stable/2958083 (accessed on 1 January 2021).

- 8. Tawn, J.A. An extreme-value theory model for dependent observations. J. Hydrol. 1988, 101, 227–250. [CrossRef]
- 9. Weibull, W. A statistical theory of the strength of materials. *Proc. Royal Swedish Inst. Eng. Res.* **1939**, *151*, 1–45.
- 10. Balkema, A.A.; De Haan, L. Residual life time at great age. Ann. Probab. 1974, 2, 792–804. [CrossRef]
- 11. Soares, C.G.; Scotto, M. Application of the r largest-order statistics for long-term predictions of significant wave height. *Coast. Eng.* **2004**, *51*, 387–394. [CrossRef]
- 12. Goda, Y.; Konagaya, O.; Takeshita, N.; Hitomi, H.; Nagai, T. Population distribution of extreme wave heights estimated through regional analysis. In *Coastal Engineering 2000. Presented at the 27th International Conference on Coastal Engineering (ICCE), Copenhagen, Sydney, Australia, 16–21 July 2000;* American Society of Civil Engineers: Sydney, Australia, 2001; pp. 1078–1091. [CrossRef]
- 13. Davies, G.; Callaghan, D.P.; Gravois, U.; Jiang, W.; Hanslow, D.; Nichol, S.; Baldock, T. Improved treatment of non-stationary conditions and uncertainties in probabilistic models of storm wave climate. *Coast. Eng.* **2017**, *127*, 1–19. [CrossRef]
- 14. Mazas, F.; Hamm, L. A multi-distribution approach to POT methods for determining extreme wave heights. *Coast. Eng.* **2011**, *58*, 385–394. [CrossRef]
- 15. Shamji, V.; Aboobacker, V.; Vineesh, T. Extreme value analysis of wave climate around Farasan Islands, southern Red Sea. *Ocean Eng.* **2020**, 207, 107395. [CrossRef]
- Coles, S.; Bawa, J.; Trenner, L.; Dorazio, P. An Introduction to Statistical Modeling of Extreme Values; Springer: London, UK, 2001; Volume 208, Available online: https://www.springer.com/gp/book/9781852334598 (accessed on 1 January 2021).
- 17. Hill, B.M. A simple general approach to inference about the tail of a distribution. Ann. Stat. 1975, 3, 1163–1174. [CrossRef]
- 18. Beirlant, J.; Goegebeur, Y.; Segers, J.; Teugels, J.L. *Statistics of Extremes: Theory and Applications*; John Wiley and Sons Ltd.: Chichester, UK, 2004.
- 19. Carreau, J.; Bengio, Y. A hybrid Pareto model for asymmetric fat-tailed data: The univariate case. *Extremes* **2009**, *12*, 53–76. [CrossRef]
- 20. Eastoe, E.F.; Tawn, J.A. Statistical models for overdispersion in the frequency of peaks over threshold data for a flow series. *Water Resour. Res.* **2010**, *46*, 1–12. [CrossRef]
- 21. Solari, S.; Losada, M. A unified statistical model for hydrological variables including the selection of threshold for the peak over threshold method. *Water Resour. Res.* 2012, *48.* [CrossRef]
- 22. Chen, B.-Y.; Zhang, K.-Y.; Wang, L.-P.; Jiang, S.; Liu, G.-L. Generalized extreme value-pareto distribution function and its applications in ocean engineering. *China Ocean Eng.* 2019, 33, 127–136. [CrossRef]
- Niroomandi, A.; Ma, G.; Ye, X.; Lou, S.; Xue, P. Extreme value analysis of wave climate in Chesapeake Bay. Ocean Eng. 2018, 159, 22–36. [CrossRef]
- 24. Vanem, E. Uncertainties in extreme value modelling of wave data in a climate change perspective. J. Ocean Eng. Mar. Energy 2015, 1, 339–359. [CrossRef]
- Thompson, P.; Cai, Y.; Reeve, D.; Stander, J. Automated threshold selection methods for extreme wave analysis. *Coast. Eng.* 2009, 56, 1013–1021. [CrossRef]
- 26. Liang, B.; Shao, Z.; Li, H.; Shao, M.; Lee, D. An automated threshold selection method based on the characteristic of extrapolated significant wave heights. *Coast. Eng.* **2019**, *144*, 22–32. [CrossRef]
- 27. Shao, Z.; Liang, B.; Gao, H. Extracting independent and identically distributed samples from time series significant wave heights in the Yellow Sea. *Coast. Eng.* **2020**, *158*, 103693. [CrossRef]
- Scarrott, C.; MacDonald, A. A review of extreme value threshold estimation and uncertainty quantification. *REVSTAT-Stat. J.* 2012, 10, 33–60. [CrossRef]
- 29. Ozger, M. Scaling characteristics of ocean wave height time series. Phys. A Stat. Mech. Its Appl. 2011, 390, 981–989. [CrossRef]
- 30. Kantelhardt, J.W.; Zschiegner, S.A.; Koscielny-Bunde, E.; Havlin, S.; Bunde, A.; Stanley, H.E. Multifractal detrended fluctuation analysis of nonstationary time series. *Phys. A Stat. Mech. Its Appl.* **2002**, *316*, 87–114. [CrossRef]
- Du, H.; Wu, Z.; Zong, S.; Meng, X.; Wang, L. Assessing the characteristics of extreme precipitation over northeast China using the multifractal detrended fluctuation analysis. *J. Geophys. Res. Atmos.* 2013, 118, 6165–6174. [CrossRef]
- 32. Liu, B.; Chen, J.; Chen, X.; Lian, Y.; Wu, L. Uncertainty in determining extreme precipitation thresholds. *J. Hydrol.* **2013**, *503*, 233–245. [CrossRef]
- 33. Zhang, J.; Gao, G.; Fu, B.; Wang, C.; Gupta, H.V.; Zhang, X.; Li, R. A universal multifractal approach to assessment of spatiotemporal extreme precipitation over the Loess Plateau of China. *Hydrol. Earth Syst. Sci.* **2020**, *24*, 809–826. [CrossRef]
- Du, H.; Wu, Z.; Li, M.; Jin, Y.; Zong, S.; Meng, X. Characteristics of extreme daily minimum and maximum temperature over Northeast China, 1961–2009. *Theor. Appl. Climatol.* 2013, 111, 161–171. [CrossRef]
- 35. Naseef, T.M.; Kumar, V.S. Influence of tropical cyclones on the 100-year return period wave height—A study based on 39-year long ERA5 reanalysis data. *Int. J. Climatol.* 2020, 40, 2106–2116. [CrossRef]
- Sreelakshmi, S.; Bhaskaran, P.K. Wind-generated wave climate variability in the Indian Ocean using ERA-5 dataset. *Ocean Eng.* 2020, 209, 107486. [CrossRef]
- 37. Wang, J.; Liu, J.; Wang, Y.; Liao, Z.; Sun, P. Spatiotemporal variations and extreme value analysis of significant wave height in the South China Sea based on 71-year long ERA5 wave reanalysis. *Appl. Ocean Res.* **2021**, *113*, 102750. [CrossRef]

- 38. Méndez, F.J.; Menéndez, M.; Luceño, A.; Losada, I.J. Estimation of the long-term variability of extreme significant wave height using a time-dependent peak over threshold (pot) model. *J. Geophys. Res. Ocean.* **2006**, *111*, 1–13. [CrossRef]
- 39. Luceño, A.; Menéndez, M.; Méndez, F.J. The effect of temporal dependence on the estimation of the frequency of extreme ocean climate events. *Proc. R. Soc. A Math. Phys. Eng. Sci.* 2006, 462, 1683–1697. [CrossRef]

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.