Article

Comparison of Extreme Wind and Waves Using Different Statistical Methods in 40 Offshore Wind Energy Lease Areas Worldwide

Saravanan Bhaskaran [1], Amrit Shankar Verma [1,*], Andrew J. Goupee [1], Subhamoy Bhattacharya [2], Amir R. Nejad [3] and Wei Shi [4]

1   Department of Mechanical Engineering, University of Maine, Orono, ME 04473, USA; saravanan.bhaskaran@maine.edu (S.B.); agoupe91@maine.edu (A.J.G.)
2   Department of Civil and Environmental Engineering, University of Surrey, Guildford GU2 7XH, UK; s.bhattacharya@surrey.ac.uk
3   Department of Marine Technology, Norwegian University of Science and Technology, 7491 Trondheim, Norway
4   Deepwater Engineering Research Center, Dalian University of Technology, Dalian 116024, China; weishi@dlut.edu.cn
*   Correspondence: amrit.verma@maine.edu

Abstract: With the ongoing global drive towards renewable energy, several potential offshore wind energy lease areas worldwide have come into focus. This study aims to estimate the extreme wind and wave conditions across several newly designated offshore wind lease sites spanning six continents that are crucial for risk assessment and the design of offshore wind turbines. Firstly, the raw data of wind speeds and wave heights prevailing in these different lease areas were obtained. Following this, an in-depth extreme value analysis was performed over different return periods. Two principal methodologies were applied for this comparative study: the block-maxima and the peaks-over-threshold (POT) approaches. Various statistical techniques, including the Gumbel method of moments, Gumbel maximum likelihood, Gumbel least-squares, and the three-parameter GEV, were employed under the block-maxima approach to obtain the distribution parameters. The threshold for the POT approach was defined using the mean residual life method, and the distribution parameters were obtained using the maximum likelihood method. The Gumbel least-squares method emerged as the most conservative estimator of extreme values in the majority of cases, while the POT approach generally yielded lower extreme values compared to the block-maxima approach. However, the results from the POT approach showed large variations based on the selected threshold. This comprehensive study's findings will provide valuable input for the efficient planning, design, and construction of future offshore wind farms.

Keywords: offshore wind turbines; extreme value analysis; block maxima; peaks-over-threshold

1. Introduction

1.1. Background and Motivation

The offshore wind energy market is expanding rapidly in recent years, with a global installed capacity of 64.3 GW at the end of 2022 [1]. It is reported that 8.8 GW of offshore wind energy were newly added to the grid during the calendar year 2022 [1]. Figure 1 shows the offshore wind installed capacity targets set by various countries for 2030. It is notable that China's target of 200 GW of installed capacity by 2030 is far higher than all the other countries in the world. The USA has the highest potential for offshore wind energy in North America, with more than 26 GW worth of projects in the pipeline [2]. The International Renewable Energy Agency (IRENA) expects that that total installed capacity of offshore wind will be 2000 GW by 2050 [3].
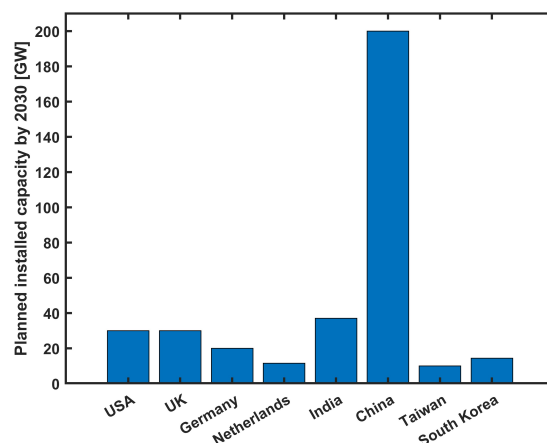
Figure 1. Offshore wind targets for different markets [1].

This means that at least 70 GW of offshore energy needs to be installed every year to achieve that target. This would require more than 5000 new turbines to be installed yearly, thereby occupying around 500,000 km^2 of space in the world's oceans [4]. As a result, new offshore wind lease areas are being designated in the territorial waters of countries worldwide to meet their respective targets. Furthermore, the size of wind turbines has been increasing steadily to maximize the use of the available wind energy potential. Large wind turbines can have a rotor diameter reaching up to 260 m and correspondingly large hub heights [4], thus posing new design challenges [5].

Offshore wind turbines (OWTs) typically have an operational lifetime of 20–25 years [6] and are subjected to extreme wind and waves during their service lives. As a result, when designing offshore wind turbines (OWTs), accurately estimating extreme wind and wave conditions is crucial to ensure that the OWT structure can survive these loads. The 50-year wind speed is one of the important environmental inputs, which is required during the selection of wind turbine classes as specified by the International Electrotechnical Commission (IEC) [7]. There are several methods in the literature that can be used to estimate sea states corresponding to different return periods. For instance, the traditional method of using reference wind speed (U_{ref}) in wind turbine standards, which is taken as five times the mean wind speed, tends to result in conservative designs, thus leading to larger and more expensive structures [8]. To mitigate this, more sophisticated statistical approaches such as block maxima and peak over thresholds can be used. However, these methods require long periods of metocean data, which may not be readily available for newly designated offshore lease areas. Another major issue is that estimating extreme values using these sophisticated methods is associated with high uncertainty. Some major causes include the following: (1) extrapolating wave heights and wind speeds to large return periods with limited durations of metocean data, (2) the chosen data set can have varying levels of accuracy, time duration, and inhomogeneities, (3) the probability distribution which is used to model the extreme value populations can vary, and (4) the method used to estimate the distribution parameters can be different. The sound understanding and judgment of risk are required to develop offshore wind energy in new locations around the world.

1.2. Literature Review

Several studies have been conducted on the estimation of extreme wind and waves. Barthelmie et al. [2] estimated the extreme wind and waves along the east coast of the USA using the ERA5 dataset. This study compared several different block-maxima methods, including the Gumbel graphical method, Gumbel–Weibull method, Gumbel method of moments (MOM), Gumbel maximum likelihood (ML) and three-parameter generalized extreme value distribution (GEVD) using the ML method. The study found that the 50-year wind speeds at 100 m above sea level were ranging between 30 ms^{-1} and 40 ms^{-1} along

the east coast of the US. The three-parameter GEVD provided the highest estimate among all of the methods used in this paper. However, these results cannot be extrapolated to longer return periods, since the distribution parameters were not provided. Pryor et al. [8] performed an assessment of extreme wind speeds specifically for applications related to wind energy. This study concentrated only on 50-year wind speeds obtained from the ERA5 dataset using four different methods. This paper provides a digital atlas of 50-year wind speeds at specific grid points around the world. The results for the 50-year wind speeds were within 9–13% of the buoy measurements. This paper is limited to the block-maxima approach only. Izaguirre et al. [9] presented the variability in global extreme significant wave height to spatial and temporal factors based on satellite data. This study used the GEV distribution with a block period of one month. Lee et al. [10] estimated the extreme wind speeds in fifteen wind farm locations in South Korea using the Gumbel and Weibull methods. This paper also compared the predicted values with numerical analysis and field measurement data for verification. Sacre [11] assessed the extreme wind speeds in France with a special focus on the 1999 storms. Torrielli et al. [12] focused exclusively on predicting the number of independent extreme wind events in a year, which is particularly useful in the block-maxima approach. Palutikof et al. [13] published a review of the methods available for extreme wind speed estimation. The paper also presented the strategies available for generating synthetic wind speed datasets. Wang et al. [14] proposed the R^2 and proper error (PE) criteria for evaluating the goodness of fit of various block-maxima methods for extreme wind estimation. Hong et al. [15] analyzed the performance of different fitting methods for the Gumbel distribution. This paper compared the 50-year wind speeds obtained at different meteorological stations in Canada. The Gumbel maximum likelihood, Gumbel method of moments, Gumbel method of L-moments, and Gumbel generalized least-squares methods were used for estimating the extreme wind speed. Based on efficiency, the maximum likelihood was rated as the best method, followed by the least-squares method. Lombardo [16] provided an improved method for estimating extreme wind speeds, which takes into account the spatial data resolution and micrometeorological characteristics. Afzal et al. [17] developed a novel method for the prediction of extreme significant wave height using a machine learning algorithm based on the GEV parameters.

Simiu et al. [18] developed a peaks-over-threshold (POT) approach for modeling extreme wind distribution tails, which has been widely used in subsequent studies. Viselli et al. [19] estimated the extreme wind and wave design parameters in the Gulf of Maine using a POT approach. An et al. [20] presented a comparison between four methods for the estimation of extreme wind speeds. Two of the methods used the block-maxima approach, while the other two methods used the POT approach. Kang et al. [21] conducted an extreme wind assessment on Jeju Island, South Korea. This study compared the block-maxima and POT approaches and the suitability of the Gumbel distribution regarding both the approaches. Rivas et al. [22] made a comparison between the block-maxima and POT approaches applied to the pitting corrosion of low carbon steel. This study found that the POT approach to be more suitable for pitting corrosion experiments provided that enough data are available. Vinoth and Young [23] provided global extreme wind speed and wave height data for a $1° \times 1°$ grid. This paper used the initial distribution and POT approaches for estimating the 100-year extreme values. The results were found to agree well with buoy measurements. Jonathan and Ewans [24] investigated the uncertainties associated with extreme wave height estimates due to hurricanes using the POT approach. Pandey et al. [25] estimated the extreme wind speeds using L-moments in the POT approach. This study found that the estimates of the shape parameter using the L-skewness of the values exceeding the threshold and provided a stable upper bound for the quantiles of the wind speed. Karpa and Naess [26] developed a novel average conditional exceedance rate (ACER) method for estimating extreme wind speeds. The main advantage of the ACER method is that it represents the exact extreme value distribution in a nonparametric form. The ACER method also does not require any initial declustering, as there is no requirement for independent data. Some nontraditional methods have also

been developed for extreme wind speed prediction, which do not involve the use of any statistical distribution. Gaidai et al., proposed an innovative method for extreme wind speed prediction using deconvolution [27]. The main advantage of this method is that it does not involve any extrapolation, which helps in reducing the associated errors. This paper also compared the predicted results with the Naess–Gaidai extrapolation method to determine the efficacy of the proposed method.

It is evident that plenty of work has been accomplished regarding the estimation of extreme wind and waves in general, but their applications in the context of offshore wind energy are quite limited. Since the offshore wind industry has recently emerged in several countries, newer studies are required to provide the relevant information. There is currently only one paper on the estimation of extreme wind and waves for new offshore wind lease areas, which is limited to the east coast of the US. This paper also does not consider the POT approach. Hence, there is a scope for more research to be performed on this topic.

1.3. Research Objectives and Novelties

The primary aim of this study is to estimate extreme wind speeds and wave heights using different statistical methods for several new offshore wind lease areas around the world as of August 2023. There are two traditional approaches for extreme value analysis, namely, the block-maxima approach and the POT approach. Both approaches have been used in this study for comparative analysis. Further, the extreme values were estimated using five different methods, namely, The Gumbel least-squares (LS), Gumbel ML, Gumbel MOM, GEVD, and the generalized Pareto distribution (GPD) methods. A sensitivity analysis has also been carried out to study the effect of different input parameters on the results from the POT approach. The final objective is to obtain the extreme wind speeds and wave heights for different return periods, as well as their corresponding distribution parameters for all the sites.

1.4. Novelties

This paper presents three novelties in the domain of OWT design and extreme value analysis: (a) First, it offers a comprehensive estimation of extreme wind and wave conditions tailored specifically to newly designated offshore wind lease areas. This approach serves as a vital tool in understanding and preparing for the unique challenges presented by these novel regions. This would be crucial for risk assessment and the design of offshore wind turbines at these new sites; (b) Second, the paper underscores the disparities among various statistical methods by conducting comparative studies. In addition, by utilizing a new Waveclimate dataset, it aims to substantially reduce the uncertainties that typically arise due to variations in datasets and analysis techniques. This integrated approach enhances the reliability and accuracy of the estimations; and (c) Last, to encourage transparency and foster further research in the field, the distribution parameters for the extreme values for the analyzed offshore lease areas are meticulously detailed. Such provision not only ensures the easy reproduction of our findings, but also facilitates the broader application and reuse of the results by the scientific community. It is expected that the extreme values provided in this paper will serve as crucial input during the design of OWTs for new offshore wind lease areas.

2. Methodology

2.1. Criteria for Choosing Offshore Wind Lease Areas

Table 1 shows the offshore wind lease areas selected for this study, along with their geographical coordinates and the water depths prevalent in those locations. Firstly, the offshore wind lease areas which are in the concept or early planning stage were selected from different continents around the world. There are some sites where the consent application has been authorized and construction is about to begin. All the names of the wind lease areas and their corresponding geographical coordinates were obtained from the global offshore wind map provided by the marine consultancy company 4C Offshore [28],

which is located in the United Kingdom. It can be noted that there is only one site chosen from the African continent due to a lack of information on sites in this region. There is also a scarcity of information on sites in the Middle East and South Asia. There are a total of 9 sites where the average water depth is greater than 50 m where floating wind turbines are to be installed. This indicates that the majority of the planned offshore wind farms in the near future are still in shallow water depths under 50 m, where bottom-fixed turbines are to be installed.

Table 1. Offshore wind lease areas selected for this study [28,29].

Lease Area #	Name	Country	GPS Coordinates	Water Depth (m)
1	Maine research array	USA	43°23′ N, 69°21′ W	≈175
2	Revolution wind	USA	41°8′ N,71°4′ W	≈35
3	Ocean wind	USA	39°6′ N, 74°17′ W	≈37.5
4	Garden State offshore energy	USA	38°40′ N, 74°42′ W	≈20
5	Empire wind	USA	40°17′ N, 73°19′ W	21.9–41.14
6	OCS-A 0545	USA	33°27′ N, 77°58′ W	≈26
7	CVOW Commercial Project	USA	36°54′ N, 75°20′ W	21.9–38.1
8	Cascadia wind	USA	46°46′ N, 124°39′ W	≈150
9	Morro Bay E	USA	35°31′ N, 121°41′ W	≈150
10	Allan array	Canada	51°37′ N, 128°43′ W	≈35
11	Sea-Breeze Tech	Canada	46°2′ N, 61°49′ W	≈50
12	UY01	Uruguay	34°14′ S, 51°40′ W	≈50
13	Projeto Acu	Brazil	22°8′ S, 40°44′ W	≈50
14	Farol wind	Brazil	28°51′ S, 48°41′ W	≈50
15	Sopros do RJ	Brazil	21°37′ S, 40°25′ W	≈27
16	Projeto Ubu	Brazil	20°51′ S, 40°23′ W	≈27
17	Voyage	Ireland	51°21′ N, 7°21′ W	≈85
18	Inch Cape	United Kingdom	56°29′ N, 2°11′ W	≈25
19	Nordlicht I	Germany	54°17′ N, 6°13′ E	≈35
20	Baltic offshore alpha	Sweden	58°17′ N, 18°21′ E	≈36
21	Bornholm bassin syd	Denmark	54°50′ N, 15°34′ E	≈57
22	Vigso bay	Denmark	57°10′ N, 8°39′ E	≈14
23	Calabria	Italy	38°26′ N, 16°52′ E	≈475
24	Normandie	France	49°52′ N, 0°49′ W	≈45
25	GoliatVIND	Norway	71°49′ N, 22°34′ E	300–400
26	Nao Victoria	Spain	36°17′ N, 4°43′ W	≈300
27	Genesis Hexicon	South Africa	30°2′ S, 31°38′ E	≈500
28	E3	India	7°50′ N, 77°49′ E	≈50
29	Miaoli	Taiwan	24°39′ N, 120°38′ E	≈50
30	Huaneng Hainan Wenchang 1	China	19°58′ N, 111°3′ E	≈120
31	Huaneng Daishan I	China	30°18′ N, 121°42′ E	≈10
32	Minyang Jieyang Qianzhan III	China	22°38′ N, 116°27′ E	≈40
33	Boryeong	South Korea	36°14′ N, 126°4′ E	≈6
34	Satsuma	Japan	31°49′ N, 130°8′ E	≈40
35	Southern Mindoro	Philippines	11°52′ N, 121°28′ E	≈26
36	Leeuwin	Australia	33°1′ S, 115°17′ E	≈40
37	Mid West	Australia	29°32′ S, 114°35′ E	≈50
38	Southern winds	Australia	38°9′ S, 140°47′ E	≈35
39	Barwon	Australia	38°44′ S, 142°18′ E	≈78
40	South Taranaki	New Zealand	39°32′ S, 173°40′ E	≈36

2.2. Source of Raw Data

The raw time series of wind and wave data were then obtained for each location using WaveClimate infoplaza [30]. Waveclimate infoplaza is a third party web-based portal developed to provide offshore operators with quick and easy access to wind and wave statistics data. The portal provides instant access to wind and wave statistics and is designed

to support exploration studies and to assess operability for marine sites globally. The portal offers point-oriented data based on the wave model WaveWatchIII for offshore and a wave-ray models (SWRT) for nearshore translations [31]. The model has data calibrated with satellites, and the satellites have been calibrated with buoy measurements. The quality of the data has been validated and compared in previous studies [31]. The raw data were obtained for an area of 50 km × 50 km. The Waveclimate data set has a latitude/longitude grid with a resolution of $30' \times 30'$. The latitude and longitude coordinates were rounded off to $\pm 0.5°$ ($30'$) by the website, since the raw data are not available at all coordinates. The raw data contain the mean wind speed (U_W) and significant wave height (H_S) data, among others, for every 3 h starting from 1 January 1992 to 31 December 2022. The wind data contain the one-hour mean wind speed (U_W) at 10 m above sea level. The significant wave height (U_W) data were obtained from the average of the highest one-third of the measured wave heights within a three-hour period.

2.3. Block-Maxima Approach

Figure 2 shows the methodology used in the block-maxima approach. The block-maxima approach consists of dividing the raw data into block periods, which are nonoverlapping periods of equal size, and the maximum observation in each block period is recorded. The block period was taken as one year in this study. Figure 3a,b show a sample of H_S and U_W data recorded, respectively, for 1 year, along with the selected annual maximum values. It can be noted that the maximum U_W and H_S occurred at around the same time during the year. The annual maximum U_W and H_s were extracted for each calendar year from the raw data and were fitted to the GEVD and Gumbel distributions to obtain the corresponding distribution parameters.
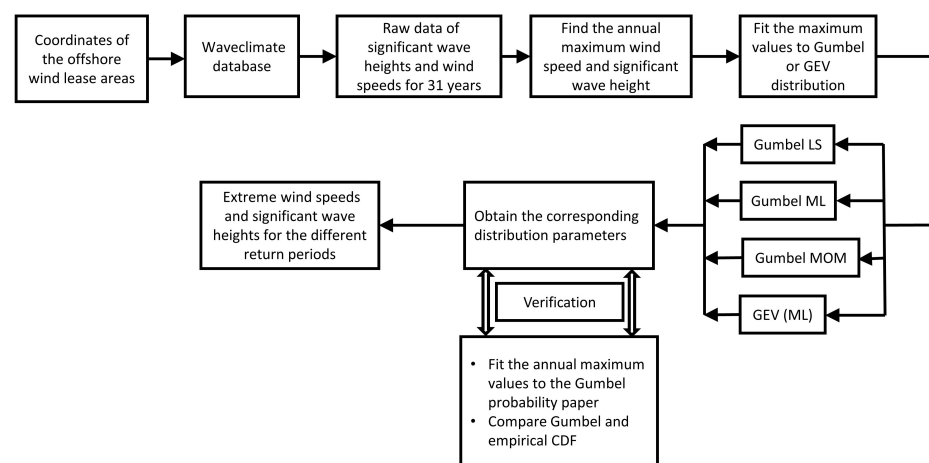

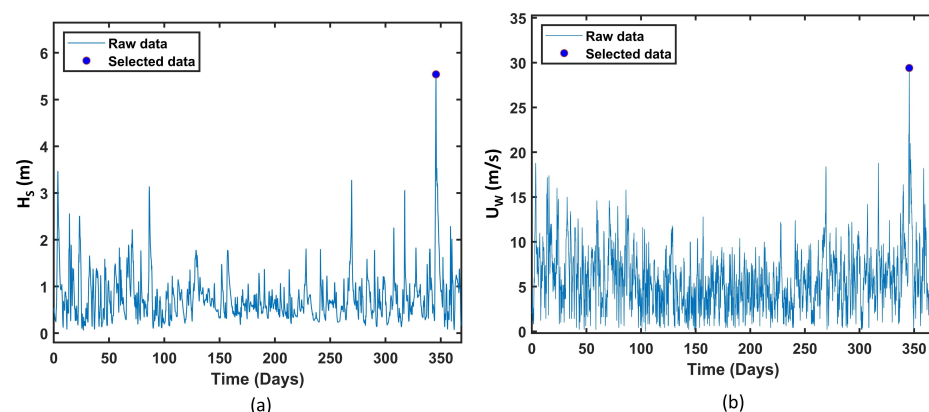**Figure 2.** Methodology used in the block-maxima approach.

Figure 3. Sample annual maximum selection for the block-maxima approach: (**a**) H_S. (**b**) U_W.

The cumulative distribution function (CDF) of the GEVD is given by the following [32]:

$$F(x) = e^{-(1-k\frac{x-\mu}{\sigma})^{1/k}} \tag{1}$$

where k is the shape parameter, μ is the location parameter, and σ is the scale parameter. The GEVD parameters are estimated using the ML method. The Gumbel distribution is also known as the type I GEVD. The Gumbel distribution is a special case of the GEVD, with the shape parameter being zero. The CDF of the Gumbel distribution is given the following by [33]:

$$F(x) = e^{-e^{-(x-\mu)/\beta}} \tag{2}$$

where μ is the location parameter, and β is the scale parameter. Firstly, the Gumbel parameters are estimated using the least-squares (LS) method. The least-squares method is a mathematical regression technique in which the sum of the squares of the difference between the observed values and the fitted values from the Gumbel distribution are minimized. The second method which is used to estimate the Gumbel parameters is the ML method, in which the joint probability density for obtaining the actual data points from the Gumbel distribution is maximized. The Gumbel LS, Gumbel ML, and GEVD (ML) methods were implemented using the WAFO toolbox [34] in MATLAB. The Gumbel parameters were also estimated using the empirical relations obtained by the MOM. This method involves equating sample moments with theoretical moments. In this method, the Gumbel parameters are estimated using the mean ($\bar{U}$) and standard deviation (σ) of the annual maximum U_W using the following relations [35]:

$$\beta = \frac{\sqrt{6}}{\pi}\sigma \tag{3}$$

$$\mu = \bar{U} - 0.577\beta \tag{4}$$

where μ is the location parameter, and β is the scale parameter. Once the Gumbel parameters are obtained, the extreme values of U_W and wave heights are obtained using the inverse Gumbel CDF, which is given by the following formula [33]:

$$x_q = \mu + \beta(-ln(-ln(1-p))) \tag{5}$$

where x_q is the required observation, and p is the probability of exceedance. Similarly, the extreme values are also obtained by using the inverse of the GEVD CDF, which is given by [32]

$$x_q = \mu - \frac{\sigma}{k}[(\ln p)^k - 1] \tag{6}$$

where x_q is the required observation, and p is the probability of exceedance. A verification study is also performed using probability paper and empirical CDF, which are detailed in Section 2.5.

2.4. POT Approach

Figure 4 shows the methodology used in the POT approach. The POT approach uses a threshold to seclude values considered extreme to the rest of the data and creates a model for the extreme values by modeling the tail of all the values that exceed this threshold. The optimal threshold is found using the mean residual life method in this study. In this approach, the mean of the excesses over a range of a range of thresholds is plotted as shown in Figure 5. The red lines in the figure show the 95% confidence interval values. The confidence interval denotes the probability that the mean excess will fall within the specified lower and upper bounds. The optimal threshold is chosen as the maximum of the region where the mean excess varies linearly with the change in the threshold.
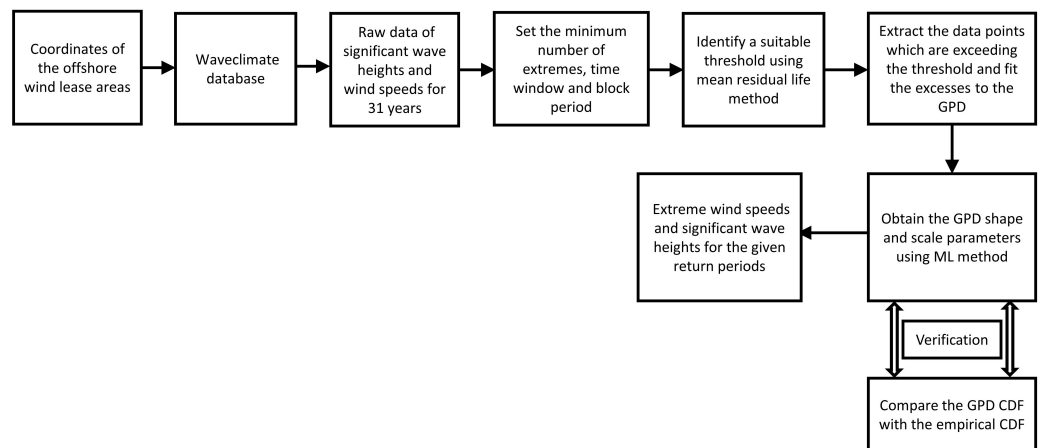
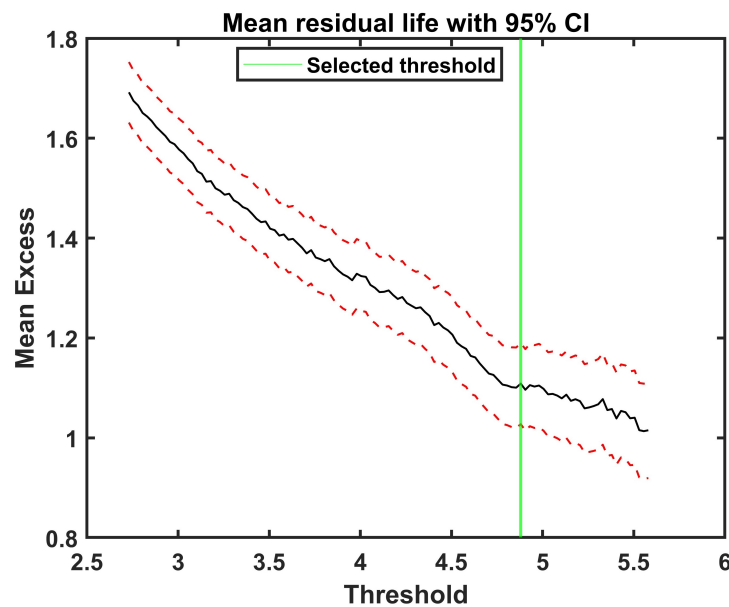**Figure 4.** Methodology used in the peaks over threshold approach.

Figure 5. Sample H_S threshold selection using the mean residual life approach.

Once the threshold is obtained, the U_W and H_S exceeding values are located along with the corresponding times at which they occurred. Figure 6a,b shows a sample of H_S and U_W data recorded for 1 year, along with the selected values. The zoomed-in view shows the selected values from the first 30 days. The circled area shows two peaks exceeding the threshold within a time window of 4 days. In such cases, only the highest peak within the time window is chosen, and other peaks are neglected. This declustering process is carried out in order to avoid bias in the selected data points [36]. The exceeding values are then fit to the generalized Pareto distribution (GPD). The CDF of the GPD is given by [37]

$$F(x) = 1 - (1 - kx/\sigma)^{1/k} \tag{7}$$

where k is the shape parameter, and σ is the scale parameter. The ML method was used for the GPD parameter estimation. The extreme values are then found by using the inverse of GPD CDF, which is given by [37]

$$x_q = \frac{-\sigma}{k}[(1-p)^k - 1] \tag{8}$$

where x_q is the required observation, and p is the probability of exceedance.
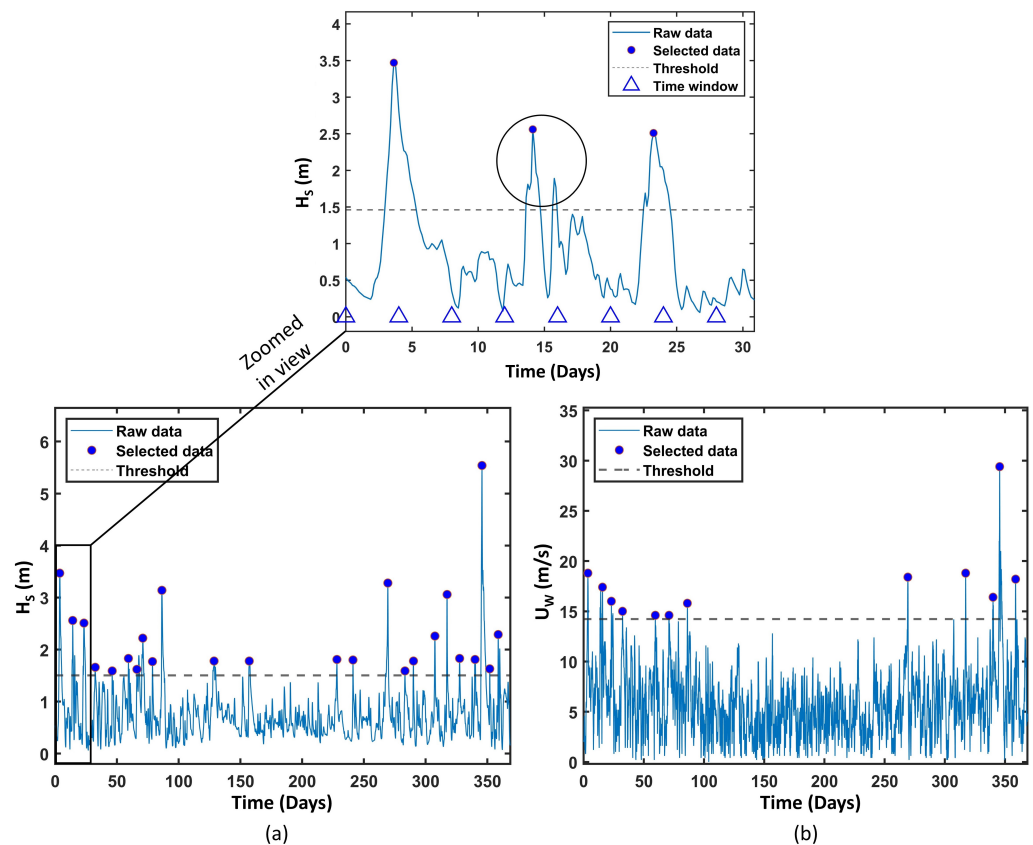
Figure 6. Sample annual maximum selection for the peaks over threshold approach: (**a**) H_S. (**b**) U_W.

2.5. Verification

The verification study for the block-maxima approach is aimed at ensuring that the yearly maximum values are true to the Gumbel distribution. The goodness of fit of the data to the Gumbel distribution was verified by performing a Pearson's χ^2 test. The χ^2 test statistic is given by the following equation [38]:

$$\chi^2 = \sum_{i=1}^{N} \frac{(O_i - E_i)^2}{E_i} \tag{9}$$

where N is the number of samples, O_i is the observed value, and E_i is the expected value. The χ^2 value is compared with the $\chi^2_{critical}$ value, which is obtained from the standard χ^2 curve. The χ^2 curve is dependent on the significance level, which is assumed to 0.05 and the number of degrees of freedom, which are determined using the expression $n - 1 - m$. The number of bins is denoted by n, while m refers to the number of distribution parameters, which is 2 for the Gumbel distribution. The null hypothesis that the observed data follows the Gumbel distribution cannot be rejected when χ^2 is less than $\chi^2_{critical}$. A visual comparison of the fitted Gumbel CDF with an empirical CDF was also performed as an additional check for the goodness of fit. The empirical CDF is defined as [39]

$$F_n(x) = \frac{\#(x_j <= x)}{n} \tag{10}$$

where x_n is a random sample, which is drawn from a distribution with a CDF of F.

Similarly, the U_W and H_S excess values were fit to the GPD distribution, and the goodness of fit to the GPD was verified using the Anderson–Darling test. The Anderson–Darling test statistic is obtained using the folloeing expression [40]:

$$A_n^2 = n \int_{-\infty}^{\infty} \frac{[F(x) - F_n(x)]^2}{F(x)[1 - F(x)]} dF(x) \tag{11}$$

where n is the number of samples, F is the CDF of hypothesized GPD, and F_n is the empirical CDF. This test measures the distance between F and F_n while giving a higher weight to the tail of the distribution [41]. A smaller value of the statistic means that the data follows the GPD closely. In this study, the significance level was chosen as 0.05, and the p value must be less than this value to fulfil the criteria for goodness of fit. The fitted GPD CDF was also compared with the empirical CDF for additional verification.

2.6. Assumptions and Limitations

The Waveclimate dataset contains spectral smoothing and variance underestimation, which are a characteristic features of numerical models for wind and wave data [31]. The Waveclimate dataset does not account for hurricanes or tropical storms, which are known to occur at some of the sites selected in this study. These effects must be accounted for separately using storm tracks and intensity. The 1-hour mean wind speed was used in this study instead of the IEC standard of the 10-minute mean wind speed, which leads to an underestimation of the maximum wind speed [2]. The extreme wind speeds provided in this study were at a reference height of 10 m above the sea level. Therefore, appropriate wind shear must be taken into account in order to find the wind speeds at the hub heights of OWTs. This study does not consider the statistical implications of comparison between a two-parameter distribution and a three-parameter distribution. The relative error between the different methods used in this study has not been considered. The goodness-of-fit tests used in study do not necessarily confirm to the selected data following the hypothesized distribution. They simply imply that the hypothesis cannot be rejected at the chosen significance level. The suitability of the methods used and associated error involved must be evaluated separately for each site. In addition, the extreme values provided in this study have not been validated by buoy measurements. The methods used in this study are not exhaustive. Other distributions, as well as parameter estimation methods, are available for the estimation of extreme values.

3. Results and Discussion

The results section consists of the following parts: (1) the block-maxima approach, (2) the POT approach, (3) verification, and (4) the comparison of extreme values. Firstly, the annual maximum values are presented, which are then fitted to the GEV and Gumbel distributions. Next, the number of data points available for different thresholds is shown for the POT approach. The sensitivity of the extreme U_W and H_S to the chosen threshold is then demonstrated. In the verification study, the GPD CDF and Gumbel CDF are compared to the empirical CDF. Finally, the comparison of the extreme values obtained from both approaches is discussed for several new offshore lease areas.

3.1. Block-Maxima Approach

The annual maximum U_W and H_S were extracted from the raw data in order to be used in the block-maxima approach. Figure 7 shows the annual maximum U_W at 10 m above the sea level for different lease areas around the world.
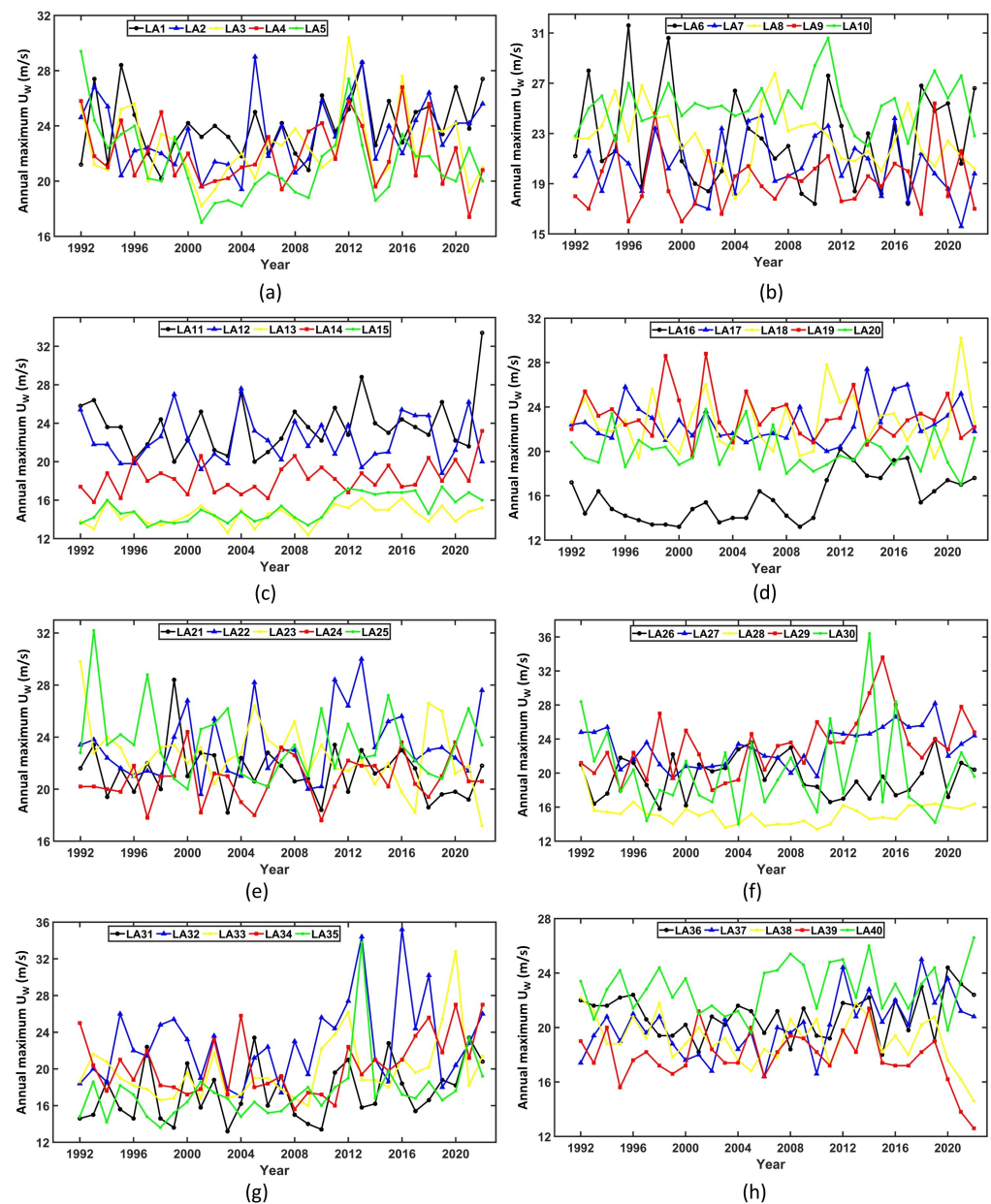
Figure 7. Annual maximum U_W at 10 m above the sea level in different lease areas around the world: (**a**) LA1–5. (**b**) LA6–10. (**c**) LA11–15. (**d**) LA16–20. (**e**) LA21–25. (**f**) LA26–30. (**g**) LA31–35. (**h**) LA36–40.

It is worth remembering that that the maximum U_W will be higher at the hub heights, and appropriate wind shear must be considered. It is noted that for LA10 (Allan array, Canada), the research has recorded the highest mean annual maximum U_W of 25.2 m/s. However, the outright highest U_W recorded between 1992 and 2022 is 36.4 m/s, which occurred in LA30 (Huaneng Hainan Wenchang I, China) in the year 2014. LA3 (Ocean wind, USA) has recorded the highest wind speed off the East Coast of USA, with 30.4 m/s in the year 2012. U_W values greater than 33 m/s are classified as hurricanes [42]. From Figure 7c (South America) and Figure 7h (Australia), it is visible that there was only one hurricane level wind speed recorded in the Southern hemisphere. Hurricanes are much more common in the Northern hemisphere than the Southern hemisphere due to low wind shear and warmer oceans [43]. Figure 8 shows the annual maximum H_S values for different lease areas around the world.
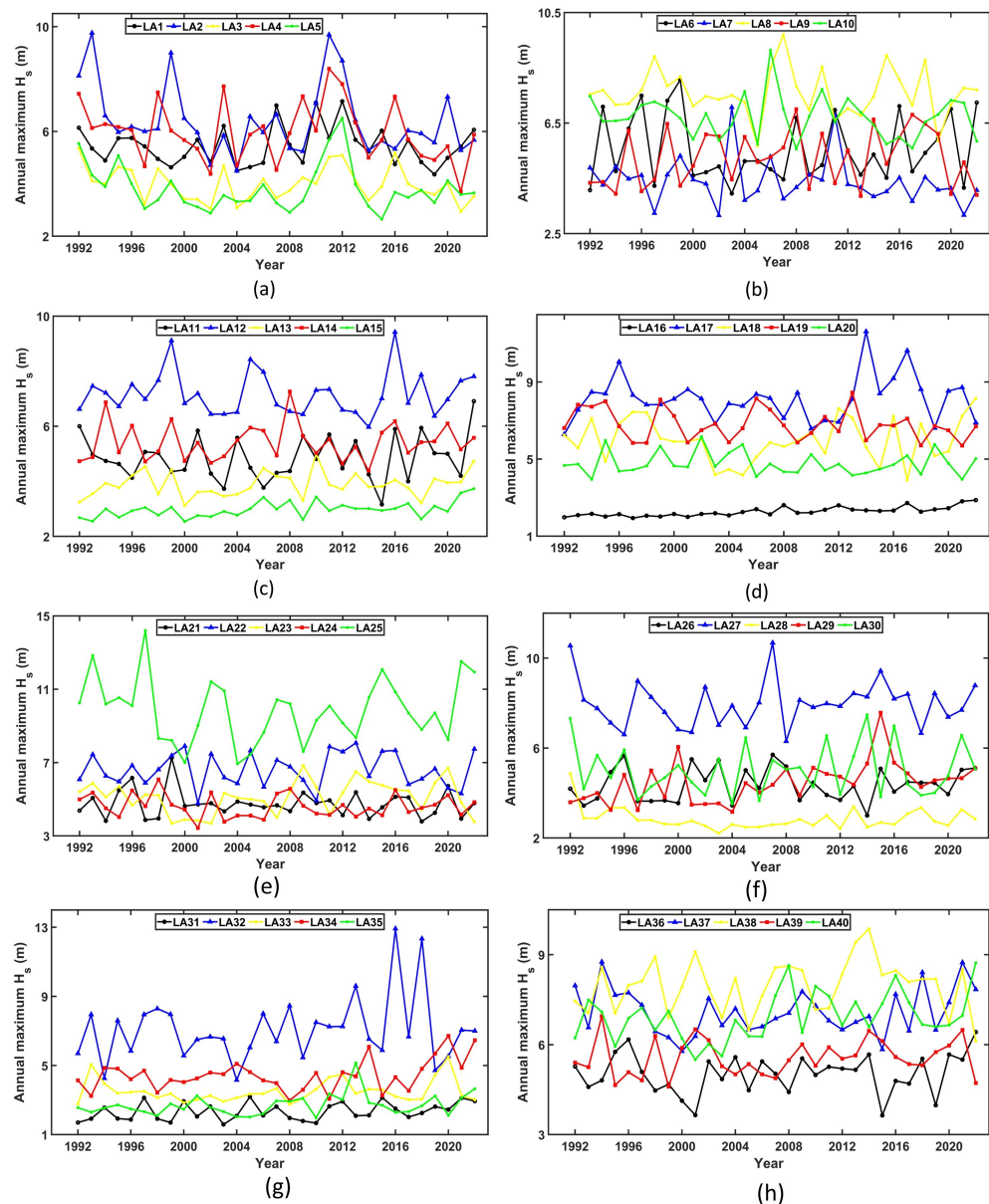
Figure 8. Annual maximum H_S in different lease areas around the world: (**a**) LA1–5. (**b**) LA6–10. (**c**) LA11–15. (**d**) LA16–20. (**e**) LA21–25. (**f**) LA26–30. (**g**) LA31–35. (**h**) LA36–40.

The highest mean annual maximum H_S of 9.85 m and the outright highest H_S of 14.2 m were recorded in LA25 (GoliatVIND, Norway). The North Sea off the coast of Norway is known for its large waves and rough weather in general [44]. LA16 (Projeto Ubu, Brazil) had the least mean annual maximum H_S of 2.27 m.

The annual maximum U_W and H_S were fitted to the Gumbel distribution for the block-maxima approach. The χ^2 test was performed for all of the annual maximum values selected from each site. The value of χ^2 was found to be less than the $\chi^2_{critical}$ in all the cases. Hence, the null hypothesis that the annual maximum values follow the Gumbel distribution cannot be rejected at the chosen significance level of 5%. Figure 9a,b show the comparison between the Gumbel and empirical CDF for the LA1 U_W and H_S, respectively. It is visible that both the annual maximum U_W and H_S followed the Gumbel distribution. Figure 10a,b show the annual maximum U_W and H_S, respectively, fitted to the Gumbel probability paper. The data aligns closely with the straight line, thus indicating a good fit with the Gumbel distribution. The Gumbel distribution parameters were then obtained for each lease area for both the U_W and H_S, which are shown in Table 2 and Table 3, respectively.
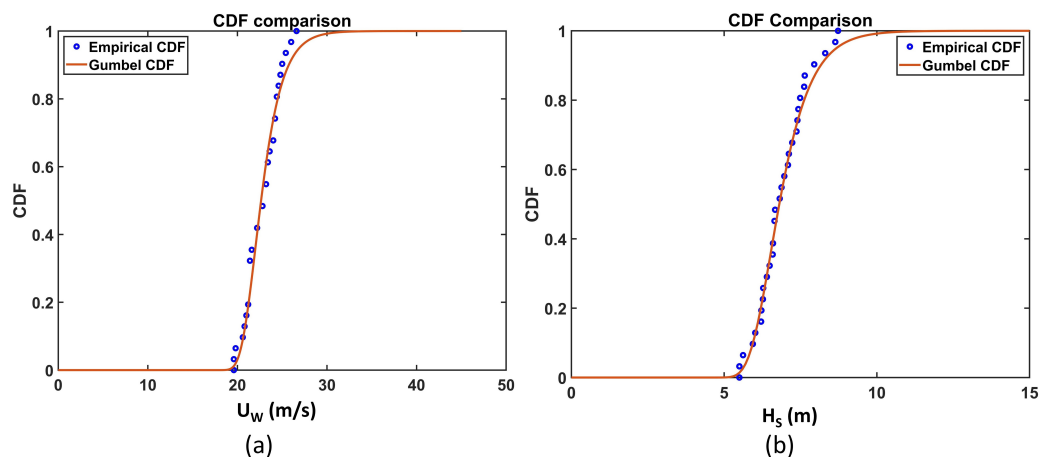
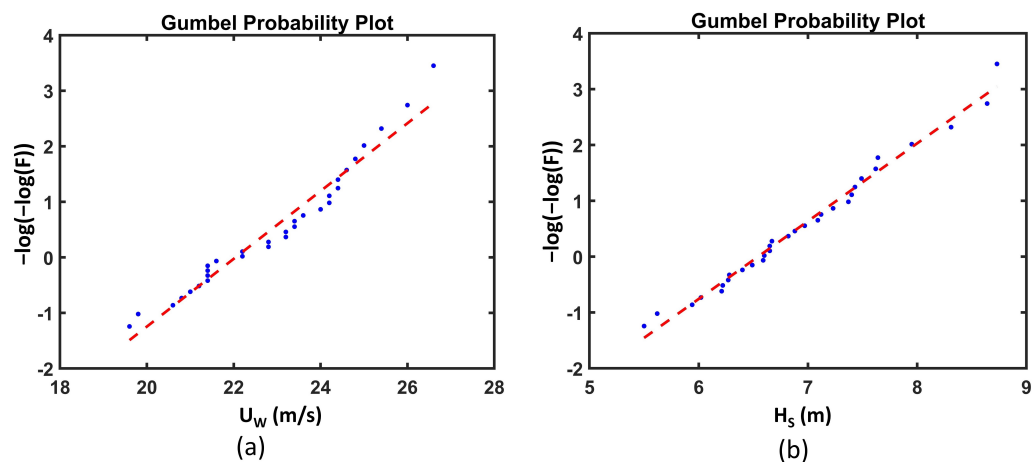**Figure 9.** Comparison of Gumbel and empirical CDF for LA1: (**a**) U_W. (**b**) H_S.

Figure 10. Gumbel probability paper for LA1: (**a**) U_W. (**b**) H_S.

Table 2. Parameters obtained for U_W for each lease area.

Lease Area #	Gumbel LS		Gumbel ML		Gumbel MOM		GEVD			POT	
	μ	β	μ	β	μ	β	k	σ	μ	k	σ
1	22.97	1.98	22.98	1.89	23.03	1.74	−0.16	1.98	23.14	0.21	3.31
2	22.24	2.19	22.27	2.04	22.31	1.92	−0.11	2.12	22.39	0.15	3.13
3	21.26	2.32	21.34	2.03	21.34	2.03	−0.02	2.04	21.36	0.05	2.73
4	20.90	2.06	20.93	1.96	20.97	1.79	−0.15	2.02	21.08	0.13	2.81
5	20.01	2.40	20.12	2.02	20.10	2.09	0.02	2.01	20.10	0.12	2.90
6	20.82	3.42	20.87	3.06	20.92	3.00	−0.02	3.08	20.90	0.02	2.70
7	19.28	2.11	19.28	2.14	19.37	1.82	−0.36	2.33	19.69	0.16	2.99
8	21.49	2.03	21.50	2.05	21.56	1.78	−0.19	2.10	21.71	0.23	3.62
9	18.12	2.06	18.19	1.77	18.19	1.81	0.03	1.74	18.16	0.001	1.73
10	24.28	1.75	24.30	1.69	24.34	1.52	−0.14	0.65	24.43	0.22	3.11
11	22.34	2.55	22.48	2.07	22.43	2.21	0.06	2.02	22.42	0.12	3.12
12	21.16	2.15	21.20	1.90	21.23	1.89	0.020	1.88	21.18	0.14	2.94
13	13.97	0.92	13.96	0.98	14.01	0.78	−0.41	1.04	14.18	0.17	1.15
14	17.55	1.47	17.60	1.29	17.60	1.28	−0.02	1.30	17.61	0.12	1.62
15	14.46	1.23	14.49	1.08	14.52	1.05	−0.05	1.11	14.52	0.24	1.50
16	14.83	1.86	14.86	1.63	14.89	1.61	0.08	1.56	14.79	0.14	1.68
17	21.74	1.62	21.81	1.30	21.79	1.41	0.12	1.24	21.73	0.20	2.70
18	21.48	2.29	21.55	1.96	21.55	2.01	0.06	1.90	21.49	0.15	3.20
19	22.13	1.90	22.21	1.63	22.20	1.66	−0.005	1.64	22.22	0.16	2.90
20	19.18	1.54	19.23	1.37	19.24	1.35	−0.03	1.38	19.25	0.20	2.62

Table 2. *Cont.*

Lease Area #	Gumbel LS		Gumbel ML		Gumbel MOM		GEVD			POT	
	μ	β	μ	β	μ	β	k	σ	μ	k	σ
21	20.33	1.84	20.43	1.59	20.42	1.56	−0.04	1.61	20.47	0.17	2.90
22	22.18	2.40	22.25	2.03	22.25	2.10	0.08	1.97	22.17	0.09	2.76
23	21.42	2.24	21.46	2.30	21.51	1.93	−0.16	2.29	21.66	0.20	3.89
24	19.93	1.52	19.92	1.57	19.99	1.32	−0.26	1.63	20.15	0.25	2.67
25	22.20	2.42	22.33	1.93	22.28	2.10	0.12	1.83	22.20	0.14	2.92
26	18.46	2.10	18.46	2.03	18.54	1.83	−0.29	2.25	18.78	0.07	2.49
27	21.98	2.05	21.98	1.97	22.06	1.76	−0.23	2.12	22.24	0.19	2.45
28	14.54	1.38	14.70	0.98	14.64	1.12	0.08	0.95	14.66	−0.06	0.84
29	21.54	3.16	21.61	2.86	21.64	2.77	−0.05	2.91	21.69	0.11	2.87
30	17.83	4.44	18.10	3.46	17.99	3.85	0.15	3.23	17.83	0.03	3.04
31	15.98	2.78	16.03	2.43	16.08	2.42	0.05	2.37	15.97	−0.01	1.95
32	20.68	4.08	20.87	3.32	20.82	3.55	0.12	3.15	20.66	0.02	2.57
33	18.17	3.22	18.51	2.16	18.33	2.72	0.28	1.86	18.20	0.02	2.08
34	18.89	2.89	18.95	2.53	18.99	2.53	0.02	2.51	18.92	0.04	2.65
35	15.62	3.66	16.33	1.97	15.97	2.80	0.16	1.82	16.15	0.03	2.16
36	20.13	1.44	20.12	1.47	20.19	1.23	−0.30	1.56	20.36	0.33	3.38
37	19.11	1.90	19.11	1.90	19.18	1.65	−0.20	1.98	19.32	0.16	2.09
38	18.25	1.62	18.23	1.83	18.33	1.37	−0.44	1.84	18.64	0.28	2.42
39	16.89	1.75	16.87	2.11	17.00	1.45	−0.40	1.95	17.30	0.25	2.36
40	22.04	1.64	22.03	1.64	22.10	1.42	−0.30	1.78	22.31	0.16	2.36

Table 3. Parameters obtained for H_S for each lease area.

Lease Area #	Gumbel LS		Gumbel ML		Gumbel MOM		GEVD			POT	
	μ	β	μ	β	μ	β	k	σ	μ	k	σ
1	5.09	0.67	5.11	0.57	5.11	0.58	0.05	0.56	5.10	0.13	1.22
2	5.76	1.22	5.84	0.94	5.81	1.05	0.11	0.90	5.79	0.09	1.19
3	3.68	0.58	3.68	0.54	3.70	0.51	−0.12	0.56	3.72	0.08	0.74
4	5.43	1.01	5.44	1.01	5.47	0.87	−0.22	1.05	5.56	0.07	1.13
5	3.39	0.79	3.45	0.58	3.43	0.68	0.16	0.54	3.41	0.05	0.71
6	4.96	1.10	4.99	0.91	5.00	0.94	0.20	0.82	4.90	0.02	0.89
7	3.98	0.79	4.04	0.63	4.02	0.67	0.09	0.63	4.04	0.07	0.80
8	7.18	0.78	7.17	0.83	7.20	0.67	−0.24	0.83	7.28	0.22	1.48
9	4.74	0.92	4.75	0.85	4.79	0.78	−0.36	1.00	4.93	0.03	0.83
10	6.35	0.69	6.37	0.64	6.38	0.60	−0.08	0.65	6.40	0.18	1.21
11	4.45	0.73	4.46	0.73	4.48	0.64	−0.17	0.75	4.52	0.10	0.88
12	6.80	0.71	6.83	0.57	6.82	0.62	0.08	0.56	6.81	0.15	1.21
13	3.68	0.40	3.68	0.38	3.69	0.35	−0.14	0.40	3.71	0.11	0.47
14	5.10	0.59	5.12	0.51	5.12	0.52	0.02	0.51	5.11	0.13	0.79
15	2.84	0.26	2.85	0.24	2.85	0.23	−0.05	0.24	2.85	0.13	0.30
16	2.16	0.22	2.17	0.18	2.17	0.19	0.09	0.17	2.16	0.15	0.27
17	7.56	1.06	7.60	0.93	7.60	0.91	−0.04	0.94	7.62	0.17	1.71
18	5.43	1.03	5.42	1.04	5.47	0.88	−0.36	1.13	5.63	0.06	1.13
19	6.40	0.71	6.41	0.64	6.42	0.62	−0.07	0.66	6.43	0.23	1.33
20	4.49	0.53	4.50	0.44	4.50	0.46	0.07	0.43	4.49	0.15	0.88
21	4.40	0.67	4.43	0.55	4.42	0.58	0.05	0.54	4.42	0.12	0.87
22	6.15	0.87	6.15	0.88	6.20	0.73	−0.59	1.04	6.43	0.13	1.31
23	4.69	0.80	4.69	0.82	4.73	0.68	−0.32	0.87	4.84	0.09	1.00
24	4.32	0.56	4.33	0.53	4.34	0.48	−0.15	0.55	4.38	0.20	0.90
25	9.02	1.56	9.03	1.51	9.07	1.37	−0.16	1.58	9.16	0.10	1.82
26	4.00	0.67	4.00	0.64	4.03	0.58	−0.23	0.69	4.09	0.10	0.90
27	7.49	0.93	7.51	0.84	7.52	0.80	−0.07	0.86	7.54	0.18	1.41
28	2.60	0.47	2.66	0.29	2.63	0.38	0.14	0.27	2.64	0.06	0.36
29	4.01	0.84	4.05	0.71	4.05	0.72	−0.003	0.72	4.05	0.09	0.72
30	4.49	1.01	4.52	0.87	4.53	0.89	0.11	0.83	4.47	−0.001	0.70

Lease Area #	Gumbel LS		Gumbel ML		Gumbel MOM		GEVD			POT	
	μ	β	μ	β	μ	β	k	σ	μ	k	σ
31	2.08	0.45	2.09	0.41	2.10	0.39	−0.09	0.42	2.11	0.01	0.32
32	6.14	1.77	6.25	1.44	6.21	1.51	0.01	1.44	6.24	−0.04	0.86
33	3.19	0.59	3.24	0.42	3.21	0.51	0.24	0.37	3.19	0.07	0.55
34	4.00	0.82	4.02	0.75	4.03	0.71	−0.09	0.77	4.06	0.06	0.79
35	2.40	0.58	2.46	0.42	2.43	0.49	0.15	0.39	2.42	0.07	0.55
36	4.70	0.61	4.69	0.67	4.73	0.52	−0.35	0.68	4.82	0.22	1.05
37	6.73	0.69	6.74	0.65	6.75	0.60	−0.12	0.67	6.78	0.14	1.12
38	7.52	0.80	7.51	0.84	7.56	0.68	−0.31	0.88	7.66	0.20	1.34
39	5.28	0.54	5.28	0.52	5.30	0.47	−0.17	0.54	5.33	0.19	0.86
40	6.54	0.72	6.55	0.69	6.57	0.63	−0.14	0.71	6.60	0.17	1.19

3.2. POT Approach

The peaks over the threshold approach involve various parameters such as the minimum number of extremes, the minimum time window between the extremes, and the block period. The minimum number of extremes was set to 10, as fewer points would lead to a bad fit to the GPD [18]. The minimum time window between extremes is required in order to decluster the extreme values and is usually in the range of 2 to 14 days [18]. The minimum time window between extremes was set to 4 days, as it provided a tighter confidence interval compared to 8 or 12 days [19]. The block period was set to 1 year (8760 h). Figure 11a shows the H_S threshold selection using the mean residual life method for LA1. An H_S threshold of 4.31 m was chosen for LA1. Figure 11c shows the U_W threshold selection using the mean residual life method for LA1. A U_W threshold of 20.91 m/s was chosen for LA1. Figure 11b,d show a sample of H_S and U_W data, respectively, for 30 days from LA1, with one selected observation which exceeded the chosen threshold. It should be kept in mind that one value of U_W or H_S which is exceeding the threshold in one month will lead to roughly 372 data points over a period of 31 years. The data points exceeding higher thresholds are known to have a better fit to the GPD than those exceeding lower thresholds. Hence, it is recommended to choose a threshold that is as high as possible while ensuring that a minimum number of data points is available for analysis [19].

Similarly, Figure 12a shows the H_S threshold selection using the mean residual life method for LA22. An H_S threshold of 4.60 m was chosen for LA22. Figure 12c shows the U_W threshold selection using the mean residual life method for LA22. A U_W threshold of 17.95 m/s was chosen for LA1. Figure 12b,d show a sample of H_S and U_W data, respectively, for 30 days from LA22, with one selected observation that exceeded the chosen threshold.

A sensitivity study was performed in order to study the effect of the selected threshold and time window on the extreme values obtained from the POT approach. Figure 13a,b show the effect of the threshold on the number of data points for U_W for ten selected sites. It should be noted that U_W thresholds ranging between 14 to 24 m/s will provide a sufficient number of data points for all of the sites. It is desirable to select the threshold in such a way that the number of data points extracted is close to the minimum requirement of 10. Figure 13c,d show the effect of the threshold on the number of data points for H_S for ten selected sites. The H_S threshold is expected to lie between between 2 to 5 m based on the number of data points available. The requirement set for a minimum number of data points ensures that thresholds providing too few points are not considered for analysis.
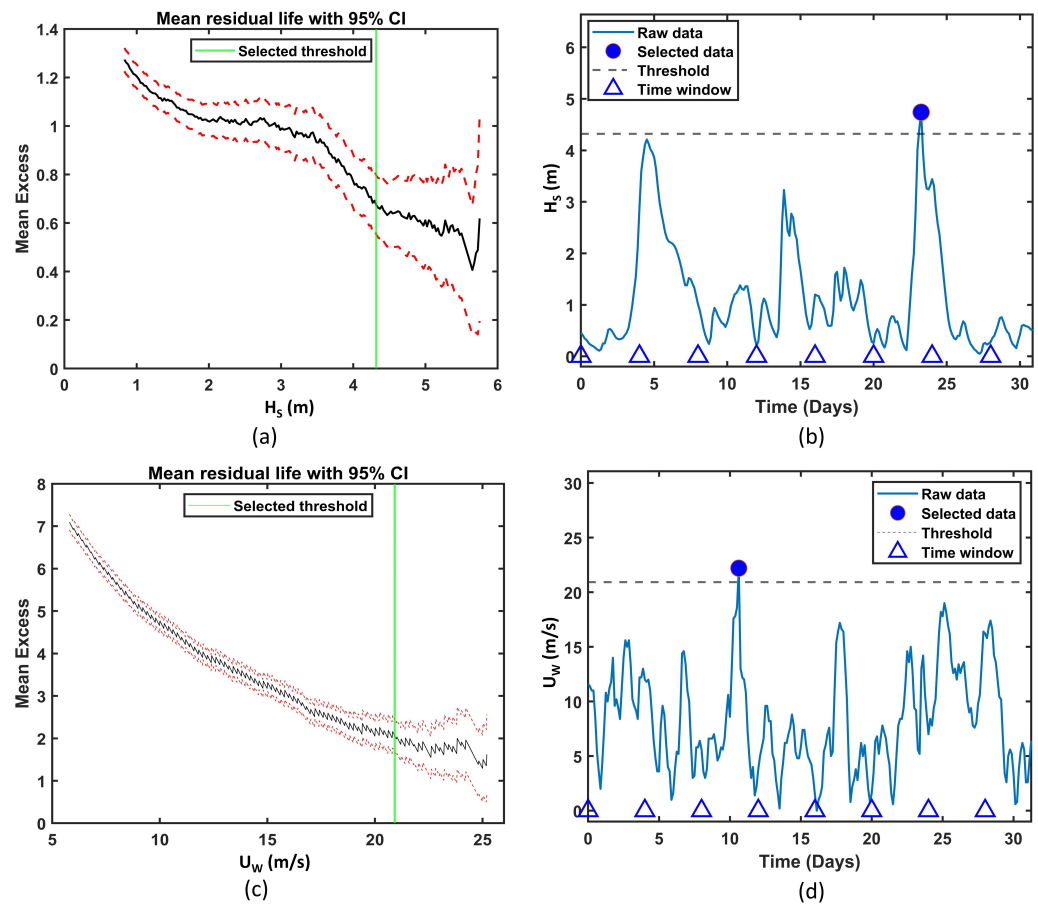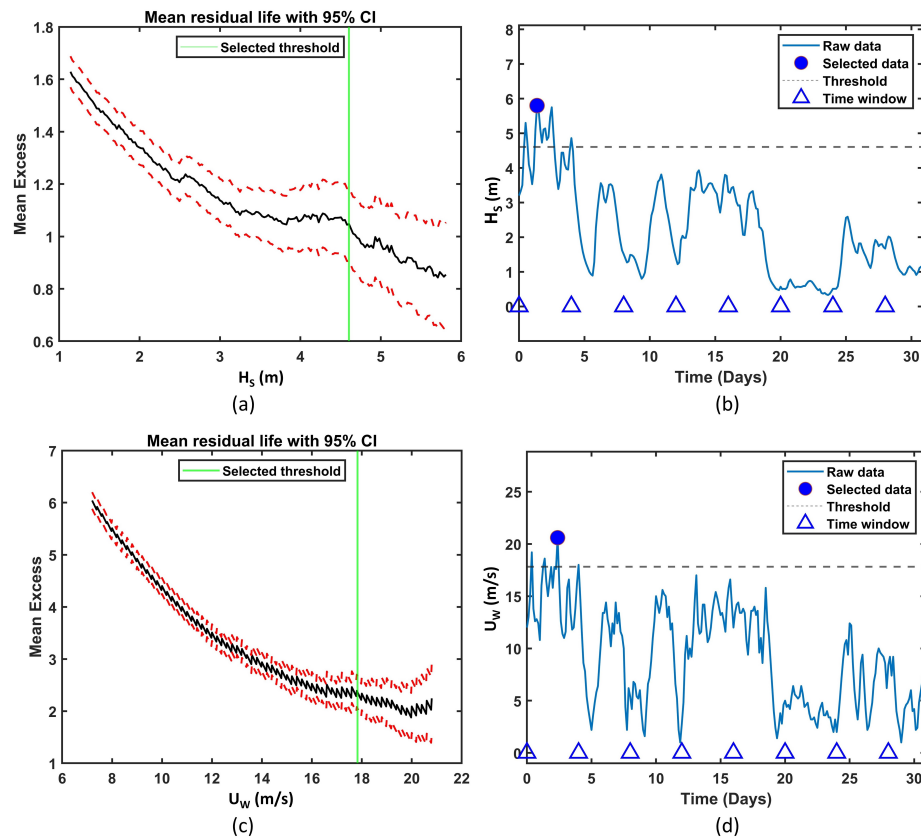
Figure 11. Threshold selection for LA1: (**a**) H_S threshold selection using mean residual life method. (**b**) Sample data reduction using the selected threshold for U_W. (**c**) Wind speed threshold selection using mean residual life method. (**d**) Sample data reduction using the selected threshold for U_W.

Figure 14a,b show the effects of the threshold and the time window between extremes on the 50-year H_s and the 50-year wind speed for LA1, respectively. In both the cases, there is a slight variation for smaller values of the threshold. However, above a certain limit, the time window between extremes did not have any effect on the predicted extreme values of H_S for LA1. Figure 14c,d show the effects of the threshold and the time window between extremes on the 50-year H_s and 50-year wind speed for LA22, respectively. There was no variation in the results for different time windows for the selected threshold in all of the cases. The Anderson–Darling test was performed for all of the excesses selected from each site. The *p* value was found was found to be less than 0.05 in all the cases. Hence, the null hypothesis that the excesses follow the GPD cannot be rejected at the chosen significance level of 5%. Figure 15a,b show the comparison of the U_W and H_S excesses, respectively, which are fitted to the GPD CDF and compared with the empirical CDF for LA1. It is clearly evident that the data closely followed the GPD distribution. The same trend was observed for all of the sites. Thus, the data points extracted were fit for the POT analysis. The GPD parameters were then obtained for each lease area for U_W and H_S, respectively.

Figure 12. Threshold selection for LA22: (**a**) H_S threshold selection using mean residual life method. (**b**) Sample data reduction using the selected threshold for H_S. (**c**) U_W threshold selection using mean residual life method. (**d**) Sample data reduction using the selected threshold for U_W.
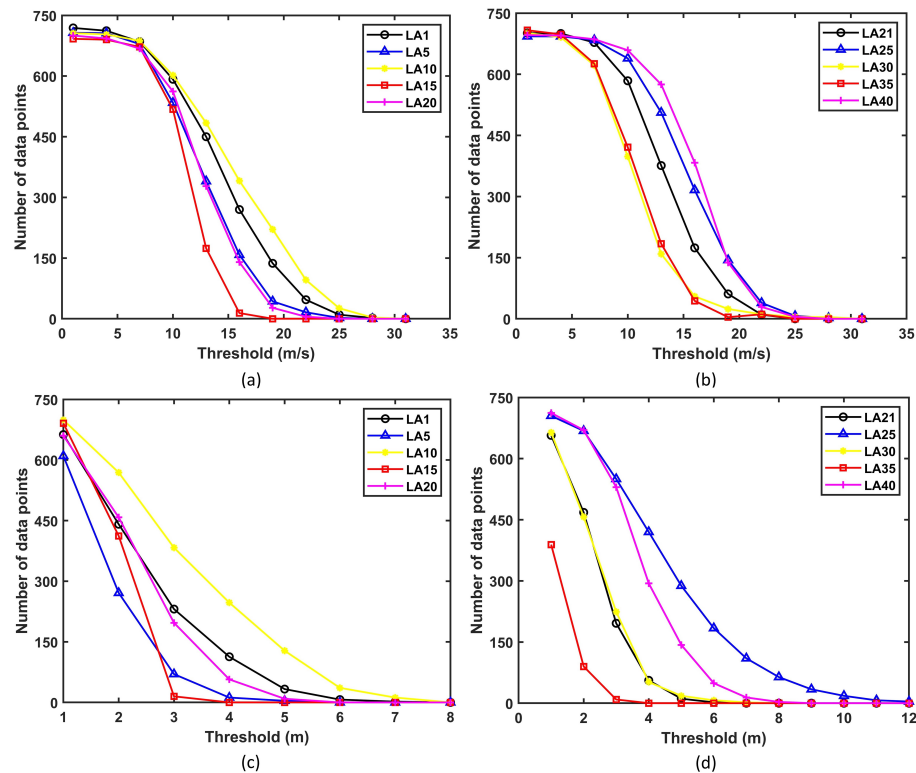
Figure 13. Effect of threshold on the number of data points for (**a**) U_W LA1–20; (**b**) U_W for LA21–40; (**c**) H_S for LA1–20; and (**d**) H_S for LA21–40.
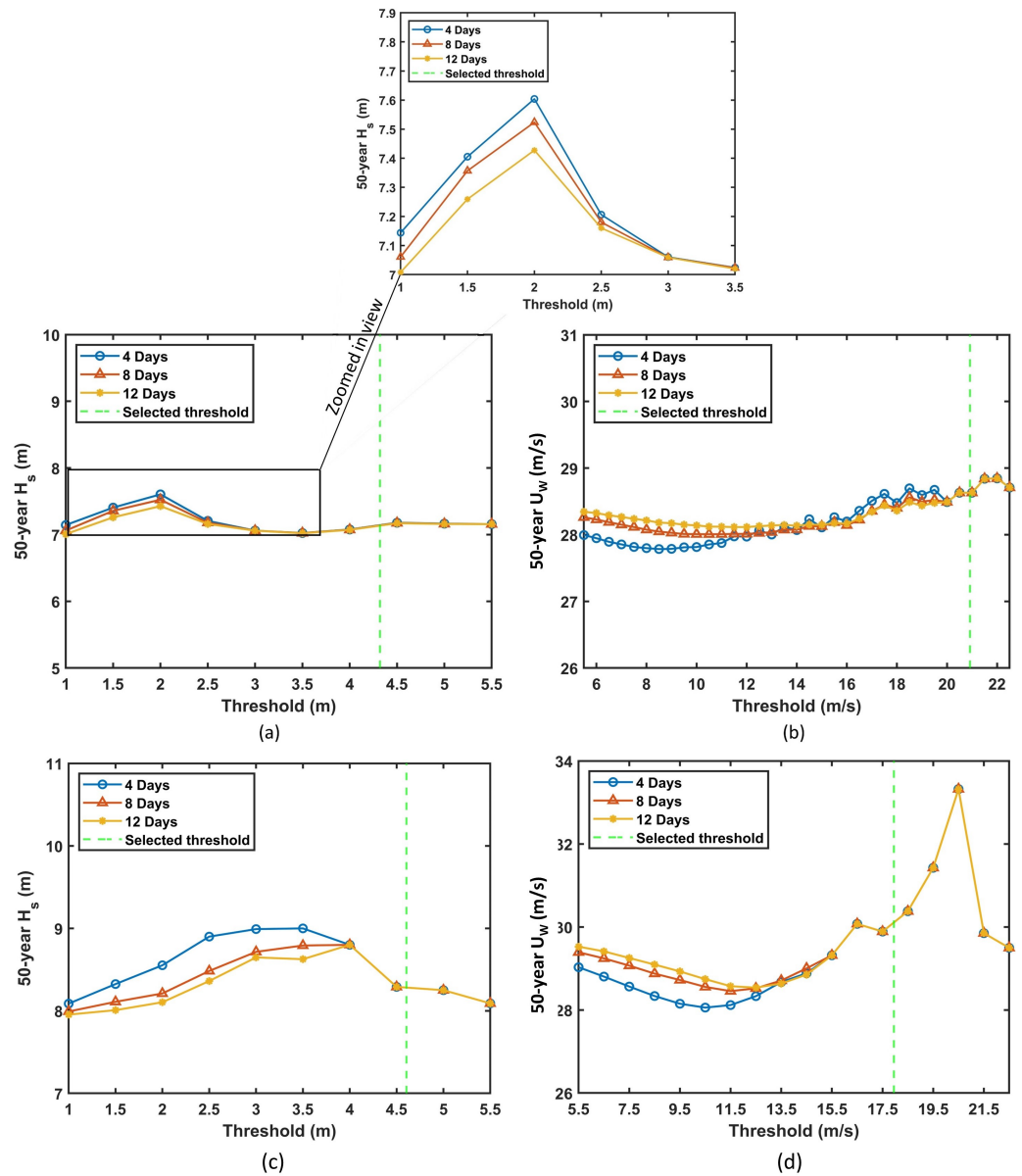
Figure 14. Sensitivity analysis: Effects of threshold and time window on (**a**) 50-year H_S for LA1; (**b**) 50-year wind speed for LA1; (**c**) 50-year H_S for LA22; and (**d**) 50-year wind speed for LA22.
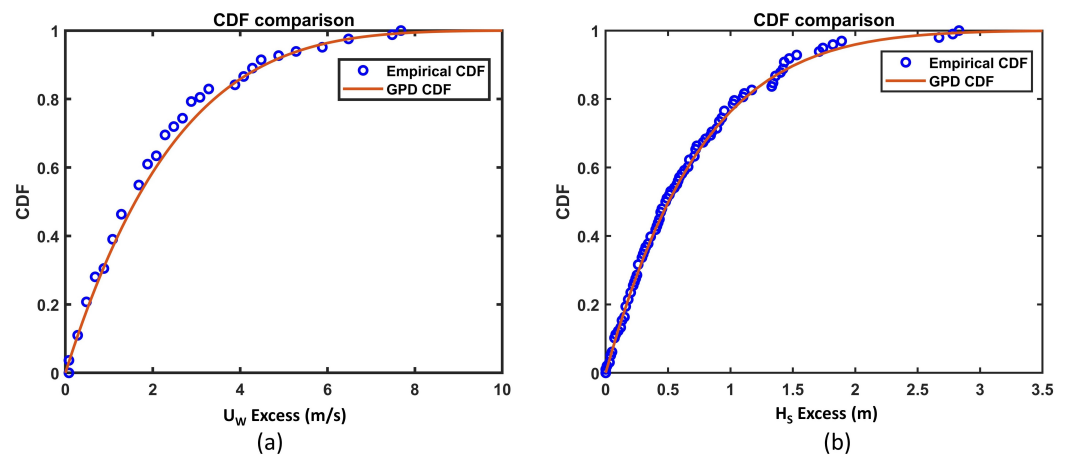
Figure 15. Comparison of GPD and empirical CDF for LA1: (**a**) U_W. (**b**) H_S.

3.3. Comparison of Extreme Values from the Block-Maxima and POT Approaches

The extreme values were obtained for U_W and H_S using a total of five different methods. Table 2 shows the parameters obtained for U_W for all of the lease areas. The location parameter, μ, for the POT approach was found to be zero for all of the sites and is not considered in the table. It was observed that several sites had a location parameter in the range of 20–23. However, LA32 (Minyang Jieyang Qianzhan III, China) had the highest scale parameter from all of the methods. This indicates that LA32 has the highest number of occurrences of extreme U_W. Table 3 shows the parameters obtained for H_S for all of the lease areas. It is observed that the location and scale parameters of LA25 (GoliatVIND, Norway) were comparatively very large for all of the methods. This indicates that this site possesses the highest average wave height, as well as a large number of occurrences of extreme wave heights.

Figure 16a,b show the comparison of 50-year H_s and 500-year H_s, respectively, obtained from the block-maxima and POT approaches for LA1. It is observed that a threshold of around of 2 m for the H_S produced extreme values, which are in the same range as the block-maxima methods. However, it is important to note that such a low threshold was not optimal according to the mean residual life method. Figure 16c,d show the comparison of 50-year U_W and 500-year U_W, respectively, obtained from the block-maxima and POT approaches for LA1. It is observed that the POT approach largely provided low estimates for the 50 and 500-year extreme U_W values, which has also been observed in previous studies [20]. Figure 17a,b show the comparison of 50-year H_s and 500-year H_s, respectively, obtained from the block-maxima and POT approaches for LA22. It is once again noted that the POT approach provided slightly lower estimates for extreme H_s values. Figure 17c,d show the comparison of 50-year U_W and 500-year U_W, respectively, obtained from the block-maxima and POT approaches for LA22. For LA22, the POT approach provided very high estimates for higher thresholds. However, it is worth remembering that when very large threshold values are selected, there might not be enough data points exceeding the threshold. This would result in a bad fit to the GPD distribution and ultimately to erroneous results. For the selected threshold, the estimates from the POT approach were very close to the lower bound of the estimates from the block-maxima methods.

Figure 18a,b show the comparison of the hazard curves obtained from different methods for the U_W and H_S, respectively, for LA1 (Maine research array, USA). It is observed that the results from the POT approach showed very little variation with increasing return periods. Figure 18c,d show the comparison of the hazard curves obtained from different methods for the U_W and H_S, respectively, for LA22 (Vigso bay, Denmark). The shape of the hazard curve was dependent on the distribution parameters obtained. It is observed that the wave height hazard curve obtained using the GEVD method for LA22 showed less variation (i.e., it was flat) with the return periods compared to other methods. The U_W hazard curve obtained using the GEVD method for LA22 was exceeding the Gumbel LS for large return periods. Figure 19a,b show the comparison of hazard curves obtained from different methods for the U_W and H_S, respectively, for LA25. It is evident that LA25 (GoliatVIND, Norway) is prone extremely large wave heights. This site also has a considerably high extreme U_W. Figure 19c,d show the comparison of the hazard curves obtained from different methods for the U_W and H_S respectively for LA32 (Minyang Jieyang Qianzhan III, China). The 500-year U_W at 10 m above sea level was found to be close to 50 m/s using the GEVD method for LA32. Both the GEVD method and the Gumbel LS method provided conservatively high estimates for all of the four sites.
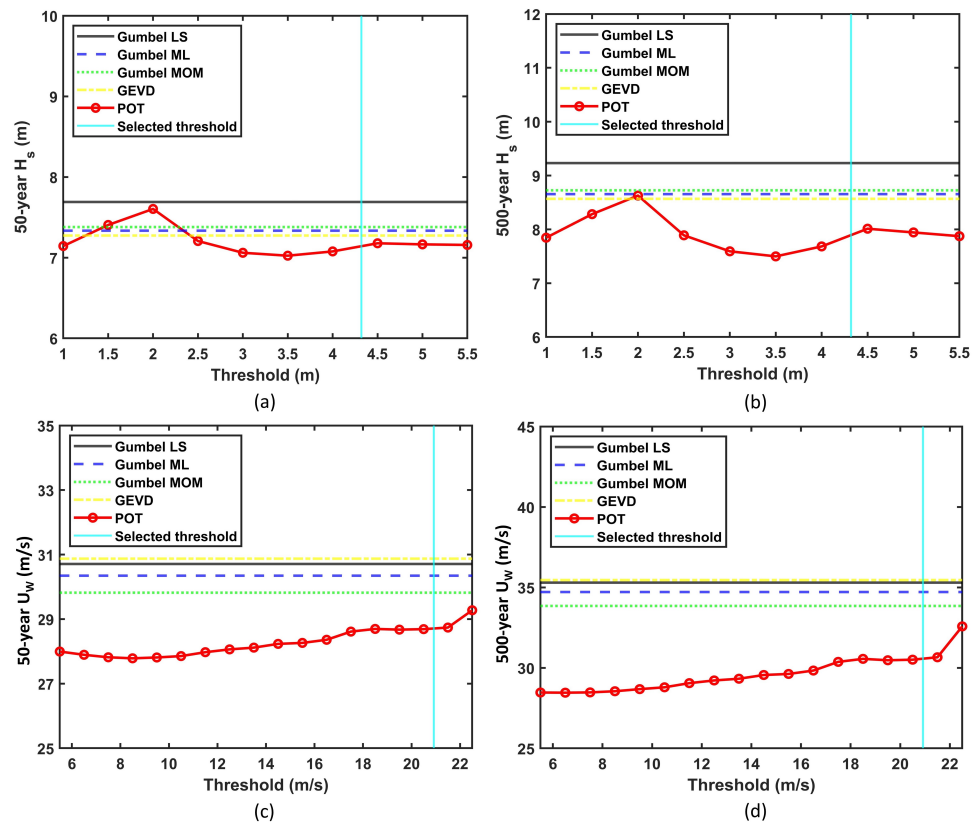
Figure 16. Comparison between block-maxima and POT approaches for LA1: (**a**) 50-year H_S. (**b**) 500-year H_S. (**c**) 50-year U_W. (**d**) 500-year U_W.
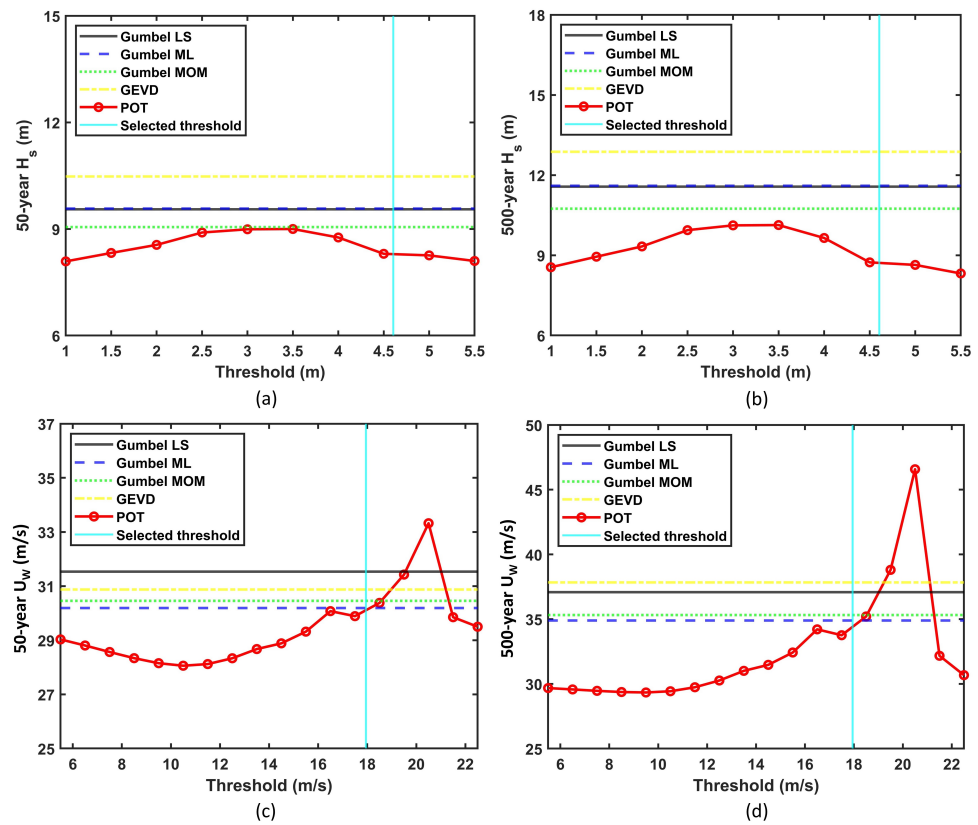

Figure 17. Comparison between block-maxima and POT approaches for LA22: (**a**) 50-year U_W. (**b**) 500-year U_W. (**c**) 50-year H_S. (**d**) 500-year H_S.
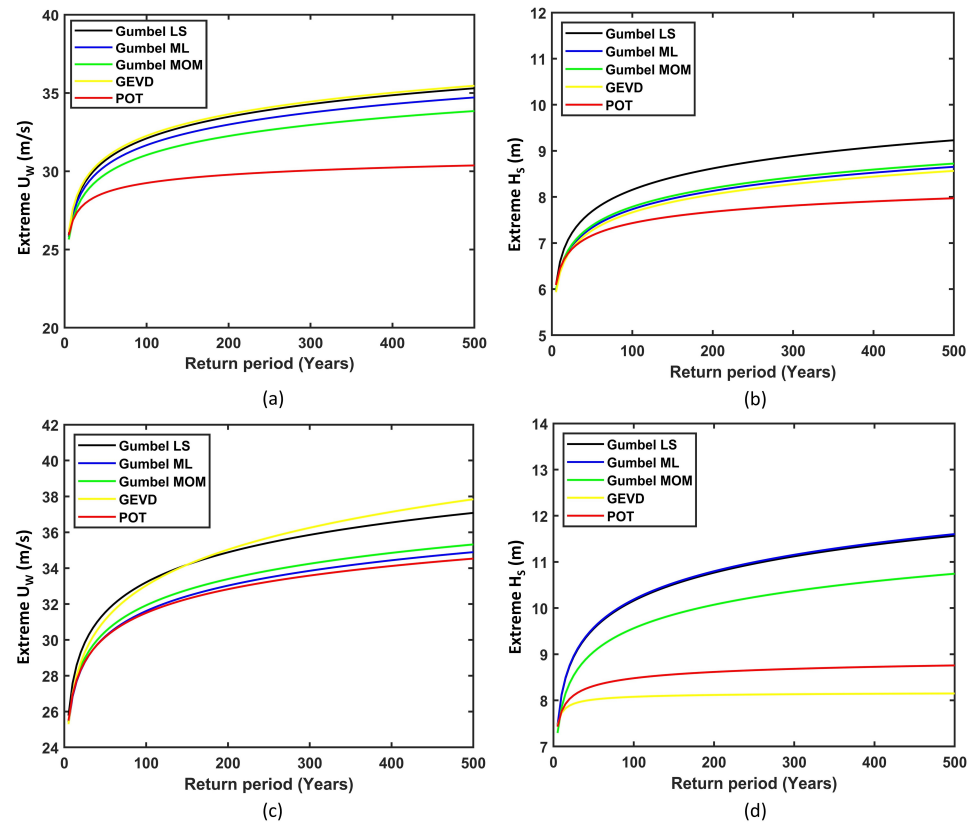
Figure 18. Comparison of hazard curves: (**a**) LA1 Extreme U_W. (**b**) LA1 Extreme H_S. (**c**) LA22 Extreme U_W. (**d**) LA22 Extreme H_S.
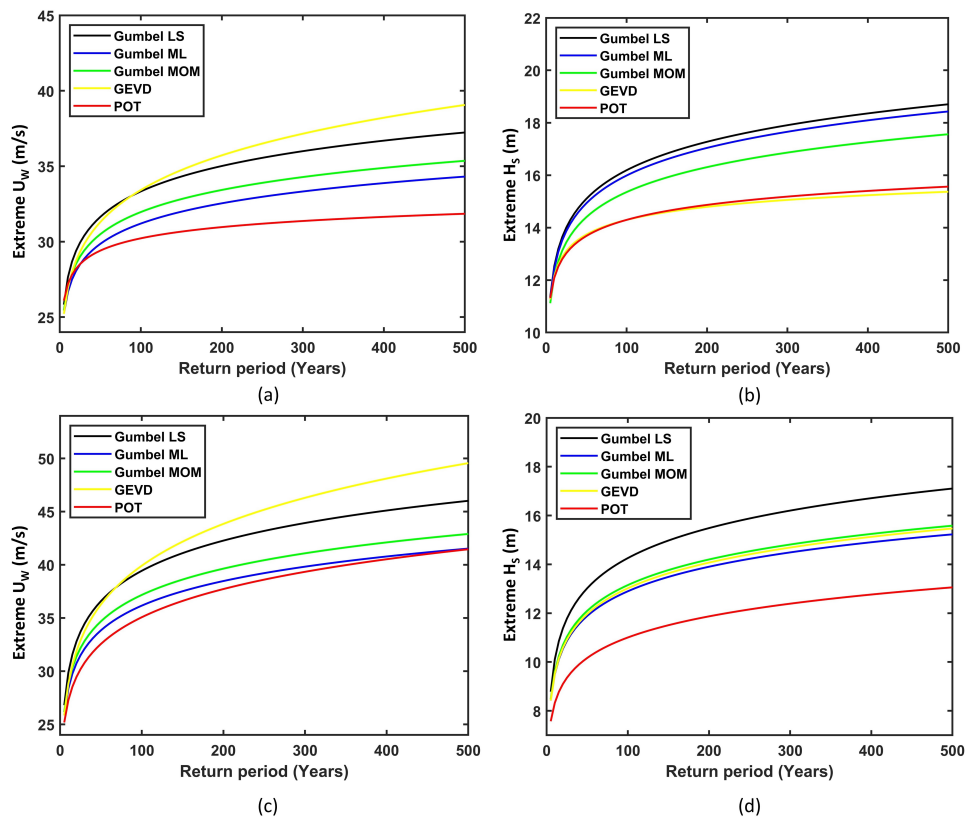
Figure 19. Comparison of hazard curves: (**a**) LA25 Extreme U_W. (**b**) LA25 Extreme H_S. (**c**) LA32 Extreme U_W. (**d**) LA32 Extreme H_S.

Figure 20a,b show the comparison of the 50- and 500-year U_W values, respectively. The sites are demarcated based on the continent in which they are located. It is observed that Asian sites, especially those in China, were more prone to the highest U_W. For certain sites such as LA32 (Minyang Jieyang Qianzhan III, China), the upper bounds of the 50- and 500-year U_W values were obtained from different methods. They were once again dependent on the distribution parameters, as seen in the hazard curve of the same site earlier. LA32 also had the highest 50- and 500-year U_W values among all the sites. Figure 21a,b show the comparison of 50- and 500-year H_S values, respectively. It is noted that the extreme H_s was comparatively lower in South American and Asian sites. LA16 (Projeto Ubu) had the lowest 50- and 500-year H_s values, and LA25 (GoliatVIND, Norway) had the highest 50- and 500-year H_s values, which is consistent with the findings in the annual maximum values. LA31 (Huaneng Daishan I, China) has a relative high wind potential, but low wave heights are prevailing in this area. Once again, the Gumbel LS and the GEVD methods provided the highest estimates for the extreme values for all of the sites. The results from the POT approach were less than the lower bound of the block-maxima results by around 3% on average. However, there were some sites where the results from the POT were significantly lower, which is due to the selected threshold.
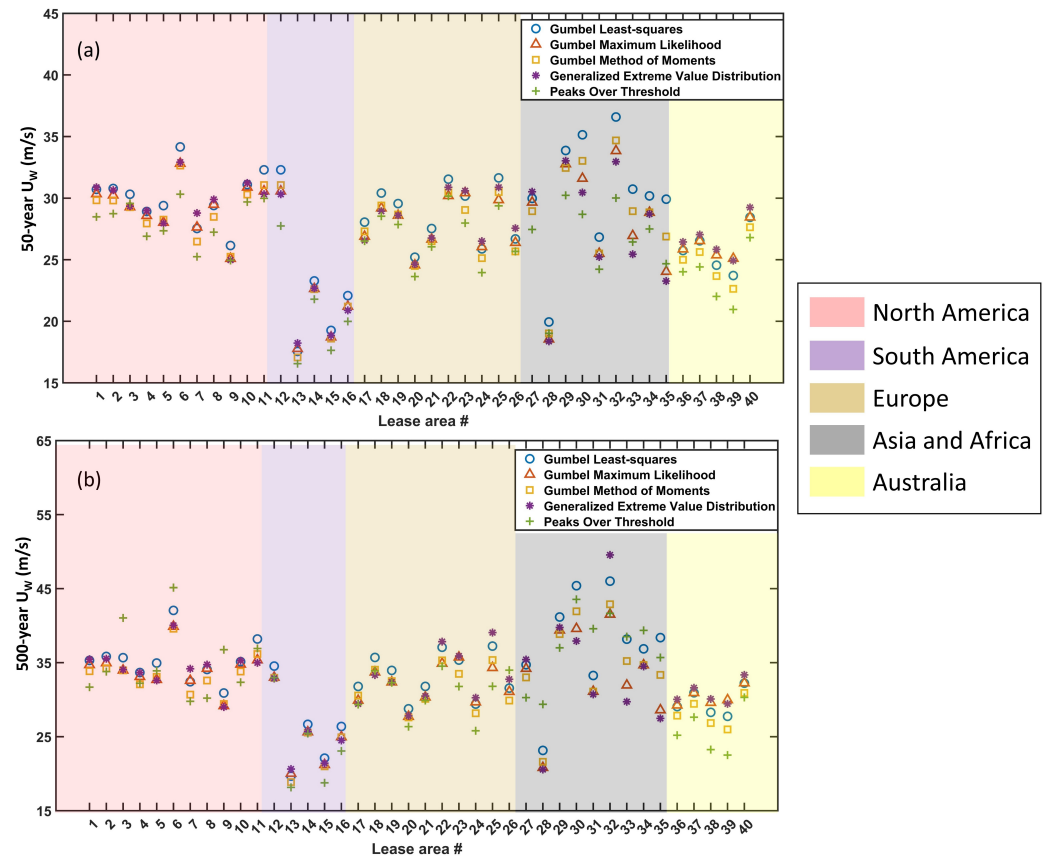

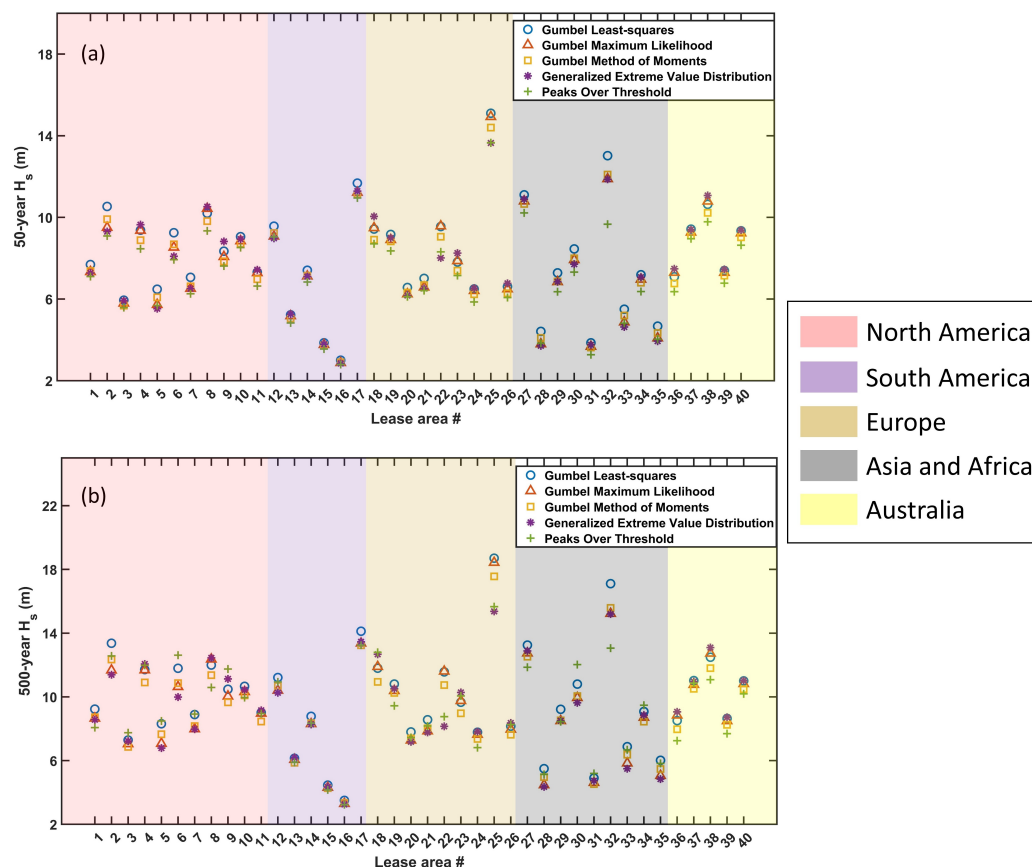**Figure 20.** Comparison of extreme U_W values obtained from different statistical methods: (**a**) 50-year U_W. (**b**) 500-year U_W.

Figure 21. Comparison of extreme H_S values obtained from different statistical methods: (**a**) 50-year H_S. (**b**) 500-year H_S.

4. Conclusions

A comparison of five different statistical techniques for estimating the extreme U_W and wave heights is presented in this paper. The Gumbel LS, Gumbel ML, Gumbel MOM, and the GEVD methods employed the block-maxima approach to estimate the extreme values. Additionally, the extreme values were also estimated using the peaks-over-threshold method. Verification studies for the block-maxima approach were carried out by fitting the annual maximum values to the Gumbel probability paper, as well as by comparing them with the empirical CDF. Verification for the POT approach was done by fitting the excesses to the GPD distribution. The visual goodness of fit was found to be satisfactory in all the cases. All of the distribution parameters obtained for the forty sites have been presented. The extreme U_W and H_S values for different return periods have been presented. The results obtained from this study will aid in the planning, design, and risk assessment of new offshore wind farms around the world. The following conclusions were drawn from the results:

- LA30 (Huaneng Hainan Wenchang I, China) recorded the outright highest annual maximum U_W of 36.2 m/s, while LA10 (Allan array, Canada) recorded the highest mean annual maximum U_W of 25.2 m/s.
- LA25 (GoliatVIND, Norway) recorded the highest mean annual maximum H_S of 9.85 m, as well as the outright highest annual maximum H_S of 14.2 m.
- The GPD CDF and Gumbel CDF showed good agreement with the empirical CDF for both the U_W and H_S values for all of the sites.
- The results from the POT approach varied significantly based on the chosen threshold.
- For smaller thresholds, the results from the POT approach were sensitive to the time window chosen. However, the time window did not have an impact on the results for larger thresholds, which were generally obtained from the mean residual life method.

- The POT approach is only effective for a small range of thresholds. Smaller thresholds lead to a poor fit to the GPD, while larger thresholds may not provide sufficient data points.
- The 50-year U_W at 10 m above the sea level ranged between 16.5 m/s and 36.5 m/s. The 500-year U_W ranged from 18.1 m/s to 49.5 m/s.
- The 50-year H_S lay between 2.8 m and 15.1 m. The 500-year H_S varied from 3.2 m to 18.7 m.
- It is found that the block-maxima approach using the Gumbel LS and GEV distributions provides upper bound estimates for the 50- and 500-year extreme values for both the U_W and H_S.
- The estimates from the POT approach were generally lower by around 3% on average, although there were some outliers.
- European sites are more prone to extreme H_S values in general. LA25 (GoliatVIND, Norway) produced the highest 50- and 500-year H_S values of 15.1 m (Gumbel LS) and 18.7 m (Gumbel LS), respectively.
- Sites along the east coast of China have high estimates of extreme U_W values. LA32 (Minyang Jieyang Qianzhan III, China) was prone to the highest 50- and 500-year U_W values of 36.5 m/s (Gumbel LS) and 49.5 m/s (GEVD), respectively.
- The distribution parameters were provided for all of the methods, which would be helpful for extrapolating the extreme values to longer return periods.
- The mean residual life method used for estimating the optimal threshold has yielded results that lie close to or within the bounds of the estimates from the block-maxima approach.

Future Work

Since the extreme values obtained using the POT approach were highly sensitive to the chosen threshold, it is worth investigating the dispersion index method for finding the optimal threshold. A comparison between the results from the mean residual life method and the dispersion index method would add more insights to the POT approach. The extreme values obtained using the Waveclimate and ERA5 datasets can be compared for verification. It is also recommended to include other statistical methods such as the ACER method and other parameter estimation methods such as the methods of L-moments in these comparative studies. In addition, the statistical methods used in this study need to be investigated more rigorously in order to find the most suitable distribution and parameter estimation method for the prediction of extreme wind and waves. Novel statistical methods can be developed to provide more accurate extreme values. New metrics can also be developed to quantify the error and compare each of the methods. The extreme wind and wave predictions must be validated with buoy measurements for at least a few sites. It would also be interesting to study the response of the OWTs subjected to extreme conditions obtained through different statistical methods. The effect of hurricanes on offshore wind turbines is an area of particular interest, especially in countries such as the USA. Such studies will help in identifying the bottlenecks associated with extreme environmental conditions. More offshore wind lease areas can also be included in future studies.

References

1. GWEC. Global Wind Report. 2023. Available online: https://gwec.net/wp-content/uploads/2023/04/GWEC-2023_interactive.pdf (accessed on 22 July 2023).
2. Barthelmie, R.J.; Dantuono, K.E.; Renner, E.J.; Letson, F.L.; Pryor, S.C. Extreme wind and waves in US east coast offshore wind energy lease areas. *Energies* **2021**, *14*, 1053. [CrossRef]
3. IRENA. World Energy Transitions Outlook 2023: 1.5 °C Pathway. 2023. Available online: https://mc-cd8320d4-36a1-40ac-83cc-3389-cdn-endpoint.azureedge.net/-/media/Files/IRENA/Agency/Publication/2023/Jun/IRENA_World_energy_transitions_outlook_v1_2023.pdf?rev=261b3ae18f70429ea8cf595d5a4bee18 (accessed on 22 July 2023).
4. Putuhena, H.; White, D.; Gourvenec, S.; Sturt, F. Finding space for offshore wind to support net zero: A methodology to assess spatial constraints and future scenarios, illustrated by a UK case study. *Renew. Sustain. Energy Rev.* **2023**, *182*, 113358. [CrossRef]
5. Zhao, Z.; Li, X.; Wang, W.; Shi, W. Analysis of dynamic characteristics of an ultra-large semi-submersible floating wind turbine. *J. Mar. Sci. Eng.* **2019**, *7*, 169.
6. Hopewell, P.; Castro-Sayas, F.; Bailey, D. Optimising the design of offshore wind farm collection networks. In Proceedings of the IEEE 41st International Universities Power Engineering Conference, Newcastle Upon Tyne, UK, 6–8 September 2006; Volume 1, pp. 84–88.
7. *IEC 61400-3-1:2019*; Wind Turbines Part 3-1: Design Requirements for Fixed Offshore Wind Turbines; Edition 1.0 2019-04; IEC FDIS: Geneva, Switzerland, 2019. p. 314.
8. Pryor, S.C.; Barthelmie, R.J. A global assessment of extreme wind speeds for wind energy applications. *Nat. Energy* **2021**, *6*, 268–276.
9. Izaguirre, C.; Méndez, F.J.; Menéndez, M.; Losada, I.J. Global extreme wave height variability based on satellite data. *Geophys. Res. Lett.* **2011**, *38*. [CrossRef]
10. Lee, B.H.; Ahn, D.J.; Kim, H.G.; Ha, Y.C. An estimation of the extreme wind speed using the Korea wind map. *Renew. Energy* **2012**, *42*, 4–10. [CrossRef]
11. Sacré, C. Extreme wind speed in France: The'99 storms and their consequences. *J. Wind. Eng. Ind. Aerodyn.* **2002**, *90*, 1163–1171. [CrossRef]
12. Torrielli, A.; Repetto, M.P.; Solari, G. The annual rate of independent events for the analysis of the extreme wind speed. *J. Wind. Eng. Ind. Aerodyn.* **2016**, *156*, 104–114. [CrossRef]
13. Palutikof, J.P.; Brabson, B.; Lister, D.H.; Adcock, S. A review of methods to calculate extreme wind speeds. *Meteorol. Appl.* **1999**, *6*, 119–132.
14. Wang, J.; Qin, S.; Jin, S.; Wu, J. Estimation methods review and analysis of offshore extreme wind speeds and wind energy resources. *Renew. Sustain. Energy Rev.* **2015**, *42*, 26–42.
15. Hong, H.; Li, S.; Mara, T. Performance of the generalized least-squares method for the Gumbel distribution and its application to annual maximum wind speeds. *J. Wind. Eng. Ind. Aerodyn.* **2013**, *119*, 121–132. [CrossRef]
16. Lombardo, F.T. Improved extreme wind speed estimation for wind engineering applications. *J. Wind. Eng. Ind. Aerodyn.* **2012**, *104*, 278–284. [CrossRef]
17. Afzal, M.S.; Kumar, L.; Chugh, V.; Kumar, Y.; Zuhair, M. Prediction of significant wave height using machine learning and its application to extreme wave analysis. *J. Earth Syst. Sci.* **2023**, *132*, 51. [CrossRef]
18. Simiu, E.; Heckert, N.A. Extreme wind distribution tails: A "peaks over threshold" approach. *J. Struct. Eng.* **1996**, *122*, 539–547. [CrossRef]
19. Viselli, A.M.; Forristall, G.Z.; Pearce, B.R.; Dagher, H.J. Estimation of extreme wave and wind design parameters for offshore wind turbines in the Gulf of Maine using a POT method. *Ocean. Eng.* **2015**, *104*, 649–658. [CrossRef]
20. An, Y.; Pandey, M. A comparison of methods of extreme wind speed estimation. *J. Wind. Eng. Ind. Aerodyn.* **2005**, *93*, 535–545. [CrossRef]
21. Kang, D.; Ko, K.; Huh, J. Determination of extreme wind values using the Gumbel distribution. *Energy* **2015**, *86*, 51–58. [CrossRef]
22. Rivas, D.; Caleyo, F.; Valor, A.; Hallen, J. Extreme value analysis applied to pitting corrosion experiments in low carbon steel: Comparison of block maxima and peak over threshold approaches. *Corros. Sci.* **2008**, *50*, 3193–3204. [CrossRef]
23. Vinoth, J.; Young, I. Global estimates of extreme wind speed and wave height. *J. Clim.* **2011**, *24*, 1647–1665. [CrossRef]
24. Jonathan, P.; Ewans, K. Uncertainties in extreme wave height estimates for hurricane-dominated regions. *J. Offshore Mech. Arct. Eng.* **2007**, *129*, 300–305. [CrossRef]
25. Pandey, M.D.; Van Gelder, P.; Vrijling, J. The estimation of extreme quantiles of wind velocity using L-moments in the peaks-over-threshold approach. *Struct. Saf.* **2001**, *23*, 179–192. [CrossRef]

26. Karpa, O.; Naess, A. Extreme value statistics of wind speed data by the ACER method. *J. Wind. Eng. Ind. Aerodyn.* **2013**, *112*, 1–10. [CrossRef]
27. Gaidai, O.; Xing, Y.; Balakrishna, R.; Xu, J. Improving extreme offshore wind speed prediction by using deconvolution. *Heliyon* **2023**, *9*, e13533. [CrossRef] [PubMed]
28. 4COffshore. Offshore Wind Energy Map. 2023. Available online: https://map.4coffshore.com/offshorewind/ (accessed on 21 June 2023).
29. NCEI. The Multibeam Bathymetry Database (MBBDB). 2023. Available online: https://www.ncei.noaa.gov/maps/bathymetry/ (accessed on 21 June 2023).
30. WaveClimate Infoplaza. Available online: http://www.waveclimate.com/ (accessed on 20 June 2023).
31. Peter Groenewoud, S.H. *Validation of the BMTA 35-Year Hindcast Database v361*; Technical Report, BMT ARGOSS; Houten, The Netherlands, 2016. Available online: http://www.waveclimate.com/clams/redesign/help/docs/I113_Validation_BMTA_35 -year_Hindcast_17jun2016.pdf (accessed on 21 June 2023).
32. Bali, T.G. The generalized extreme value distribution. *Econ. Lett.* **2003**, *79*, 423–427. [CrossRef]
33. Cooray, K. Generalized gumbel distribution. *J. Appl. Stat.* **2010**, *37*, 171–179. [CrossRef]
34. Brodtkorb, P.A.; Johannesson, P.; Lindgren, G.; Rychlik, I.; Rydén, J.; Sjö, E. WAFO-a Matlab toolbox for analysis of random waves and loads. In Proceedings of the ISOPE International Ocean and Polar Engineering Conference (ISOPE 2000), Seattle, WA, USA, 27 May–2 June 2000.
35. Mahdi, S.; Cenac, M. Estimating Parameters of Gumbel Distribution using the Methods of Moments, probability weighted Moments and maximum likelihood. *Rev. Mat. Teor. Apl.* **2005**, *12*, 151–156. [CrossRef]
36. Gibson, R.; Grant, C.; Forristall, G.Z.; Smyth, R.; Owrid, P.; Hagen, O.; Leggett, I. Bias and uncertainty in the estimation of extreme wave heights and crests. In Proceedings of the International Conference on Offshore Mechanics and Arctic Engineering, Honolulu, HI, USA, 31 May–6 June 2009; Volume 43420, pp. 363–373.
37. Castillo, E.; Hadi, A.S. Fitting the generalized Pareto distribution to data. *J. Am. Stat. Assoc.* **1997**, *92*, 1609–1620. [CrossRef]
38. Greenwood, P.E.; Nikulin, M.S. *A Guide to Chi-Squared Testing*; John Wiley & Sons: Hoboken, NJ, USA, 1996; Volume 280.
39. D'Agostino, R. *Goodness-of-Fit-Techniques*; Routledge: Oxford, UK, 2017.
40. Anderson, T.W.; Darling, D.A. A test of goodness of fit. *J. Am. Stat. Assoc.* **1954**, *49*, 765–769. [CrossRef]
41. Chu, J.; Dickin, O.; Nadarajah, S. A review of goodness of fit tests for Pareto distributions. *J. Comput. Appl. Math.* **2019**, *361*, 13–41. [CrossRef]
42. NOAA. Tropical Cyclone Climatology. 2023. Available online: https://www.nhc.noaa.gov/climo/ (accessed on 6 July 2023).
43. Colbert, A. A Force of Nature: Hurricanes in a Changing Climate. 2022. Available online: https://climate.nasa.gov/news/3184 /a-force-of-nature-hurricanes-in-a-changing-climate/ (accessed on 20 July 2023).
44. Breivik, Ø.; Aarnes, O.J.; Abdalla, S.; Bidlot, J.R.; Janssen, P.A. Wind and wave extremes over the world oceans from very large ensembles. *Geophys. Res. Lett.* **2014**, *41*, 5122–5131. [CrossRef]