EXTREME WAVE HEIGHT ANALYSIS IN NATUNA SEA USING PEAK-OVER THRESHOLD METHOD

Ismail Abdul Jabbar¹ and Nining Sari Ningsih²,*

¹Study Program of Oceanography, Faculty of Earth Sciences and Technology, Bandung Institute of Technology
²Research Group of Oceanography, Faculty of Earth Sciences and Technology, Bandung Institute of Technology
Jl. Ganesha 10, Bandung 40132, Indonesia, Fax. +62-22-2534139, Phone. +62-22-2505253

*Corresponding author: nining@fitb.itb.ac.id

Abstract. The extreme wave height analysis has been conducted in Natuna Sea, Indonesia, using 25 years (1991-2015) significant wave height (SWH) data from WAVEWATCH 3 (WW3) with a spatial resolution of 1/8°. Natuna Sea is geographically connected to the South China Sea (SCS) which is frequently crossed by tropical cyclones. These cyclones may have contributed to existences of high waves in the SCS, which could propagate as swell waves to the Natuna Sea and lead to extreme waves in this region. The extreme analysis has been done by classifying extreme events of SWHs using Peak-Over Threshold (POT) method with a fixed threshold level at quantile 0.93 and a minimum time separation of 48 hours between two successive extreme events. Furthermore, Generalized Pareto (GP) distribution has been applied to estimate return values of extreme SWHs for several return periods. Parameters of the GP distribution have been estimated by Maximum Likelihood Estimation (MLE) method. The characteristics of extreme SWHs in the Natuna Sea have been explained by maximum and seasonal distribution plots. Maximum values of extreme SWHs in the SCS could reach 13 m and around 3-5 m in the Natuna Sea. The seasonal distributions of extreme waves indicated that the occurrences of extreme waves in the SCS during Northeast Winter Monsoon (NWM) were higher than those during Southwest Summer Monsoon (SSM). Seasonal mean and maxima of extreme SWHs in the Natuna Sea were also high during the NWM and low during the SSM. To examine the effects of swell waves from the SCS and also future extreme waves in the Natuna Sea, we have analyzed the characteristics and calculated return values of extreme waves in front and behind of Bunguran Island, where the former faces directly to the SCS. There were 172 (331) extreme waves from 1991-2015 in the front (behind) of Bunguran Island and mainly from the northeast (southwest). Most of them were around 2 – 4 (0.5 – 2) m with mean periods of 6 – 10 (3 – 6) s. Based on the return values in the front (behind) of the Bunguran Island, there are possibilities of extreme waves with values 4.70, 4.87, and 4.96 (2.08, 2.20, and 2.27) m for return periods of 25, 50, and 75-year, successively.
1. Introduction

Natuna Sea is adjacent to and directly connected to the South China Sea (SCS). The SCS is frequently crossed by tropical cyclones with intensity reaching a tropical storm or even stronger [23]. These storms can generate high waves and propagate as swell waves which could lead to extreme waves in another region. The SCS is under influences of monsoonal winds and the tropical cyclones in this region were also well associated with the monsoon seasons in the SCS [23] [11]. The tropical cyclone occurrence during May-September was higher than during October-December [23]. Over the last century, as the global climate change becomes an issue, the tropical cyclones have become a concern because climate change is presumed to affect the frequency and intensity of tropical cyclones. Nevertheless, the frequency of tropical cyclone in the SCS has a negative trend even though the trend was not statistically significant [14].

Meanwhile, Bunguran Island, located in the Natuna Sea, is one of the outer islands of Indonesia and directly affected by waves from the SCS. In 2018, there were approximately 53,473 people in total living in this island (districts of West Bunguran, Bunguran Batubi, North Bunguran, East Bunguran, Northeast Burungan, Central Burungan, and South Burungan) [19] and might be under threats of extreme waves. Moreover, the Bunguran Island has also become a military base of Indonesia [23] and been considered as a potential location of wave power plant [12]. Therefore, the analysis and characterization of extreme waves in the vicinity of the Bunguran Island can be valuable information for marine and coastal activity and protection.

Extreme wave height analysis is well recommended in a location commonly affected by extreme waves or extreme events such as tropical cyclones. One of its applications is to predict return values of extreme waves that can happen in certain return periods. This information is useful for coastal protection and management such as for safety in onshore or offshore activity, design of a ship, and design of an offshore and onshore structure [7] [18] [20]. Annual Maximum (AM) Approach and Peak-Over Threshold (POT) method are some common methods that can be used to analyze the extreme waves. The AM approach has several disadvantages since its reliability depends on the number of samples and classification of “a year” since it only uses one data for each “year”. The POT method has been proven to overcome this problem since this method includes all extreme cases [1] [5] [6]. Some studies which have investigated the extreme waves using the POT are [3] and [21] with a spatial resolution of 1.5° and 1/4°, respectively.

Previous studies in the Natuna Sea have been conducted by [2] and [12] to investigate the wave characteristics for naval base purposes and assessment of wave energy, consecutively. In this paper, we have investigated extreme wave characteristics and distributions in the Natuna Sea related to the SCS monsoon seasons using the POT method. WAVEWATCH 3 (WW3) significant wave height (SWH) data with a finer spatial resolution of 1/8° has been used to give more accurate results. The effects of swell waves from the SCS to the characteristics of the extreme waves in the Bunguran Island have also been examined. In addition, we have also calculated the return values to predict future possibilities of extreme waves in the Bunguran Island.

2. Data

The WW3 SWH data was obtained from Geospatial Information Agency of Indonesia (BIG) and can be accessed in http://tides.big.go.id/. The WW3 SWH data was available in the range 90°E – 145°E and 20°S – 20°N with 6-hour time step. However, domain coverage in this study is in the range of 99°E – 119°E and
1°S – 15°N encompassing the Natuna Sea and the SCS. We also have SWH data from some buoy measurements in several locations to verify the WW3 SWH data. These data were provided by the Agency for the Assessment and Application of Technology (BPPT) in collaboration with Oceanographic of Norway and Bandung Institute of Technology (ITB) on a real-time wave observation project named SEAWATCH [15]. Location and bathymetry of domain study and also the locations of the buoys are presented in Figure 1.

The WW3 SWH data has been verified in several locations to check the quality of the simulated SWHs. Data verification was conducted by calculating bias, root means square error (RMSE), and correlation coefficient (CC) between the WW3 SWH data and the SWHs from buoy measurements (Figure 1b and Table 1). The bias, RMSE, and CC were calculated using equations as follows [4]:

\[
\text{bias}(x, y) = \frac{1}{M} \sum_{i=1}^{M} (H_m(x, y, t_i) - H_o(x, y, t_i))
\]

\[
\text{RMSE}(x, y) = \frac{1}{M} \sum_{i=1}^{M} (H_m(x, y, t_i) - H_o(x, y, t_i))^2
\]

\[
\text{CC}(x, y) = \frac{\sum_{i=1}^{M} (H_m(x, y, t_i) - \bar{H}_m(x, y))) (H_o(x, y, t_i) - \bar{H}_o(x, y)))}{\sqrt{\sum_{i=1}^{M} (H_m(x, y, t_i) - \bar{H}_m(x, y))^2} \sqrt{\sum_{i=1}^{M} (H_o(x, y, t_i) - \bar{H}_o(x, y))^2}}
\]

where an overbar indicates time averaging, \(H_m\) is the WW3 SWH data, \(H_o\) is the observed SWH data from in-situ measurement, \(M\) is the number of samples, \(t\) is time, and \(x\) and \(y\) is longitude and latitude, respectively. From Table 1, the resulting CCs showed a good agreement between the WW3 and the observed SWH data. The estimated errors by RMSE were also quite low with values around 0.17 – 0.29 m, while the values of bias were slightly higher than zero. A positive value of bias indicates that the WW3 SWH data was generally higher than the observed SWH data. Additionally, [16] have also verified the same WW3 SWH data using a similar method in Baron, Indonesia, and the resulting bias, RMSE, and CC were 0.17 m, 0.44 m, and 0.59, respectively (see Figure 3 of [16]).

**Table 1.** Computed bias, RMSE, and CC between the WW3 SWH data and SWH from buoy observations

| Marker | Location | Coordinates | Time of Measurement* | bias (m) | RMSE (m) | CC |
|--------|----------|-------------|----------------------|---------|----------|----|
| P1 Jepara | 6.6°S and 110.6°E | 1998-10-28 to 1999-08-15 | 0.08 | 0.29 | 0.75 |
| P2 Karawang | 5.78°S and 107.77°E | 1999-09-01 to 1999-09-16 | 0.04 | 0.21 | 0.82 |
| P3 Masalembo | 5.57°S 114.62°E | 1998-10-30 to 1998-11-12 | 0.03 | 0.17 | 0.76 |

*Time of measurement is written in a format of year-month-day
3. Peak-Over Threshold (POT) Method

The POT method has been used widely to estimate the return values of extreme waves. In this method, an extreme event is defined as successive SWHs above a threshold value. This extreme event is often called a “storm” and the peak or the highest value of this “storm” is referred to as an extreme wave. For convenience, we will use a term of the extreme wave instead of the storm as not to be confused with the storm which may correspond to an extreme condition.

Consider we have an ensemble of extreme waves \( x_1, x_2, \ldots, x_n \) where \( x_1 < x_2 < \cdots < x_n \). If this ensemble is statistically independent and can be assumed to be identically distributed, then the probability distribution of the ensemble can be approximated by Generalized Pareto (GP) distribution. The independence of the extreme waves can be achieved by separating each extreme waves in a particular minimum distance. [3] separated each extreme waves with a minimum separation of 48 hours to minimize the chance of extreme waves coming from the same extreme event. The cumulative distribution function (CDF) of GPD was found to be [10]:

![Figure 1](image-url)

**Figure 1.** (a) Domain coverage of full WW3 SWH data, (b) locations of buoy measurements, (c) bathymetry (in m) and location of the domain study, and (d) location of the Bunguran Island. P1, P2, and P3 in (b) are the locations of buoys for data validation in Jepara, Karawang, and Masalembo, respectively, while A1 and A2 in (d) are the locations for return values estimation in front of and behind of the Bunguran Island.

[3] separated each extreme waves with a minimum separation of 48 hours to minimize the chance of extreme waves coming from the same extreme event. The cumulative distribution function (CDF) of GPD was found to be [10]:
\[ P(x) = \Pr\{x < x\} = F(x; k, \alpha, \xi) = \begin{cases} 1 - \left( 1 - \frac{k(x - \xi)}{\alpha} \right)^{\frac{1}{\alpha}}, & k \neq 0 \\ 1 - e^{-\frac{x-\xi}{\alpha}}, & k = 0 \end{cases} \] (4)

where \( k, \alpha, \) and \( \xi \) are shape, scale, and location parameters, successively. In case of \( k = 0 \), this distribution is also known as exponential distribution and depends only on variable \( \alpha \), while in case of \( k < 0 \) and \( k > 0 \), they depend on variable \( \alpha \) and \( k \) and refer to Pareto distribution and a special case of beta distribution, respectively [3]. The value of \( \xi \) in this study was equal to the threshold value. We will note the case of \( k \neq 0 \) as the GPD and the case of \( k = 0 \) as the exponential distribution later for convenient. Consider a return period of \( x_l > x_{\text{threshold}} \) which can be expressed as [10]:

\[ R \tau_{x_l>x_{\text{threshold}}} = \frac{\Delta \tau}{1 - P(x_l)} \] (5)

where \( \Delta \tau \) is a mean time interval between two successive extreme waves. By substituting Equation (4) to Equation (5), a return value of extreme waves with a return period of \( m \)-year \((x_m)\) can be now written as follows:

\[ x_m = \begin{cases} \xi - \frac{\alpha}{k} \left( \frac{m}{\Delta \tau} \right)^{-k} - 1, & k \neq 0 \\ \xi + \alpha \ln \frac{m}{\Delta \tau}, & k = 0 \end{cases} \] (6)

Meanwhile, the shape and parameter of the GPD have been calculated using the maximum likelihood estimation (MLE) and has been solved numerically as has been done by [8]. This method has been proven to give a good estimation for a large number of samples and a shape parameter greater than 0.2 [10]. On the other hand, the parameter of the exponential distribution could be readily derived also using MLE without numerical computation. The exponential distribution has been tested to the data using the Anderson-Darling (AD) test (see [13], for example) with a significance level of 5 %. The null hypothesis that we used in this test was “the extreme waves data is from a population with an exponential distribution” and the alternative hypothesis was “the extreme waves is not from a population with an exponential distribution”. Qualities of the fitted GPD and the exponential distribution were assessed using quantile-quantile (QQ) plot and return value plot.

4. Results

The POT method with a fixed threshold level at the quantile 0.93 has been applied to the WW3 SWH data in the SCS and the Natuna Sea. The quantile 0.93 was an appropriate threshold and it gave good results even though in some places the quantiles have to be increased up to quantile 0.97 [3]. Furthermore, seasonal distributions of extreme waves in the domain of study and influences of swell waves from the SCS to the extreme waves in the Natuna Sea at several locations have been examined. Additionally, the QQ plots and the return value plots of extreme waves were presented afterwards to provide information about future extreme waves.
The 25 years (1991-2015) maximum value plot is provided in Figure 2 to exhibit the severest extreme waves which happen from 1991-2015. The maximum values were high in the SCS and low in the Natuna Sea as tropical cyclones were frequently generated in the SCS. The figure also depicts tracks of some of the severest tropical cyclones in the SCS with values up to 13 m. On the other hand, the chosen threshold values are also displayed in Figure 2. The threshold values represent minimum values of SWHs to be considered as extreme events. Although the threshold levels in each location were the same, the resulting threshold value in each location could be different and depended on wave characteristics in each location. Since SWHs in the SCS were typically larger than in the Natuna Sea [4] [17], the threshold values would also likely to be higher in the SCS than in the Natuna Sea.

Spatial distributions of seasonal extreme waves have also been examined in the domain study. Since the SCS and the Natuna Sea are influenced by the monsoon seasons, the seasons may affect distributions of extreme waves and generations of tropical cyclones in the SCS which can travel to the Natuna Sea. [17] introduced seasonal analyses of winds and wind-waves in general. They showed that the mean values of wind speeds in the SCS and the Natuna Sea during December – February, were higher than during June – August. This led to higher waves in the former season than the latter season. Using a similar method, the seasonal occurrence, mean, and maximum extreme waves have been investigated. The seasons in this area has been divided into two; Northeast Winter Monsoon (NWM) season, which happens from November to March, and Southwest Summer Monsoon (SSM) season, from April to September. This classification was based on seasons in the SCS from [4] and also to cover the tropical cyclone seasons in the SCS based on [23].

Figure 2. Spatial distribution of (a) maximum value and (b) quantile 0.93 of SWH data (in m).
Figure 3. Frequency occurrences of extreme waves during (a) the NWM and (b) the SSM.

Figure 4. Spatial distribution of mean extreme waves (in m) during (a) the NWM and (b) the SSM.

Figure 5. Spatial distribution of maximum extreme waves (in m) during (a) the NWM and (b) the SSM.
Firstly, seasonal occurrences of extreme waves in the Natuna Sea and the SCS during NWM and SSM are presented in Figure 3. From the Figure, the extreme waves in Natuna Sea and the SCS were more likely to happen during the NWM than during SSM, though [23] showed that the tropical cyclones in the SCS were stronger and happened more often during May – September (correspond to the SSM) than during November – March (correspond to the NWM). This might roughly tell us that tropical cyclones occurrences were not well correlated to the occurrences of the extreme waves in the Natuna Sea and the SCS, but more influenced by the monsoon seasons instead. The extreme waves occurred more than 125 times in the Natuna Sea and the SCS during NWM, whereas during SSM the extreme waves occurred less than 50 and 100 times in the Natuna Sea and the SCS, consecutively.

In addition, comparisons between the mean and maximum extreme waves during NWM and SSM in the SCS and the Natuna Sea are presented in Figure 4 and 5. The mean values of extreme waves in the SCS and the Natuna Sea were slightly higher during the NWM than during the SSM. These discrepancies became more obvious in the seasonal maximum plots of extreme waves. The maximum values were much higher during the NWM than during the SSM, especially in the Natuna Sea. The values could reach 6 m during the NWM and only around 1 – 3 m during the SSM in the Natuna Sea.

As it is already discussed, the extreme waves in the Natuna Sea could be caused by swell waves from the SCS. The effects of swell waves from the SCS to the extreme waves in the Natuna Sea have been examined in front of (A1) and behind of (A2) the Bunguran Island (see Figure 1b) by bivariate and rose diagrams (Figs. 6 and 7). At A1, the incoming swell waves from the SCS would directly affect the characteristics of extreme waves in this location, while at A2, the island tended to block and deflect the swell waves from the SCS and reduced their influences to the extreme waves in this area. The bivariate diagram shows that most of the extreme waves at A1 (A2) have mean periods around $6−10$ ($3−6$) s with heights around $2−4$ ($0.5−2$) m (Figure 6). The swell waves from the SCS were not only increasing the wave height at A1 but also shifting the mean period to higher periods. Moreover, the extreme waves at A1 mainly propagated southwestward, whereas the extreme waves at A2 mainly propagated northeastward (Figure 7).

Furthermore, the QQ plot and the return value plot have been provided to assess the quality of a fitted GPD and exponential distribution at A1 and A2. The GPDs (blue circles) fitted well to the extreme wave data in both locations, but the exponential distribution (red circles) seemed not to agree well in the former location, especially for extreme waves greater than 4 m (Figure 8). The return value plots also exhibited that the GPDs (blue line) fitted well to the data in both locations. Moreover, the null hypothesis of the AD test with a significance level of 5% was accepted in both locations hence the exponential distribution was still can be applied in both locations. Furthermore, the return value plots also revealed the possibility of future extreme waves in both locations. At A1, the estimated return values by the GPDs were 4.70, 4.87, and 4.96 m and at A2 were 2.08, 2.20, and 2.27 m for return periods of 25, 50, and 75-year, successively, whereas the exponential distributions tended to produce higher return values for return period greater than 3 years at A1, and slightly higher values over the GPD at A2.
Figure 6. Bivariate diagram of extreme waves (a) behind of and (b) in front of the Bunguran Island. Colors represent the frequency or probability occurrence of extreme waves (in percent).

Figure 7. Wave rose of extreme waves (a) behind of and (b) in front of the Bunguran Island. Percentage (circle) represents frequency of occurrence of extreme waves and N, E, S, and W indicate North, East, South, and West, respectively.
5. Conclusions

Extreme wave height analysis using peak-over threshold (POT) method has been conducted in the Natuna Sea. The WW3 SWH data with finer spatial resolution than the simulated SWH data from [3] and [21] was expected to provide more accurate results. The return values of extreme waves have also been provided to predict future extreme waves. We have also presented the characteristics of the extreme waves in term of direction and mean period of the extreme waves due to the influences of the swell waves from the SCS. From the results, we can conclude that the extreme waves were higher and occur more often during NWM than during SSM in the SCS and the Natuna Sea. The swell waves from the SCS affected the extreme waves in the Natuna Sea by increasing the heights and mean periods of the extreme waves. Moreover, the calculated return values of extreme waves using the GPD were higher in the front of the Bunguran Island than behind the Bunguran Island. Although the AD test was accepted the exponential distribution in both locations, the QQ plot and the return value plot has shown that the GPDs fitted better to the extreme wave data in both locations. Based on the GPDs, the resulting return values in front of (behind of) the Bunguran
Island are 4.7, 4.87, 4.96 (2.08, 2.20, and 2.27) m for return periods of 25, 50, and 75-year, respectively. Further study of the threshold level selection can be conducted since it is not covered in this study. Moreover, the use of another method to estimate the shape and scale parameters and its produced return value is also intriguing and can be conducted for comparison to the MLE.

6. Acknowledgement

Parts of this research were funded by the Ministry of Research, Technology, and Higher Education (Kemenristekdikti) of the Republic of Indonesia under Basic Research Grant 2019. The authors would like to thank the supports of Kemenristekdikti. The authors also would like to thank Dr. Ibnu Sofian from the Geospatial Information Agency (BIG) for providing the WW3 SWH data.

References

[1] Abild, J., E. Y. Andersen and D. Rosbjerg 1992 The Climate of Extreme Winds at the Great Belt J. of Wind Engineering and Industrial Aerodynamics 41-44 521-532
[2] Anggara, P. D., T. M. Alam, D. Adrianto and W. S. Pranowo 2018 The Wave Characteristics in Natuna Sea and Its Adjacent for Naval Operation Base Purposes IOP Conf. Series: Earth and Environmental Science 176
[3] Caires, S. and A. Sterl 2005 100-Year Return Value Estimates for Ocean Wind Speed And Significant Wave Height From The ERA-40 Data J. of Climate 18(7) 1032-1048
[4] Chu, P. C., Y. Qi, Y. Chen, P. Shi and Q. Mao 2004 South China Sea Wind-Wave Characteristics. Part I: Validation of Wavewatch-III Using Topex/Poseidon Data J. of Atmospheric and Oceanic Technology 21 1718-1733
[5] Ferreira, J. A. and C. G. Soares 1998 An Application of the Peaks Over Threshold Method to Predict Extremes of Significant Wave Height J. of Offshore Mechanics and Arctic Engineering 120 165-176
[6] Fischer, S. and A. Schumann 2014 Comparison Between Classical Annual Maxima and Peak Over Threshold Approach Concerning Robustness SFB 823 26
[7] Goda, Y., M. Kudaka and H. Kawai 2010 Incorporation of Weibull Distribution in L-Moments Method for Regional Frequency of Peaks-Over-Threshold Wave Heights Coastal Engineering 2010 (Proc. 32th International Conference, Shanghai, ASCE) Waves 62 1-11
[8] Grimshaw, S.D. 1993 Computing Maximum Likelihood Estimates for the Generalized Pareto Distribution Technometrics 29(1) 185-191
[9] Holthuijsen, L. H. 2007 Waves in Oceanic and Coastal Waters Cambridge University of Press Cambridge
[10] Hosking, J. R. M., J. R. Wallis and E. F. Wood 1985 Parameter and Quantile Estimation for the Generalized Pareto Distribution Technometrics 29(3) 339-349
[11] Huangfu, J., R. Huang and W. Chen 2017 Relationship between the South China Sea Summer Monsoon Onset and Tropical Cyclone Genesis over the Western North Pacific International Journal of Climatology 1-5
[12] Idris, K. and D. A. Gammaranti 2018 Assessment of Wave Energy Resources in the Vicinity of Natuna Islands International Journal of GEOMATE 52 137-145
[13] Laio, F. 2004 Cramer-von Mises and Anderson-Darling Goodness of Fit Tests for Extreme Value Distributions with Unknown Parameter Water Resources Research 40
[14] Lee, T., C. Y. Leung, M. Kok and H. Chan 2012 The long Term Variations of Tropical Cyclone Activity in the South China Sea and the Vicinity of Hongkong Tropical Cyclone Research and Review 1(3) 277-292
[15] Purwanta, W. 2002 Pengolahan Raw Data Parameter Perairan Laut Menjadi Usable Data Dalam Upaya Diseminasi Data Kelautan Seawatch Indonesia Jurnal Teknologi Lingkungan 4(1) 19-20 (in Indonesian)

[16] Rizal, A. M., N. S. Ningsih, I. Sofian, F. Hanifah and I. Hilmi 2019 Preliminary Study of Wave Energy Resource Assessment and Its Seasonal Variation along the Southern Coasts of Java, Bali, and Nusa Tenggara Waters J. Renewable Sustainable Energy 11 1-13

[17] Semedo, A., K. Suselj, A. Rutgersson and A. Sterl 2011 A Global View on the Wind Sea and Swell Climate and Variability from ERA-40 J. of Climate 24.

[18] Smith, C. B. 2007 Extreme Waves and Ship Design Proceeding of the 10th Symposium on Practical Design of Ships and Other Floating Structures 1-8

[19] Statistics Indonesia https://natunakab.bps.go.id/ last accessed 27 July 2019.

[20] Teena, N. V., V. S. Kumar, K. Sudheesh and R. Sajeev 2012 Statistical Analysis on Extreme Wave Height. Natural Hazards 64(1) 223-236

[21] Timmermans, B., D. Stone, M. Wehne’r and H. Krishnan 2017 Impact of Tropical Cyclones on Modelled Extreme Wind-Wave Climate Geophysical Research Letters 44 1393–1401

[22] VOA: Remote Military Outpost will Help Indonesia Resist Chinese Maritime Expansion https://www.voanews.com/east-asia/remote-military-outpost-will-help-indonesia-resist-chinese-maritime-expansion last accessed 27 July 2019

[23] Wang, G., J. Sue, Y. Ding and D. Chen 2006 Tropical Cyclone Genesis over the South China Sea J. of Marine Sciences 68 318-326