Seasonal and Non-Seasonal Generalized Pareto Distribution to Estimate Extreme Significant Wave Height in The Banda Sea

by Jusup Suprijanto

Submission date: 16-May-2022 09:21PM (UTC+0700)
Submission ID: 1837580068
File name: to_Estimate_Extreme_Significant_Wave_Height_in_The_Banda_Sea.pdf (1.01M)
Word count: 2868
Character count: 13816

IOP Conference Series: Earth and Environmental Science

PAPER • OPEN ACCESS

Seasonal and Non-Seasonal Generalized Pareto Distribution to Estimate Extreme Significant Wave Height in The Banda Sea

3rd International Conference on Tropical and Coastal Region Eco Development 2017 IOP Publishing
IOP Conf. Series: Earth and Environmental Science **116** (2018) 012066 doi:10.1088/1755-1315/116/1/012066

Seasonal and Non-Seasonal Generalized Pareto Distribution to Estimate Extreme Significant Wave Height in The Banda Sea

Nursamsiah[1,4], Denny Nugroho Sugianto[2,3], Jusup Suprijanto[1], Munasik[1], and Bambang Yulianto[1]

[1] Faculty of Fisheries and Marine Science, Diponegoro University, Jl. Prof. H. Soedarto, SH, UNDIP, Tembalang, Semarang 50275, Indonesia
[2]Center for Coastal Disaster Mitigation and Rehabilitation Studies, Diponegoro University, Jl. Prof. H. Soedarto, SH, UNDIP, Tembalang, Semarang 50275, Indonesia
[3] Departement of Oceanography, Fisheries and Marine Science Faculty, Diponegoro University
Jl. Prof. H. Soedarto, SH, UNDIP, Tembalang, Semarang 50275, Indonesia
[4]Indonesian Agency for Meteorology, Climatology and Geophysics Jl. Angkasa I No.2, Kemayoran, Jakarta10610, Indonesia
nursamsiah533@gmail.com

Abstract. The information of extreme wave height return level was required for maritime planning and management. The recommendation methods in analyzing extreme wave were better distributed by Generalized Pareto Distribution (GPD). Seasonal variation was often considered in the extreme wave model. This research aims to identify the best model of GPD by considering a seasonal variation of the extreme wave. By using percentile 95 % as the threshold of extreme significant wave height, the seasonal GPD and non-seasonal GPD fitted. The Kolmogorov-Smirnov test was applied to identify the goodness of fit of the GPD model. The return value from seasonal and non-seasonal GPD was compared with the definition of return value as criteria. The Kolmogorov-Smirnov test result shows that GPD fits data very well both seasonal and non-seasonal model. The seasonal return value gives better information about the wave height characteristics.

1. Introduction

Information on extreme waves is necessary for marine planning, managing and evaluating. The Occurrence of high waves may disturb the transportation activities [1] and destruct the conservation area of mangroves and coral reefs [2]. Maritime structures in the coastal and offshore region must be designed for an extreme condition such as high waves [3]. The structures designed to stand in extreme wave condition according to expected structure lifespan. The information of maximum waves that may occur in certain periods (return periods) such as 5, 20 or maybe 100 years usually needed for marine planning.

In analyzing extreme wave event, Peaks Over Threshold (POT) method was recommended in the IAHR Working Group on Extreme Wave Analysis [4]. POT uses data that exceeds the threshold as an extreme value. POT was naturally described using Generalized Pareto Distribution (GPD [5], then this method was used widely to predict extreme waves [6],[7],[8].

IOP Conf. Series: Earth and Environmental Science **116** (2018) 012066 doi:10.1088/1755-1315/116/1/012066

Wave seasonality or variations that repeat periodically (weekly, monthly, yearly, etc.) were often examined carefully in modeling extreme waves [9] but other researchers only consider the location of the site [6][10]. The seasonal and nonseasonal extreme wave model uses another model (GEV) gave different result according to [11]. The primary objective of this research is to identify the difference between seasonal and nonseasonal GPD model in estimating extreme wave return value for maritime planning and management.

In the Banda Sea, there were inter-island transportation and fishing activity. Good planning and management required to support this activity. Information on extreme waves needs to be estimated to determine the risk of disasters that could interfere with such activities. The wave characteristics in the Banda Sea has associated with the monsoon and vary each month with the highest peak occurring in the Australia monsoon period or in June-July-August (JJA)[12].

2. Research Methods

2.1 Data

This research used significant wave height (Hs) data in the south of Buru Island (4° 30' 0" S, 126° 7' 30" E), provided by BIG (Indonesian Geospatial Information Agency) based on Consortium of Oceanic and Atmospheric Prediction (COAP). The data were obtained from **http://tides.big.go.id/** (access on January 14, 2017). To fit the GPD we used daily maximum Hs from 1991 – 2010, while data from 2011 - 2014 was used as validation data of extreme return period results. The prediction was performed for 2, 5, 10, 25, 50 and 100 years.

Figure 1. Study location

2.2 Peaks Over threshold

The POT method employed a series of Hs above a defined significant wave height level or threshold level. If threshold value (u) is too low then the exceeded data will produce a biased estimator. On the other hand, if the selected value u was too high then there will not be enough data to fit the model, resulting in large variations. One method of determining the frequently used threshold value was the percentile method. This percentile method was easier and practical, but the resulting threshold determination is accurate [13]. The 95% percentile will be used in this study as an extreme Hs limit. The 95% percentile was also used by [14] and generates the GPD model corresponding to the Hs data used.

2.3 Generalized Pareto Distribution

In general, the parameters of GPD are known as σ scale parameter (σ > 0), *k* shape parameter ($k \in \mathbb{R}$) dan μ location parameter($\mu \in \mathbb{R}$). For threshold *u*, convergent on GPDs that have a cumulative distribution function (CDF) as follows:

$$F(x) = \begin{cases} 1 - \left(1 - k\frac{x-\mu}{\sigma}\right)^{1/k} & k \neq 0 \\ 1 - exp\left(-\frac{x-\mu}{\sigma}\right) & k = 0 \end{cases} \quad (1)$$

And probability function (pdf) of GPD :

$$F(x) = \begin{cases} 1 - \left(1 - k\frac{x-\mu}{\sigma}\right)^{1/k} & k \neq 0 \\ 1 - exp\left(-\frac{x-\mu}{\sigma}\right) & k = 0 \end{cases} \quad (2)$$

The parameter estimation was using the L-moment method in Easyfit software.

Non-seasonal GPD uses all data that exceed the threshold value. The seasonal GPD was created by splitting the model by its seasonal variation in this study we use monthly variation, reference to the study about Indonesian wave characteristics by [12]. The seasonal GPD has twelve model represent monthly variation in the study site. The determination of threshold values and parameter estimation using the same method with non-seasonal GPD model.

2.4 Goodness of fit test

2.4.1. Quantile-Quantile Plot (QQ plot). By plotting empirical and fitted distributions for Hs above the thresholds *u* against each other, we can see the consistency of the estimated parameter value of selected *u*. If the QQ plot follows a linear line then the distribution is fitted.

2.4.2. Kolmogorov-Smirnov test. The hypothesis used in Kolmogorov-Smirnov test are :

H_0 : Fn(x) = F0(x) (data follow Generalized Pareto Distribution)
H_1 : Fn(x) ≠ F0(x) (data doesn't follow Generalized Pareto Distribution)

H_0 rejected if $D > D_{1-\alpha/2}$, using a 95% confidence level (α = 5%) with a value of D derived from :

$$D = \sup_x \left| F_n(x) - F_0(x) \right| \quad (3)$$

2.5 Return Value

The return value (x_m) of the extreme wave height exceeding the threshold (*u*) at least once in the *m* observation is as follows:

$$x_m = u + \frac{\sigma}{k}\left[(m\zeta_u)^k - 1\right] \quad (4)$$

ζ_u Is the probability of events that exceed the threshold u.

Return value in *N*-years dan n_y Is the amount of data in a year then$m = N \times n_y$. The equation of *N*-years return value as follows :

$$x_N = u + \frac{\sigma}{k}\left[\left(Nn_y\zeta_u\right)^k - 1\right] \quad (5)$$

2.6 Model Validation

Validation was performed by comparing the return value with maximum data during the period of 1, 2, 3 and 4 years based on Hs data for the period 2011-2014. The criteria were made based on the

IOP Conf. Series: Earth and Environmental Science **116** (2018) 012066 doi:10.1088/1755-1315/116/1/012066

definition of return value that is the maximum Hs that occurs at least once during the return period. The return value greater than the maximum Hs was identified as incorrect and for smaller repeat periods of maximum Hs per period was identify as correct. The difference between the maximum Hs and the return period is also considered in model validation.

3. Result and Discussion

The threshold value specified using 95% percentile. Data greater than the threshold value are considered to be extreme values and are used to estimate GPD parameters. The number of extreme data used to estimate nonseasonal GPD parameter was 365 data with threshold value 1.76 meters. The threshold in seasonal GPD varies with the lowest threshold in November (0.74 m) and the highest in July (2.15 m). At the peak of monsoon Australia (JJA) the threshold value is higher than in other months, it means high waves occur in those months. The result of parameter estimation for nonseasonal and seasonal GPD using the L-moment method in Easyfit software was shown in Table 1.

Table 1. Non-seasonal and seasonal GPD parameter

Parameter	Non seasonal	Monthly variation (Seasonal)											
		Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Shape (k)	-0,14	0,03	-0,14	-0,28	-0,23	-0,47	-0,15	0,01	-0,14	-0,39	-0,51	0,10	-0,03
Scale (σ)	0,28	0,23	0,27	0,48	0,15	0,32	0,25	0,21	0,21	0,26	0,22	0,15	0,40
Location(μ)	1,77	1,27	1,21	1,14	1,21	1,66	1,40	2,15	2,01	1,49	1,06	0,72	1,02
Threshold(u)	1,76	1,29	1,25	1,09	1,21	1,69	2,12	2,15	2,01	1,52	1,09	0,74	1,05
Number of extreme Hs	365	31	28	31	30	31	30	31	31	30	31	30	31

The deviations in non-seasonal GPD (Figure 2) look significant at a high value of Hs but significant deviations also occur in low-value of Hs in seasonal GPD. QQ plot for seasonal GPD (Figure 3) seen to be around diagonal lines, but it can be seen that considerable deviations on seasonal models in certain months. These results can be taken into consideration in the selection of the best models. Generally, the QQ plots for non-seasonal and seasonal GPD are spreading close to diagonal lines.
Kolmogorov-Smirnov goodness of fit test shows that the value of $D_{1-\alpha/2}$ greater than D value (Table 2). It means that extreme Significant wave height follows GPD with confidence level 95%. According to QQ plot and Kolmogorov-Smirnov test, the non-seasonal and seasonal GPD fit the extreme significant wave height in the location.

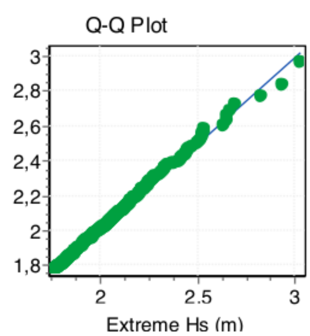
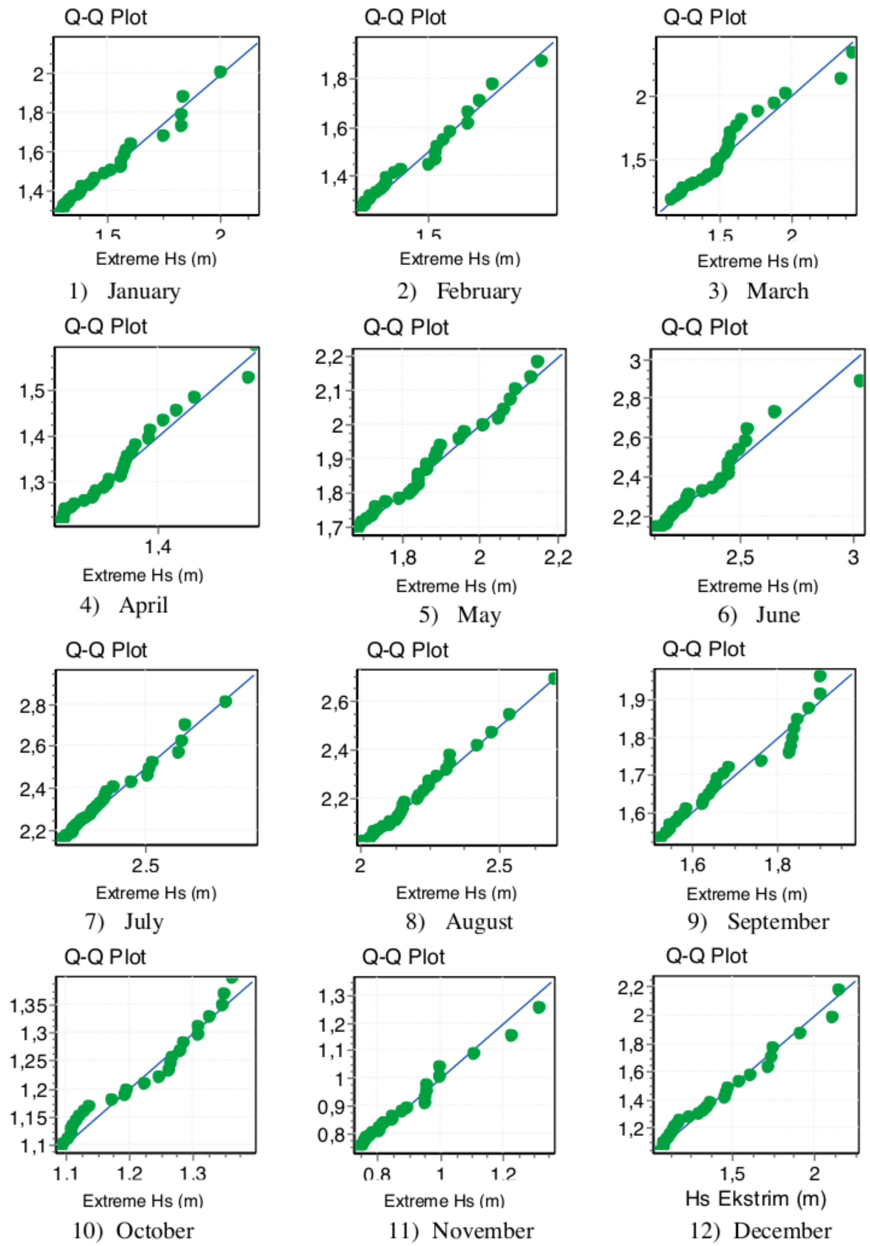


Figure 2. QQ plot for non-seasonal GPD



1) January 2) February 3) March 4) April 5) May 6) June 7) July 8) August 9) September 10) October 11) November 12) December

Figure 3. QQ plot for seasonal GPD

Table 2. Kolmogorov-Smirnov goodness of fit test

Parameter	Non seasonal	Monthly variation (Seasonal)											
		Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
D	0,03	0,10	0,13	0,12	0,09	0,09	0,09	0,07	0,09	0,13	0,14	0,11	0,12
$D_{1-\alpha/2}$	0,07	0,24	0,24	0,25	0,24	0,24	0,24	0,24	0,24	0,24	0,24	0,24	0,24

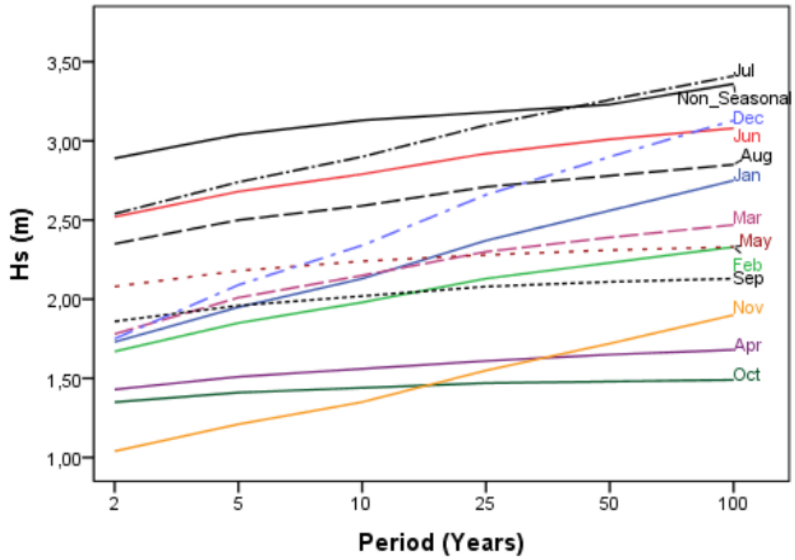


Figure 4. Non seasonal and monthly(seasonal) Return value.

Tabel 3. Model Validation by comparing Hs Return value form nonseasonal and seasonal GPD with maximum Hs. The bold number identify as incorrect.

Period (Years)	Hs Return Value		Maximum Hs (2011-2014)
	Non Seasonal	Seasonal	
1	**2,55**	2,38	2,38
2	2,66	2,54	2,72
3	2,72	2,63	2,72
4	2,76	2,69	2,93

Seasonal return value gives varying value each month while nonseasonal return value only has one value. These monthly variation of seasonal return value gives better information about the characteristics of the study location. The model has similar result with other model [11] in variation but has different in another characteristics. The maximum seasonal return value resulting a lower value in the initial period and increases steadily exceed the non-seasonal return values. The difference between nonseasonal and maximum seasonal return value was very small (less than 1 m). Return values which have positive k (Jan, Jun, and Nov) gave more increases return value in each period than negative k return value.

On comparing with maximum Hs in study location, non-seasonal and maximum seasonal return value in the early period was slightly different (< 0,2 m). The nonseasonal return value has 1 incorrect value for 1 year period, but in 2 – 4 years period Nonseasonal return value closer to maximum Hs than seasonal Hs.

4. Conclusion

The return value from seasonal GPD and nonseasonal GPD was compared to examine the difference. Monthly variation was selected as seasonal variation in study location. From the goodness of fit test, the GPD was fitted to the extreme significant wave height data both seasonal and nonseasonal. The return values from seasonal GPD can be obtained for any month while nonseasonal return value has

one value for each period. Seasonal return value could give better information for each month or season and anticipate the extreme wave height for maritime planning and management.
Model validation show all correct value for the seasonal return value and incorrect return value for the nonseasonal model in 1 years period but closer value for next period. A longer period of return level or other sample site needed to verify consistency of model validation.

References

[1] Sugianto DN, Zainuri M, Darari A, Darsono S, and Yuwono N. 2017. *Int. J. Civ. Eng. Technol.*, 8 (5) :. 604–619.
[2] World Bank. 2016. "Managing Coasts with Natural Solutions," p. 167.
[3] Jonathan P and Ewans K, 2012. *Ocean Eng.*, 62: 91–109.
[4] Mathiesen M, Goda Y, Hawkes PJ, Mansard E, Martín MJ, Peltier E, Thompson EF, and Vledder GV. 1994. *J. Hydraul. Res.*, 32 (6) : 803–814.
[5] Coles S. 2001. *Springer*.
[6] Mazas F, and Hamm L. 2011. *Coast. Eng.*, 58 (5) : 385–394.
[7] You ZJ. 2012. *Coast. Eng.*, 61 (1) : 49–52.
[8] Far SS, Khairi A, and Wahab A. 2016. *Ocean Sci. Discuss.*, pp. 1–25.
[9] Randell D, Feld G, Ewans K, and Jonathan P. 2015. *Environmetrics*, 26: 442–450.
[10] Solari S, and Losada MÁ. 2012. *Coast. Eng.*, 68: 67–77.
[11] Mendez FJ, Menendez M, Luceno A, Medina R, and Graham NE. 2008. *Ocean Eng.*, 35: 131–138.
[12] Kurniawan R, Habibie MN, and Suratno. 2012. *J. Meteorol. Dan Geofis.*, 12 (3): 221–232.
[13] Estiningrum T, Rusgiyono A, and Wilandari Y. 2015. *J. Gaussian*, 4: 141–150.
[14] Agarwal A, Venugopal V, and Harrison GP. 2013. *Renew. Sustain. Energy Rev.*, 27: 244–257.

Seasonal and Non-Seasonal Generalized Pareto Distribution to Estimate Extreme Significant Wave Height in The Banda Sea

ORIGINALITY REPORT

4%	%	4%	%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Zhuxiao Shao, Bingchen Liang, Huijun Gao. "Extracting independent and identically distributed samples from time series significant wave heights in the Yellow Sea", Coastal Engineering, 2020 Publication	1%
2	G.G.L. Nashed, Emmanuel N. Saridakis. "Stability of motion and thermodynamics in charged black holes in f(T) gravity", Journal of Cosmology and Astroparticle Physics, 2022 Publication	1%
3	Mazas, Franck, Philippe Garat, and Luc Hamm. "Questioning MLE for the estimation of environmental extreme distributions", Ocean Engineering, 2014. Publication	1%
4	Chandra R. Rupa, P.P. Mujumdar. "Modeling High-Intensity Precipitation for Urban Hydrologic Designs", Elsevier BV, 2019 Publication	<1%

5	T. Muhammed Naseef, V. Sanil Kumar. "Variations in return value estimate of ocean surface waves – a study based on measured buoy data and ERA-Interim reanalysis data", Natural Hazards and Earth System Sciences, 2017 Publication	<1%
6	Erik Vanem. "Bayesian Hierarchical Space-Time Models with Application to Significant Wave Height", Springer Science and Business Media LLC, 2013 Publication	<1%
7	T. Muhammed Naseef, V. Sanil Kumar, Jossia Joseph, B. K. Jena. "Uncertainties of the 50-year wave height estimation using generalized extreme value and generalized Pareto distributions in the Indian Shelf seas", Natural Hazards, 2019 Publication	<1%
8	Lorenzo Annulli, Vitor Cardoso, Leonardo Gualtieri. "Applications of the close-limit approximation: horizonless compact objects and scalar fields", Classical and Quantum Gravity, 2022 Publication	<1%