



Evaluation of Peaks-Over-Threshold Method

Soheil Saeed Far¹ and Ahmad Khairi Abd. Wahab^{1,2}

¹Faculty of Civil Engineering, Universiti Teknologi Malaysia (UTM), 81310, Skudai, Johor Bahru, Malaysia.

²Center for Coastal & Ocean Engineering, Research Institute for Sustainable Environment, UTM, Jalan Sultan Yahya Petra, 54100, Kuala Lumpur, Malaysia.

Correspondence to: Ahmad Khairi Abd. Wahab (Email: akhairi@utm.my)

Abstract. Two extreme wave analysis models, namely Peaks-Over-Threshold (POT) and Generalized Pareto Distribution (GPD), were developed in order to improve the POT model and highlight merits and limitations of the two models. Studies have shown that the POT model was not equipped with a suitable approach to determine a true threshold value. This paper proposed an approach to specify the most suitable threshold value for the POT model, which is called Hybrid method. In 5 addition, until now the MIR (minimum ratio of residual correlation coefficient) criterion has been used as a goodness-of-fit method in the POT model. However, the examinations on the method represented that MIR is not always a stable approach in determining true distribution function. This paper proposed an alternative approach instead of the MIR criterion method, it is called Norm of Residuals, and its credibility was examined by the Chi-Square test. The results drawn from this study also demonstrated that the Hybrid method completely matched with the POT model, and the threshold obtained by this method is 10 credible, moreover, the Norm of Residuals method is completely stable in determining the best fitting distribution for the POT model.

Keywords: POT (Goda) model, Generalized Pareto Distribution (GPD) model, threshold value, Chi-Square test.

1 Introduction

Currently, many types of structures are designed for different purposes in coastal and offshore regions. The force of storm waves usually influences the life span of these structures, compelling engineers to design structures to withstand these destructive waves. Therefore, the estimation of extreme waves is important in the design phase of structures in the marine areas. However, prior to the phase of designing, the probability of storm waves during the shelf life of structures should be projected. The estimation of a probable wave height is called return period of storms, and the employed method is generally known as extreme value analysis. Fundamentally, the aim of extreme value analysis is to specify the probabilities of exceedance and 20 non-exceedance of wave heights, and non-exceedance probabilities are known as return periods.

There are several methods to estimate extreme values; two commonly used models, namely Peaks-Over-Threshold (POT) and Generalized Pareto Distribution (GPD) were developed in this study.

One crucial step in developing the POT model is to select the best fitting distribution function for dataset. To determine a best fitting distribution, two methods were introduced by Goda (1988); namely correlation coefficient method, and the MIR



(minimum ratio of residual correlation coefficient) criterion method. However, examinations represented that the MIR criterion method has shown some instabilities in determining correct results when it comes to choosing a true distribution function. This study proposed an alternative approach instead of the MIR criterion method, it is called norm of residual (N.o.R).

Selecting an appropriate threshold value is important because the estimated extreme events by the POT model are sensitive 5 to the changes in threshold value. Goda (2000) pointed out that in selecting a suitable threshold, users should refer to previous studies and engineering judgment. Li et al. (2012) followed the same procedure to determine a true threshold value. The authors provided a list of several extreme analysis studies, which had been conducted on the same data, and came to a conclusion with one of the thresholds used by one of those studies. However, in absence of a reliable threshold method, the other authors in the list of Li et al. (2012) could claim of the credibility of their threshold values. Moreover, if there is no preceding study for a 10 particular data, engineering judgment alone may not be sufficient when selecting a suitable threshold. At present, there is still no appropriate approach for selecting a threshold value for the POT model. A hybrid method consists of a two-part procedure; a mean residual life (MRL) plot, and a graph of several thresholds, was proposed in this paper to specify a suitable threshold value for the POT model.

Based on the definition by Grimshaw (1993), the GPD is a two-parameter family of distribution, which is applied to a data 15 model exceeding a known threshold value. The model was introduced for the first time by Pickands (1975) and is known to be a stable distribution for exceedances beyond a threshold. The model has been presented by many authors; such as, Davison (1984), Smith (1985), Van Montfort & Witter (1985), Hosking & Wallis (1987), and Coles (2001).

The data that was applied in this study comprised a set of wave heights gathered from shipboard observations over a 41-year period from 1949 to 1989 in the South China Sea. Since 1989, the process of data acquisition was stopped, and the modern 20 observation was replaced by satellites. In this study, only the observed data during the 41-year period has been used, due to the focus of this study on the merits and limitations of the two models, and improving the POT model. Moreover, there is a previous study on the same data, which could be referenced and compared with the new findings.

The data were collected within the Marsden Square Numbers 2554, 2555, 2564 and 2565 (the rectangular area within the magnified section of Fig. (1)) in the Federal Territory of Labuan off the coast of Sabah, Malaysia. The Labuan metropolitan 25 coast is located at the South China Sea on the continental shelf of Malaysia. National Climatic Data Center, USA compiled the dataset, which is known as Labuan data. The Labuan data are available from the wave database of the Drainage and Irrigation Department, Malaysia.

The Labuan data are a set of wave height records in which the time intervals between the successive data are considered to be at least three days to secure the required independence of the data. To fill the large gaps between the successive data; e.g., 30 more than 4 days, the linear interpolation method was used since the data varied relatively slowly. A time interval of two to four days between successive peaks was recommended by Mathiesen et al. (1994) to maintain the condition of independency.



Figure 1. Location of Federal Territory of Labuan, and the surrounding offshore area.

The northeast monsoon waves (same direction with the smaller arrow in the magnified part of the map in (Fig. 1)) were the predominant waves among the wave heights for the four seasons in the region. About 65 percent of the waves during this season approached from between 30 to 60 degrees (Syed Abdullah, 1992). Therefore, to fulfill the requirement for homogeneity in the data, only storms in the compass direction of 30 to 60 degrees (northeast monsoon waves) were employed in the analysis.

5 The purpose of the foregoing tasks were to obtain the conditions of independent and identically distributed (iid) measurement for the Labuan data. The Labuan data was evaluated and used for the first time by Coastal and Offshore Engineering Institute, Universiti Teknologi Malaysia. (2009) as a report on extreme wave analysis for the Marine Department of Malaysia. The results of the report (Table 1) represent the estimated extreme wave heights for several return periods. The authors employed only one distribution function (Gumbel distribution) in their analysis without considering any threshold value.

Table 1. Extreme wave heights estimated by Gumbel distribution and the Labuan data (UTM Archive).

Return Periods (yr)	2	5	10	20	30	50	100
Significant Wave heights (m)	3.08	3.57	3.89	4.20	4.38	4.60	4.90

10 The following objectives were discussed in this paper: i) developing the proposed Hybrid method to specify true threshold value for the data for the POT model; ii) conducting the Norm of Residual method for the POT model to select the best fitting distribution function for the dataset; and iii) developing the modified POT (POT with the two proposed methods by this study) and GPD models, and comparing their results to evaluate the proposed methods and highlight merits and limitations of the two models.



This paper has been divided into four sections. Following this introduction, Section 2 provides summaries about the two models, and proposed methods for the POT model. Discussion about the return values and figures for both models, the goodness-of-fit tests, and the proposed methods are presented in Section 3. Finally, Section 4 contains the conclusions drawn from this study.

5 2 Methodology

2.1 Generalized Pareto Distribution Model

Generalized Pareto Distribution model could be fitted for the observed wave height data with considering the following assumptions: The data are considered as exceedances from a specific threshold, which are a sequence of independent and identically distributed (iid) measurements, $x_{(1)}, \dots, x_{(k)}$. If exceeded data over a particular threshold u are defined as $y_{(j)} = x_{(j)} - u$ for 10 $j = 1, \dots, k$, therefore, by definition, a random variable of $y_{(j)}$ may be regarded as independent realization whose distribution is approximated by one of the members of the Generalized Pareto family (Coles, 2001).

The model has been evaluated and used by many authors: Smith (1984) developed and presented the model in a form of comparison method with a statistical model, Van Montfort & Witter (1985) embedded exponential distribution in GPD to check the adequacy of the exponential distribution for a set of data, Hosking & Wallis (1987) studied about several estimators for the 15 GPD model, Moharram et al. (1993) estimated the river flood by the GPD model, Castillo & Hadi (1997) used the model for estimating wave height, Pandey et al. (2001) studied extreme wave velocity by the model, Coles et al. (2003) employed extreme analysis for estimating rainfall, Pandey et al. (2004) used the model to evaluate the sea level, Öztekin (2005) in estimating river discharge, Jagger & Elsner (2006) in studying hurricanes, and Moisello (2007) employed the model to estimate the extreme rainfall.

20 The generalized Pareto distribution function has the following formula,

$$F(x) = 1 - \left(1 - \frac{\xi x}{\sigma}\right)^{1/\xi} \quad (1)$$

In general, the two parameters of GPD are known as: the scale parameter σ ($\sigma > 0$), and the shape parameter ξ ($-\infty < \xi < \infty$). The shape parameter, ξ has remarkable effects of specifying qualitative behavior of Generalized Pareto distribution. In this paper, the shape parameter is considered as $\xi \neq 0$. The case, $\xi = 0$ considered as the exponential distribution which is not 25 discussed here.

2.1.1 Threshold Value

First Technique: With regard to the above assumptions, specifying a suitable threshold is the first step of developing the model. Two methods are available to this aim: first, mean residual life (MRL) plot is applied as an exploratory technique, which is carried out prior to the model estimation (Fig. 2). The method is described as a sequence of $x_{(1)}, \dots, x_{(n_u)}$, which are



considered as $n_{(u)}$ observations, exceeded over a threshold u , and x_{max} is the greatest value of x_i . Therefore, Eq. (2) is termed as the MRL formula, and denoted as,

$$\left\{ \left(u, \frac{1}{n_u} \sum_{i=1}^{n_u} (x_i - u) \right) : u < x_{max} \right\} \quad (2)$$

Second Technique of Determining Threshold Value: The second method to determine threshold value was introduced as the complementary technique to evaluate the stability of parameter estimation (Coles, 2001). In general, the method of work is to fit GPD for a range of thresholds, and searching for the stability of parameter estimates. Therefore, if GPD is a reasonable model for data exceeding a threshold u , then, the model should be reasonable for higher threshold u_1 . The argument suggests two graphs of shape and modified scale against different threshold values with confidence intervals for each of these quantities. The two graphs are called probability and quantile plots. The technique is carried out after the model estimation (Figs. 4,5).

10 2.1.2 Determining The Parameters

Once the threshold value has been specified, estimating the shape and scale parameters would be the next step of developing the model. Based on the size of the data, different parameters estimators methods can be applied to estimate the GPD parameters. Mackay et al. (2011) conducted a comparison of several estimators by Monte Carlo simulation, and argued that likelihood-moment (LM) estimators is the most substantial method with the lowest bias and variance. Zhang & Stephens (2009) proposed an estimation procedure, and called it modified likelihood moment estimators (LME). Dupuis & Tsao (1998) proposed a hybrid estimators of two well-known estimators: method of moments (MOM), and probability weighted moments (PWM). The method was derived by incorporating a simple auxiliary limitation of the estimates.

In addition, de Zea Bermudez & Kotz (2010) carried out an extensive study on several types of methods for estimating the GPD parameters. The authors argued that the success of GPD on a set of data depends substantially on the process of parameter estimate, and maximum likelihood (ML) estimators has been the most popular method among other estimators. However, the ML estimators exists only for the shape parameter $\xi \leq 1$ because for $\xi > 1$ the log-likelihood becomes infinity. In this study, the maximum likelihood (ML) estimators was employed to estimate the parameters. The log-likelihood formula for $\xi \neq 0$ was derived from Eq. (1) as follows,

$$\ell(\sigma, \xi) = -n \log \sigma + \left(\frac{1}{\xi} - 1 \right) \sum_{i=1}^n \log \left(1 - \frac{\xi x_i}{\sigma} \right) \quad (3)$$

25 As foregoing discussion, numerical techniques are required to estimate the shape and scale parameters.

2.1.3 Return Level and Confidence Intervals

Subsequent to the estimation of the shape and scale parameters, the return levels should be estimated. Coles (2001) argued that the interpretation of extreme value models in terms of return levels are usually more convenient than using individual parameter



values. Therefore, we assume that GPD with defined shape and scale parameters is an appropriate model for the exceeded data over a threshold u to estimate m-observation return level, x_m ,

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

where σ and ξ are the scale and shape parameters, respectively. And ζ_u is described as $\zeta_u = Pr\{x_i > u\}$, in which x_i stands for the wave heights, and u is the threshold value. Consequently, the return level x_m , which is exceeded on average once every m observations being the solution of Eq. (4). Also, the parameters m and ζ_u can be evaluated by Eqs. (5, 6) as follows,

$$m = ARI \frac{N_t}{k} \quad (5)$$

in which ARI and N_t represent, respectively, the Average Recurrence Interval (return period), and the total number of data during a period of k years. Equation (6) can be employed to compute ζ_u ,

$$10 \quad \zeta_u = \frac{N}{N_t} \quad (6)$$

By substituting Eqs. (5, 6) into Eq. (4), x_m can be computed as,

$$x_m = u + \frac{\sigma}{\xi} [(ARI \cdot \frac{N}{N_t})^\xi - 1] \quad for (\xi \neq 0) \quad (7)$$

where N is the number of events exceeding a threshold u . Subsequently, the confidence interval for the return graph should be computed. There are several methods to derive confidence intervals for x_m , variance-covariance matrix, and delta methods are reliable methods (Davison & Smith, 1990; Coles, 2001; Pickands, 1975).

To check the quality of the modeling, probability plot, return level plot and quantile plot are usually evaluated (Coles, 2001).

2.2 Peaks-Over-Threshold Model (Goda Method)

The Peaks-Over-Threshold (POT) model was introduced by Goda (1988), and has been developed in this study for further evaluation. The model can be developed with a long length dataset to obtain the results in a smaller range of confidence intervals (Goda, 2010). Following is a summary of developing the model which were described by Goda (1988, 2000, 2010), as well as the proposed methods added by this study in order to improving the model:

1. After arranging the data in descending order, specifying the threshold value is the next step of the analysis.
2. Due to the lack of a suitable method to determine a true threshold value, this study introduced a reliable approach, which is called Hybrid method. The Hybrid method consists of two graphs, as following descriptions:



At the first stage, the method of mean residual life (MRL) plot (Fig. 2) is recommended to be evaluated for determining a true threshold value. The plot is a part of the GPD model for determining threshold value which was described in the previous section. In this study, for the first time, the MRL plot was employed for determining threshold value for the POT model. The prerequisite condition to employ the MRL plot is to secure the iid measurement for the data contributing in the analysis (Coles, 2001). The MRL plot is based on the determination of deviation from mean in the population. It is obvious that the extreme value analysis methods, such as the POT model, deal with standard deviation (errors of residuals) from the mean. Accordingly, the MRL plot can be employed in the POT analysis as the first part of Hybrid method to determine a true threshold value.

Second, the second part of the Hybrid method is done after determining the best fitting distribution function. Once the best fitting distribution function has been selected, several graphs of fitting for different threshold values should be drawn. For instance, in this study, several fitted Weibull distributions with shape parameter $k = 1.4$ have been depicted for different threshold values (Fig. 12). Ultimately, the most suitable threshold candidate(s) obtained from the first part (the MRL plot), should be assessed and evaluated. The thresholds, which were used in the figure can be chosen by considering the order and type of data. It is important to select the threshold values with considering the candidate(s) obtained from the MRL plot. Usually, 5 to 7 threshold values are sufficient to guide the user to make a suitable decision.

In most cases, the results of the first part of the Hybrid method determines a true threshold, however, the second part can be used for removing any misinterpretation of the MRL plot (more details in step (10)).

3. During evaluating a sample of extreme data, sometimes, an individual data exhibits a value much greater than other data.

In the process of fitting the sample of data to a candidate distribution, that particular data is plotted at a position far above the line of the fitted distribution of return period curve. Such data is known as outlier(s). Therefore, the sample of extreme data contributing in the analysis should be examined to explore for the presence of outlier(s). The process of searching for outlier(s) can be done by several methods such as a statistical examination by Barnett & Lewis (1994) or using of DOL (Deviation of OutLier) criterion method Goda (2000); Goda & Kobune (1993).

4. Calculating the two following essential sample parameters:

(a) Mean Rate λ , this parameter determines the number of observed data per year, $\lambda = N_T/k$. Where N_T , is the number of events and k , is the period of observed wave heights, in year.

(b) Censoring ν , it determines a fraction of data that contributes in the analysis, $\nu = N/N_T$, in which N is the number of data applied in the analysis and N_T , number of the total data.

5. In this case, the POT model is applied with nine distribution candidates including FT-I, four FT-II distributions with fixed shape parameters 2.5, 3.33, 5, 10, and four Weibull distributions with fixed shape parameters 0.75, 1, 1.4, 2. Then, the



best fitting distribution for the data is selected among the nine distribution functions. Eqs. (8, 9, 10) represent the parent cumulative distribution functions (CDF),

(a) Gumbel Distribution (FT-I): $-\infty < x < \infty$

$$F(x) = \exp\left[-\exp\left(-\frac{x-B}{A}\right)\right] \quad (8)$$

5 (b) Frechét Distribution (FT-II): $B-kA \leq x < \infty$

$$F(x) = \exp\left[-(1 + \frac{x-B}{kA})^{-k}\right] \quad (9)$$

(c) Weibull Distribution: $B \leq x < \infty$

$$F(x) = 1 - \exp\left[-\left(\frac{x-B}{A}\right)^k\right] \quad (10)$$

Where A and B are the scale and location parameters, respectively; k is the shape parameter (dimensionless); and x stands for the extreme variate (wave height).

- 10 6. In the next step, the assignment of the non-exceedance probability $\hat{F}_{(m)}$, or plotting position for the nine distribution functions are determined (Eqs. 11, 12, 13).

Gringorten (1963) was the first who introduced a formula to FT-I and claimed that "It yields no bias". Goda (1988) examined the formula and concluded the same result as,

$$15 \quad \hat{F}_{(m)} = 1 - \frac{m - 0.44}{N_T + 0.12} \quad (11)$$

For the Weibull distribution, Petrauskas & Aagaard (1971) introduced a formula (known as *P&A* formula). Seventeen years later, Goda (1988) argued that the formula is biased, and proposed a modified version to rectify its biasness. Thus, the Goda or the modified *P&A* formula is written as follows,

$$\hat{F}(m) = 1 - \frac{m - \alpha}{N_T + \beta} \quad (12)$$

$$20 \quad \text{for } \begin{cases} m = 1, 2, \dots, N \\ \alpha = 0.2 + \frac{0.27}{\sqrt{k}} \\ \beta = 0.2 + \frac{0.23}{\sqrt{k}} \end{cases}$$



The plotting position empirical formula for the FT-II distribution function, was introduced by Goda & Onozawa (1990) as follows,

$$\hat{F}(m) = 1 - \frac{m - \alpha}{N_T + \beta} \quad (13)$$

$$for \begin{cases} m = 1, 2, \dots, N \\ \alpha = 0.44 + \frac{0.52}{\sqrt{k}} \\ \beta = 0.12 + \frac{0.11}{\sqrt{k}} \end{cases}$$

- 5 7. The next step is to calculate the reduced variate $y_{(m)}$ for the m^{th} ordered data. Equations (14, 15, 16) are the reduced variate for the three distribution families.

$$FT - I : \quad y_{(m)} = -\ln[-\ln\hat{F}_{(m)}] \quad (14)$$

$$FT - II : \quad y_{(m)} = k[(-\ln\hat{F}_{(m)})^{\frac{-1}{k}} - 1] \quad (15)$$

$$Weibull : \quad y_{(m)} = [-\ln(1 - \hat{F}_{(m)})]^{\frac{1}{k}} \quad (16)$$

- 10 8. After determining the threshold value, the exceeded data over a certain threshold are employed in the analysis. Then, the plotting positions in Eqs. (11, 12, 13) are evaluated to use in Eqs. (14, 15, 16). The next step would be the estimation of the parameters A and B by means of the Least Squares method. The input data used in the Least Squares method are the extreme variates with the corresponding reduced variates for each distribution function (Goda, 2000).

9. Best Fitting for the Dataset

- 15 In this step, three methods are discussed: Correlation Coefficient method and MIR criteria which are recommended by Goda (2000), and the method of Norm of Residuals (N.o.R), presented by this paper.

- 20 (a) The Correlation Coefficient of the order statistics (data) x_m and their corresponding reduced variates $y_{(m)}$ should be computed to determine the best fitting distribution for the dataset. The largest value represents a distribution function with the best fit for the dataset. In some literatures, it has been shown by R^2 . The correlation coefficient value is the square root of R^2 .
- (b) The MIR criterion is a method to evaluate the reliability of the results of correlation coefficient (the best fitting distribution is the one with the smallest ratio). However, Goda (2000) recommended that the criterion of the correlation



coefficient (CC.) produces better results than that of the MIR criterion, and concluded that further examination on the MIR criterion is required.

In this study, during several examinations, the results of MIR indicated some instabilities in obtaining the best fitting distribution function.

- 5 (c) Due to instabilities come with results of the MIR criterion, this study proposed an alternative method instead of the MIR criterion method. The method is called Norm of Residuals (N.o.R), as the following formula,

$$RSS = \sum_i^n (y_i - \hat{y}_i)^2 \quad (17)$$

$$\text{norm of residual} = \sqrt{RSS} \quad (18)$$

where RSS is the residual sum of squares, y_i and \hat{y}_i are the original data values and modeled values, respectively.

10 The value of N.o.R represents how well dataset fit with a statistical model. It varies between zero and infinity; however, values closer to zero indicate better fit.

- (d) R square (R^2) is also calculated for each distribution. R^2 is always ranged between 0 and 1 (including both numbers), the closer value to 1 indicates better fit, it has the following formula,

$$TSS = \sum_i^n (y_i - \bar{y})^2 \quad (19)$$

$$R^2 = 1 - \frac{RSS}{TSS} \quad (20)$$

where TSS is the total sum of squares, and \bar{y} stands as the mean of y .

10. **Second Part of Hybrid Method:** Once the best fitting distribution for the dataset has been verified, the second part of Hybrid method is done to examine the reliability of the obtained threshold from the first part of conducting the Hybrid method.

20 In this step, two criteria, recommended by Goda (2000) are imposed on the candidates (the most probable true threshold values) obtained from the first part of Hybrid method. The two criteria are as follows: First, the threshold would be better set to maximize the number of storm. Second, if the threshold value is set too high, the number of exceedances data becomes too small. If it is considered too low, a sequence of several storm events could be judged as one event and the number of data will be decreased.

25 The two recommendations (criteria) should be considered during the visual assessment of the graphs in (Fig. 12) to choose the most suitable candidate obtained from the first part of Hybrid method.

11. In this study, the credibility of the N.o.R method was examined by another goodness-of-fit test . The Chi-Square test has been a reliable method to examine the distribution goodness-of-fit (Romeu, 2003). It uses the following formula,

$$\chi^2 = \sum_{i=1}^k \frac{(e_i - o_i)^2}{e_i} \sim \chi^2_{k-1-nep} \quad (21)$$



Where o_i is the observed number of data in the specific bin, e_i is the expected number of data in bin i , and k stands for the total number of bins. $k - 1 - nep$, is the Chi-Square degrees of freedom, in which nep is the abbreviation of Number of Estimated Parameters.

The Chi-Square test is not part of the POT model. In this study, it has exceptionally been used to examine the reliability
 5 of the N.o.R method.

12. The return values were estimated using Eqs. (22 - 25), which is the next step of developing the POT model,

$$\hat{X}_m = \hat{B} + \hat{A} \cdot y_R \quad (22)$$

where \hat{A} and \hat{B} are the estimated scale and location parameters, respectively. \hat{X}_m is the return value, and y_R is the reduced variate, which is calculated as a function of the return period. In other words, by using the following equations,
 10 the return values are estimated by extrapolating the best-fit probability distribution.

$$FT - I : y_R = -\ln -\ln[1 - \frac{1}{\lambda R}] \quad (23)$$

$$FT - II : y_R = k[-\ln(1 - \frac{1}{\lambda R})]^{-1/k} - 1 \quad (24)$$

$$Weibull : y_R = [-\ln(\lambda R)]^{1/k} \quad (25)$$

where λ is the mean rate of the sample.

15 13. The 90% confidence intervals band was calculated by empirical formula introduced by Goda (1992). The empirical formula were categorized and described several years later as a textbook (Goda, 2000).

3 Results and Discussions

3.1 The GPD Model

3.1.1 Threshold Value

20 Coles (2001) recommended to use two techniques to determine the threshold value for the GPD model; the plot of mean excess against threshold (Fig. 2), and the complementary technique to fit GPD at a range of thresholds (Figs. 4, 5).

The first method, Fig. (2) represents that the graph follows a linear trend until the threshold exceeds $2m$. After which, from $u = 2.2m$, it is approximately linear until $u = 2.9m$. If we choose $u = 2.9m$, there will be only 43 exceedances of the threshold, not quite enough to make inferences, and the analysis leads to high variances. Moreover, the information in the plot
 25 with a large threshold value could be unreliable due to the low number of data on which the estimate of return periods and confidence interval are based. Accordingly, it is better to conclude $u = 2m$ as the threshold value for the data.

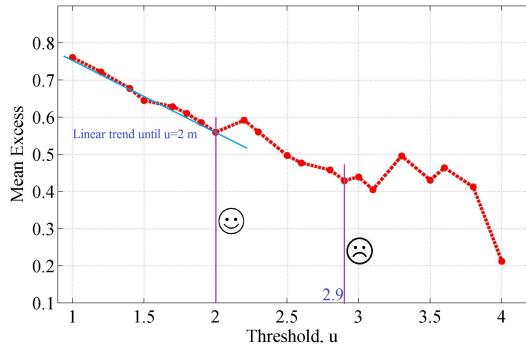


Figure 2. Mean residual life plot.

3.1.2 Parameter Estimation

The ML estimators is applied to estimate the shape and scale parameters for the exceedances data, which are equal to 263 independent storm events. The defined threshold is $u = 2 \text{ m}$ and the estimated parameters are $\hat{\xi} = -0.1442$ and $\hat{\sigma} = 0.6451$.

Numerical method was used to calculate the parameters. However, because of converging the estimation to problem, care must 5 be taken to avoid instabilities in numerical work around $\xi = 0$. As discussed in the preceding section, $\xi = 0$ is the exponential distribution, in which the shape of generalized Pareto distribution could be heavier or lighter tailed alternatives. Hence, the restricted attention should be assigned to the case $-1/2 < \xi < 1/2$, except $\xi = 0$.

3.1.3 Return Plot and Confidence Band

The return plot of the GPD model in Fig. (3) represents that the GPD model was fitted for the wave heights data. The abscissa 10 of the graph represents the average recurrence intervals (ARI) from 1 to 100 years against the significant wave height (SWH) at y axis.

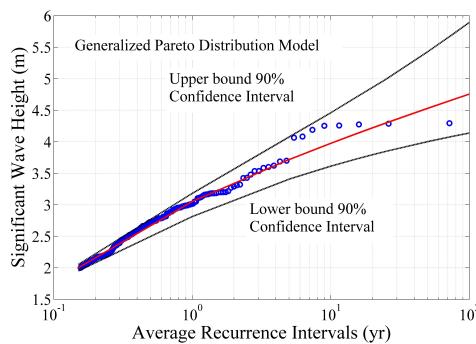


Figure 3. Return plot for the GPD model, the graph sufficiently fitted for the wave height data.



The dotted lines indicate the upper and lower confidence intervals band. In this research, the confidence intervals were calculated using the variance-covariance matrix and the delta method (Coles, 2001).

To provide the variance-covariance matrix the following evaluations are required: The first and second derivatives of the log-likelihood formula with respect to the parameter σ can be obtained as,

$$5 \quad \frac{\partial \ell}{\partial \sigma} = -\frac{k}{\sigma} + \left(\frac{1}{\xi} - 1\right) \sum_{i=1}^k \frac{\xi X}{\sigma^2 - \xi X \sigma} \quad (26)$$

$$\frac{\partial^2 \ell}{\partial \sigma^2} = \frac{k}{\sigma^2} + \left(\frac{1}{\xi} - 1\right) \sum_{i=1}^k \frac{-(2\sigma - \xi X)\xi X}{(\sigma^2 - \xi X \sigma)^2} \quad (27)$$

The first and second derivatives of the log-likelihood formula with respect to parameter ξ can also be obtained as,

$$\frac{\partial \ell}{\partial \xi} = -\frac{1}{\xi^2} \sum_{i=1}^k \log\left(1 - \frac{\xi X}{\sigma}\right) + \left(\frac{1}{\xi} - 1\right) \sum_{i=1}^k \frac{-X}{\sigma - \xi X} \quad (28)$$

10

$$\frac{\partial^2 \ell}{\partial \xi^2} = \frac{2}{\xi^3} \sum_{i=1}^k \log\left(1 - \frac{\xi X}{\sigma}\right) - \frac{2}{\xi^2} \sum_{i=1}^k \frac{-X}{\sigma - \xi X} + \left(\frac{1}{\xi} - 1\right) \sum_{i=1}^k \frac{X^2}{(\sigma - \xi X)^2} \quad (29)$$

To calculate the second derivatives of log-likelihood formula with respect to σ and ξ , the first derivative of the log-likelihood formula should be derived with respect to parameter ξ . The second derivative of Eq. (28) with respect to σ can be derived as,

$$\frac{\partial^2 \ell}{\partial \sigma \partial \xi} = -\frac{1}{\xi^2} \sum_{i=1}^k \frac{\xi X}{\sigma^2 - \xi X \sigma} + \left(\frac{1}{\xi} - 1\right) \sum_{i=1}^k \frac{X}{(\sigma - \xi X)^2} \quad (30)$$

15 The results of Eqs. (27, 29, 30) are substituted in the following matrix,

$$M = \begin{bmatrix} -\frac{\partial^2 \ell}{\partial \sigma^2} & -\frac{\partial^2 \ell}{\partial \sigma \partial \xi} \\ -\frac{\partial^2 \ell}{\partial \xi \partial \sigma} & -\frac{\partial^2 \ell}{\partial \xi^2} \end{bmatrix}$$

Ultimately, the confidence intervals band could be obtained as the following procedure: after substituting the calculated values in the M matrix, the inversion of the matrix should then be calculated. The return matrix will be the Variance-Covariance matrix for the exceeded data over a threshold u . Then, the delta method can easily be calculated based on the instructions 20 provided by Coles (2001).

3.1.4 Complementary Technique (Threshold Value)

The plots of modified scale ($\sigma^* = \sigma - \xi u$) and shape against threshold values are the complementary technique to fit GPD at a range of threshold values, and to check the stability of parameter estimates.

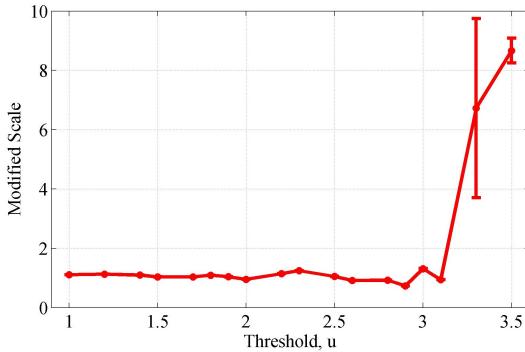


Figure 4. GPD modified scale parameter $\sigma^* = \sigma - \xi u$ against threshold u .

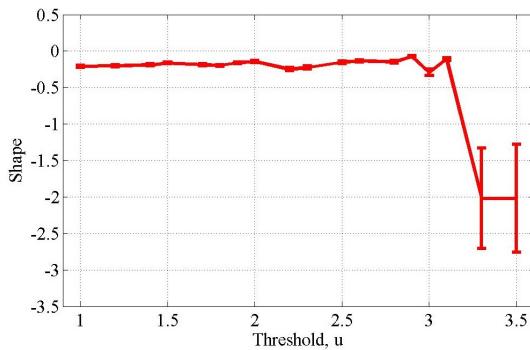


Figure 5. GPD shape parameter against threshold u .

If GPD is a reasonable model for exceedances data over a threshold u , then excesses of a higher threshold, u_1 should also follow a generalized Pareto distribution. It should be noted that the shape parameters of the two situations are identical (for u and u_1). The GPD modified scale parameter and shape parameter are plotted against different threshold values (Figs. 4, 5).

The figures represent that the modified scale and shape parameters are approximately constant until $u = 2m$, after which the graphs slightly rise and decline, respectively. Thus, the threshold value for the Labuan dataset is selected as $u = 2m$ for this GPD analysis.

3.1.5 Sensitivity of GPD With Different Threshold

The examinations on the model of Generalized Pareto Distribution indicated that the model is sensitive with any change in the value of threshold. Figure (6) represents the sensitivity of return values (SWH) of the GPD model with different thresholds (u) in a range of return periods (ARIs).

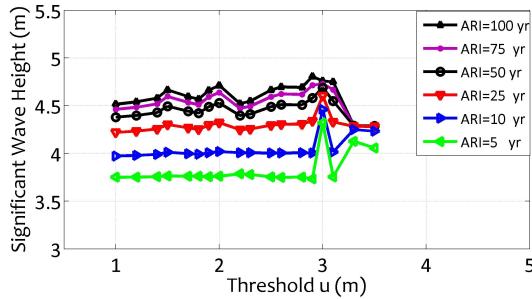


Figure 6. Sensitivity of the GPD model with different threshold u .

The threshold value at 2.9 m represents a sharp increase in graphs 5, 10, 25 and 50 years. The instability in the graphs at 2.9 m occurred due to the low number of exceeded data contributing in the analysis. In addition, low sensitivity for 5 and 10 years return periods can be observed for different threshold values.

However, for higher return periods the GPD model represents different behaviors (higher sensitivity) in estimating the 5 significant wave height.

3.1.6 Other Examinations on GPD Model

Quantile-quantile (QQ), probability-probability (PP) plots (Figs. 7, 8), and return plot (Fig. 3) leave no doubt that the generalized Pareto distribution model is fitted for the dataset. The three described plots represent the quality of the fitted GPD for the dataset. If the GPD model is reasonable for the data excesses, the points in the QQ and PP plots should be close to the diagonal 10 line (Figs. 7, 8), meaning that they should be arranged in a linear manner. In both plots, the points are sufficiently close to the diagonal line to lend support to the fitted model.

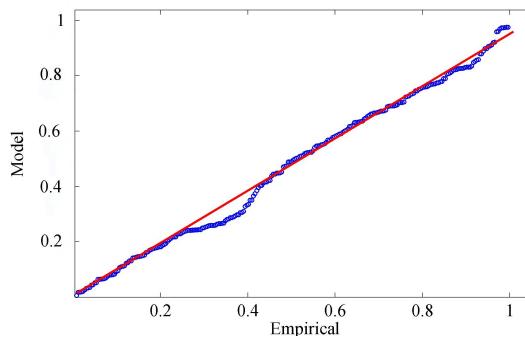


Figure 7. Probability plot for the GPD model.

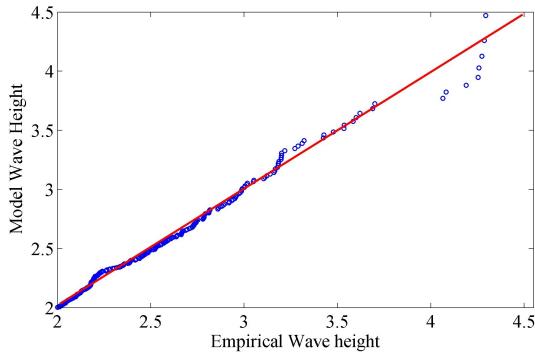


Figure 8. Quantile plot for the GPD model.

3.2 The POT Model

3.2.1 Threshold Value, Part One (Hybrid Method)

After arranging data in descending order, the first step of conducting the model is the process of determining threshold value by the Hybrid method. In Section 3.1.1, a true threshold value was selected by the MRL plot (Fig. 2) as the first part of the 5 Hybrid method. Based on the interpretation of the figure, the value of $u = 2 \text{ m}$ was concluded as a suitable candidate.

3.2.2 Best Fitting Distribution (Norm of Residual Method)

Once the threshold value has been determined, choosing the best fitting distribution is the next step. To select the best fitting distribution among the nine distribution functions, this paper proposes calculating Norm of Residuals (N.o.R) for all nine candidates. The graphs in Figs. (9, 10, 11) show the fittings of nine distribution functions for the dataset, and their residuals 10 plots, as well as calculating the R^2 for each candidate.

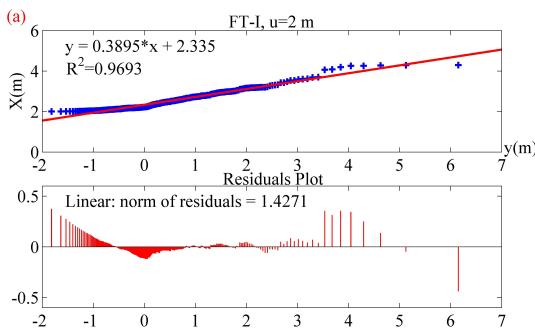


Figure 9. Fitting FT-I (Gumbel) distribution function for the data, the threshold value $u = 2 \text{ m}$, and the horizontal axis for the fitted graph is Reduced Variate (dimensionless).



The values of Norm of Residuals and R^2 in Figs. (9, 10, 11) represent that Weibull distribution with shape parameter, $k = 1.4$ obtained the least value for N.o.R and the highest value for R^2 among other distribution functions. Means that the Weibull distribution with shape parameter $k = 1.4$ is the best fitting for the dataset (Based on R^2 and N.o.R methods).

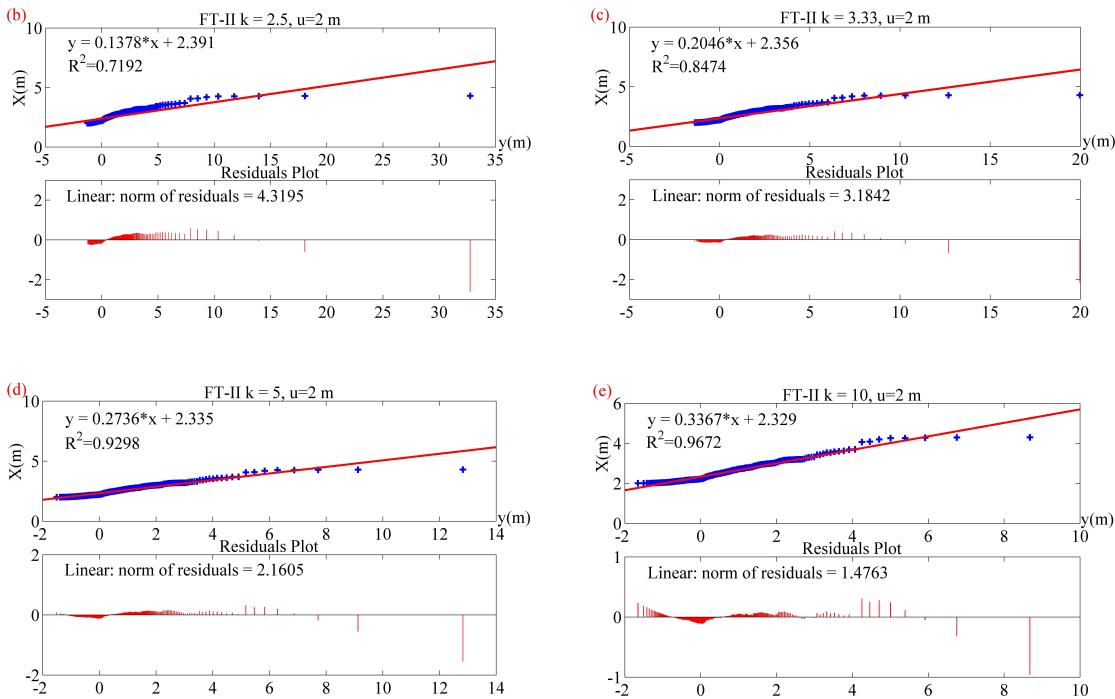


Figure 10. Fitting four FT-II (Frechét) distribution functions for the data, the threshold value $u = 2$ m, and the horizontal axis for the fitted graphs are Reduced Variates (dimensionless).

3.2.3 Best Fitting Distribution (POT model)

- 5 In this study, based on Goda (2000) recommendation, the Correlation Coefficient (C.C) and MIR criterion were calculated to find the best fitting distribution. Table (2) shows the results of Correlation Coefficient and MIR criterion, as well as the results of the Norm of Residuals (N.o.R) method. All three methods (MIR, N.o.R and C.C) show that the Weibull distribution (with $k=1.4$) is the best fitting distribution function among nine distribution candidates.

3.2.4 Determining Outliers

- 10 Usually, at this stage of developing the POT model, data should be assessed in order to find out the possible outlier(s) among the dataset. The possibility of existing any outlier among the data could early be anticipated by doing data analysis and drawing box plot to check outlier(s). For instance, if the box plot represented any outlier, the data should be re-evaluated for any plausible

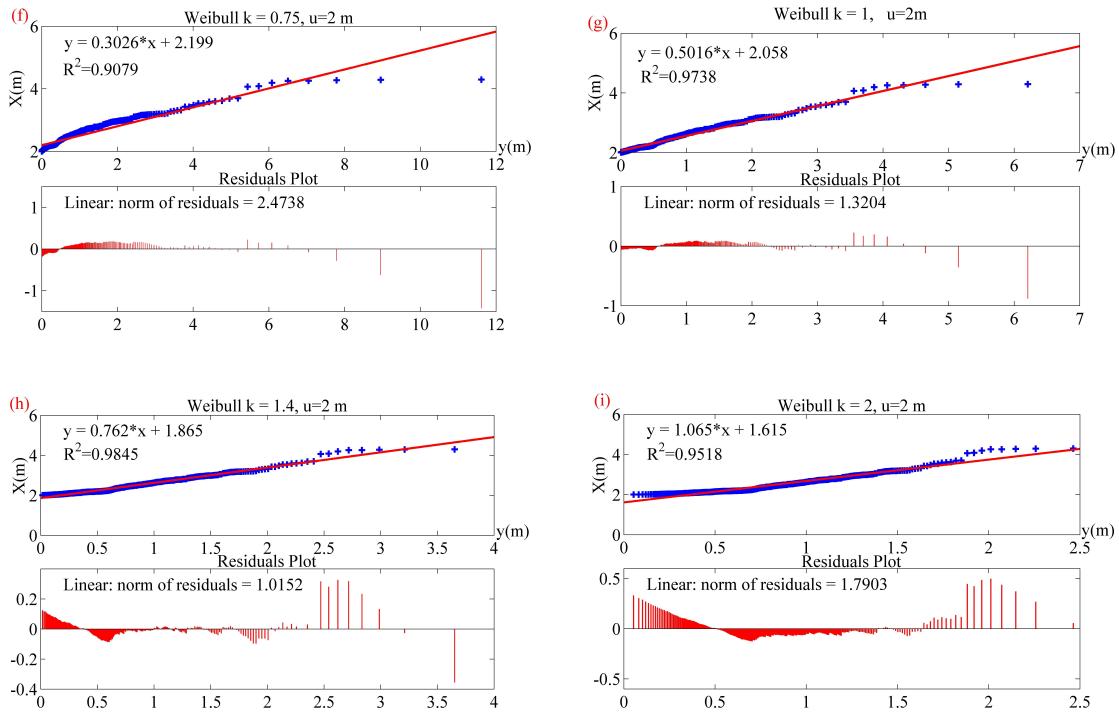


Figure 11. Fitting four Weibull distribution functions for the data, the threshold value $u = 2 \text{ m}$, and the horizontal axis for the fitted graphs are Reduced Variates (dimensionless).

Table 2. The results of three goodness of fit methods; proposed by Goda (2000), and this paper.

Applied methods		FT-I	FT-II				Weibull			
			k=2.5	k=3.33	k=5	k=10	k=0.75	k=1	k=1.4	k=2
POT method	C.C	0.9846	0.8481	0.9205	0.9642	0.9835	0.9528	0.9868	0.9922	0.9756
	MIR	2.8317	4.2103	3.4022	2.4566	1.8809	6.1692	2.2670	1.7447	6.7396
Proposed by this paper	R^2	0.9693	0.7192	0.8474	0.9298	0.9672	0.9079	0.9738	0.9845	0.9518
	N.o.R	1.4271	4.3195	3.1842	2.1605	1.4763	2.4738	1.3202	1.0152	1.7903

Correlation Coefficient (C.C), Norm of Residual (N.o.R).

outlier(s) using suitable methods. In this study the recommended DOL criterion method by Goda (2000) was employed. By definition, if the DOL method detected any outlier, it is recommended to check the process of data acquisition. If no error was occurred, therefore, instead of removing the outlier, the first fitting distribution should be removed from the list of distribution candidates. Then, the second best fitting distribution should be chosen for the analysis.

5 In this study, the DOL method has found no any outlier among the data. However, to explain the proposed methods and to show the instability of the MIR method, it is assumed that there has been an outlier among the dataset. As foregoing discussions,



if DOL detected any outlier, the data should be checked to find out human errors or instrument malfunction in the process of data collection. If the quality of data has proved that there is no error in the data acquisition process, then instead of removing the data, the first fitting distribution should be removed from the analysis. Table (2) shows that the first fitting distribution is Weibull (with $k=1.4$).

5 Therefore, the second candidate shows its importance, for instance, in this case, the C.C, R^2 and N.o.R methods represent Weibull ($k = 1$) as the second candidate (Table 2). However, the MIR criterion method represents a different candidate (FT-II with $k=10$). Means that the methods obtained different parent distribution as the best fitting parent distribution function. Therefore, another goodness-of-fit method should be used to specify a suitable parent distribution. In the next section, the Chi-Square test is developed to figure out the best parent distribution function.

10 3.2.5 Developing the Chi-Square Method

Evaluating the underlying distribution by a goodness-of-fit method such as the Chi-Square method offers a suitable process to examine the accuracy of selected distribution as the best fitting for a set of data. By definition, in extreme value analysis, a particular distribution is usually considered as the underlying distribution, and it is assumed that the data should follow this pre-specific distribution. Corder & Foreman (2009) and Romeu (2003) in separate literatures argued that this assumption could 15 be risky if the considered underlying distribution is not correct, and consequences of incorrectly identifying a distribution may prove costly. Vivekanandan (2015) employed the Chi-Square test to evaluate the adequacy of fitting of probability distributions for their data and noted that the Chi-Square method is a credible approach.

The results of Chi-Square test were listed in Table (3). First, the dataset should be divided into 9 bin-sectors ($k = 9$ sub-intervals) by end points 2.12, 2.2, 2.3, 2.5, 2.7, 2.85, 3.1 and 3.6, which are listed in the second column of the table. By applying 20 the probability density function (PDF) of Weibull distribution, the non-exceedance probabilities are calculated in third column. In fourth column, the probabilities of the individual sub-interval from the third column are listed. Then, in the column fifth, the results in column four should be multiplied by the total sample size contributed in the analysis. In the sixth column, the number of samples from each sub-intervals are listed. Finally, in the last column the test parameters are calculated.

The Chi-Square table (<http://www.unc.edu/farkouh/usefull/chi.html>) includes the information of probabilities of several 25 percentage of confidence intervals and values of freedom $DF = k - 1 - n = (9 - 1 - 2)$, where k is the number of bins and n is the number of parameters for the assumed distribution (Romeu, 2003).

According to the table of Chi-Square, the crosspoint of two values consisting of 95% probability, and $DF = 6$, spots the 30 value of 12.53. The calculated Chi-Square statistic value for Weibull distribution based on Table (3) is 10.66, which is less than 12.53. Therefore, it is assumed that the distribution of the population originating from the data is Weibull, which proves the accuracy of the Norm of Residual and Correlation Coefficient methods, and shows the instability of the MIR criterion method.



Table 3. Intermediate values for the Weibull Chi-Square test.

Bin No:	End Top of Bins	CumProb.	CellProb.	Expected	Observed	Chi-square test
1	2.12	0.170	0.170	44.710	45	0.002
2	2.20	0.271	0.101	26.640	41	7.740
3	2.30	0.366	0.095	24.990	22	0.357
4	2.50	0.539	0.173	45.470	36	1.973
5	2.70	0.679	0.140	36.790	33	0.391
6	2.85	0.761	0.082	21.620	23	0.088
7	3.10	0.860	0.099	25.960	25	0.035
8	3.60	0.958	0.098	25.720	26	0.003
9	infin	1.000	0.042	11.099	12	0.073
Total:			1.000	263.000	263	10.661

3.2.6 Sensitivity of the POT Model With Changes in Threshold Value

In this study, for the first time, the sensitivity of the POT model with different threshold values was scrutinized (Fig. 12). The figure represents the best fittings of Weibull distribution with shape parameter $k = 1.4$ for exceeded data over several thresholds versus reduced variate. The advantage of using reduced variate for abscissa rather than ARI is that the graphs can separately be drawn, consequently, they could be evaluated easily. The figure shows that lower thresholds yield higher return values due to increasing the sample standard deviation, and higher threshold yields the inverse effect. The top end of the estimated lines indicate return values of wave heights for 100 years return periods.

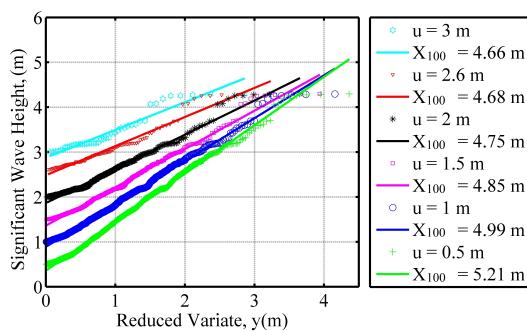


Figure 12. Sensitivity of the POT model in estimating 100 years return values when different threshold values are employed, [The top end of the estimated lines indicate the return values estimated for 100-year return periods (the graphs are fitted Weibull distribution with $k = 1.4$, and $u = 0.5, 1, 1.5, 2, 2.6$ and 3)].



3.2.7 Second Part of the Hybrid Method

Once the best fitting distribution for the dataset has been verified, it is important to examine the reliability of the threshold obtained through the first part of the Hybrid method. To select suitable threshold value, Goda (2000) noted that the threshold would be better set to maximize the number of storm wave heights. However, if the threshold is considered too low, a sequence 5 of several storm events could be judged as one event and the number of data will be decreased, and the inference leads to bias. On the other hand, if the threshold value is set too high, the number of exceedance data becomes too small, again it leads to bias. Figure (12) represents several return periods corresponding to different thresholds (0.5, 1, 1.5, 2, 2.6, and 3 m). The results from the first part of the Hybrid method can be visually judged on Fig. (12) with considering the above-mentioned two criteria for lower and higher thresholds. For instance, in this case, in the first part of Hybrid method, $u = 2.9\text{ m}$ was one of the 10 another possible threshold candidate (Fig. 2), however, based on the second limitation and visually judgment, $u = 2\text{ m}$ is not an appropriate threshold value because it is too high. Thus $u = 2\text{ m}$ is selected as the best choice.

3.2.8 Return Plot

The return plot in Fig. (13) represents that the Weibull distribution with $k = 1.4$ is sufficiently fitted for the date. The graph was depicted based on average recurrence intervals on the x axis to evaluate the significant wave heights in different return periods.

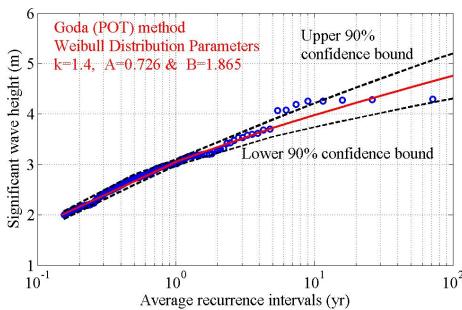


Figure 13. The return plot of the POT model with 90% confidence intervals band upper and lower the estimated line.

15 The scale ($A = 0.7260$) and location ($B = 1.865$) parameters are estimated by Least Squares method with considering the mean rate ($\lambda = 263/41 = 6.41$) for the dataset. The figure indicates another diagnostic method for evaluating of the model accuracy. The dotted lines represent a band of 90% confidence intervals around the estimated line.

3.3 Comparison of the Two Models

The higher number of data contributing in the analysis for the models of POT and GPD provide narrower confidence intervals band. Therefore, choosing lower threshold value (if possible), could be an easiest way to increase the number of data, consequently, the narrower confidence intervals band will be obtained (for both models).



Thompson et al. (2009) proposed a method to determine threshold value for the GPD model, and argued that lower threshold value yields narrower confidence intervals band. It means that the higher number of data returns lower uncertainties in the GPD model. However, if the threshold value was selected too low, it is likely to violate the asymptotic basis of the model and the model leads to bias (Coles, 2001). In contrast, if the threshold value was chosen too high, little excessive data remain to estimate the model, and the inference leads to high variance. Therefore, a suitable selection of threshold value makes a balance between bias and variance in GPD model (Scarrott & MacDonald, 2012).

The comparison of Fig. (13) with Fig. (3) discloses that the 90% confidence intervals band obtained by the POT model is narrower than the one estimated by the GPD model. It means that there are lower uncertainties in using the POT model. With regard to the importance of lower uncertainties as a significant criterion for designing maritime structures, narrower confidence intervals band is desirable. There are many reasons that can change the uncertainties of confidence intervals band (in both models) such as the number of data contributing in the analysis, the accuracy of the estimators methods (the measurement of their biasedness), the accuracy of choosing threshold value, and the compatibility of the data with the extreme model (it directly or indirectly relates to the iid measurement of the data). Therefore, since the two models have used the same dataset in this study, the narrower confidence band is an advantage for the POT model. However, with regard to the POT inferences, the fixed shape parameter can cause unrealistic narrower confidence band, but for this model, it has not been proven, yet.

The estimated return values for the GPD and POT models, as well as the report of the Coastal and Offshore Engineering Institute, Universiti Teknologi Malaysia are shown in Table (4). The obtained return values by the GPD model are remarkably close to those estimated by the POT model. The proximity of the return values represents that the proposed methods - Norm of Residual and Hybrid methods - have improved the POT model. However, the third row of Table (4) represents the return values obtained by the Gumbel distribution, and using the same dataset. It indicates a higher return value for 100 years return period. However, it should be noted that the Gumbel distribution function was one of the POT distribution candidates (Fig. 9). Based on the results of the N.o.R method, it was not a best fitting distribution for the dataset (Table 2). Thus, the use of a such stable method (norm of residual) to determine the best fitting distribution shows its importance, again.

Table 4. The results of significant wave heights for a range of several return periods obtained by this study, and the Institute of Coastal and Offshore Engineering, Universiti Teknologi Malaysia (2009).

Return Periods (yr)		1	5	10	15	20	25	50	75	100
This Study	GPD Model	3.05	3.76	4.02	4.16	4.25	4.32	4.53	4.64	4.71
This Study	POT Model	3.05	3.72	3.97	4.12	4.22	4.30	4.53	4.66	4.75
	COEI INST.	-	3.57	3.89	-	4.20	-	4.60	-	4.90

4 Conclusions

In this study, two commonly used extreme wave analysis models, namely Peaks-Over-Threshold (POT) and Generalized Pareto Distribution (GPD) were developed in order to highlight their merits and limitations, and compare them together for evaluating



their results. The results of examinations represented that the POT model introduced by Goda in 1988 was not equipped with a suitable approach to determine threshold value, moreover, the MIR criterion method introduced by the author to determine the best fitting distribution function has shown some instabilities in determining the results, especially where the parent distribution function was unknown. To deal with the limitations, two methods, namely Hybrid, and Norm of Residuals were proposed in 5 this study. The drawn results are summarized as follows: *i*) the proposed Norm of Residuals method is a stable approach with reasonable results to specify the best fitting distribution function, and its credibility has been proven by Chi-Square test. *ii*) the proposed Hybrid method to select a suitable threshold value for the model of POT is reliable, feasible and completely matched with the model. *iii*) the obtained confidence intervals band by the POT model is narrower than the one obtained by the GPD model. The narrower band provides lower uncertainties which is desirable for the designers. The lower uncertainties of the 10 POT model is related to several reasons: number of data contributing in the analysis, the employed distribution function, the estimators method, and the approach of determining threshold value. *iv*) the proximity of the obtained return values of the two models (POT and GPD) demonstrated that the two models are credible and sufficiently fitted for the data, and the proposed approaches (N.o.R and Hybrid method) have strikingly improved the POT model as well.



References

- Godav. Y. (2010). Random seas and design of maritime structures. World scientific.
- Godav. Y. (2000). Random seas and design of maritime structures. World scientific.
- Goda Y., Kobune K. (1993). Distribution function fitting for storm wave data. In Coastal Engineering Conference. (Vol. 1, pp. 18-18). ASCE
5 American Society of Civil Engineers.
- Goda Y. (1992). Uncertainty of design parameters from viewpoint of extreme statistics. Journal of offshore mechanics and arctic engineering. 1;114(2):76-82.
- Goda Y., & Onozawa M. (1990). Characteristics of the Fisher-Tippett type II distribution and their confidence intervals. In Proceedings of the Japan Society of Civil Engineers (No. 417/11, p. 13).
- 10 Goda, Y. (1988). On the methodology of selecting design wave height. Coastal Engineering Proceedings, 1(21).
- Grimshaw S.D. (1993). Computing maximum likelihood estimates for the generalized Pareto distribution. Technometrics. 1;35(2):185-91.
- Pickands J. (1975). Statistical inference using extreme order statistics. the Annals of Statistics. 1:119-31.
- Davison A.C. (1984). Modelling excesses over high thresholds, with an application. In Statistical extremes and applications. (pp. 461-482). Springer Netherlands.
- 15 Smith R.L. (1985). Maximum likelihood estimation in a class of nonregular cases. Biometrika. 1;72(1):67-90.
- Smith R.L. (1984). Threshold methods for sample extremes. In Statistical extremes and applications. (pp. 621-638). Springer Netherlands.
- Van Montfort M.A., & Witter J.V. (1985). Testing exponentiality against generalised Pareto distribution. Journal of Hydrology. 20;78(3):305-15.
- Hosking J.R., & Wallis J.R. (1987). Parameter and quantile estimation for the generalized Pareto distribution. Technometrics. 1;29(3):339-49.
- 20 Moharram S.H., Gosain A.K., Kapoor P.N. (1993). A comparative study for the estimators of the generalized Pareto distribution. Journal of Hydrology. 30;150(1):169-85.
- Castillo E., & Hadi A.S. (1997). Fitting the generalized Pareto distribution to data. Journal of the American Statistical Association. 1;92(440):1609-20.
- Pandey M.D., Van Gelder P.H., Vrijling J.K. (2004). Dutch case studies of the estimation of extreme quantiles and associated uncertainty by
25 bootstrap simulations. Environmetrics. 1;15(7):687-99.
- Pandey M.D., Van Gelder P.H., Vrijling J.K. (2001). The estimation of extreme quantiles of wind velocity using L-moments in the peaks-over-threshold approach. Structural Safety. 31;23(2):179-92.
- Coles S., Pericchi L.R., Sisson S. (2003)A fully probabilistic approach to extreme rainfall modeling. Journal of Hydrology. Mar 25;273(1):35-50.
- 30 Coles S. (2001). An introduction to statistical modeling of extreme values. London: Springer Series in Statistics;
- Öztekin, T. (2005). Comparison of parameter estimation methods for the three-parameter generalized Pareto distribution. Turkish journal of agriculture and forestry. Nov 28;29(6):419-28.
- Jagger T.H., & Elsner J.B. (2006). Climatology models for extreme hurricane winds near the United States. Journal of Climate. Jul;19(13):3220-36.
- 35 Moisello U. (2007) On the use of partial probability weighted moments in the analysis of hydrological extremes. Hydrological processes. May 15;21(10):1265-79.



- Zhang Y, & Cao Y. (2015). A fuzzy quantification approach of uncertainties in an extreme wave height modeling. *Acta Oceanologica Sinica*. Mar 1;34(3):90-8.
- Coastal and Offshore Engineering Institute, Universiti Teknologi Malaysia. *Coastal hydraulic and sedimentation study for the assessment of dredging requirements for ports and river mouths in Malaysia*. The report was submitted to the Marine Department of Malaysia. 2009.
- 5 Mathiesen M., Goda Y., Hawkes P.J., Mansard E., Martin M.J., Peltier E., Thompson E.F., & Van Vledder G. (1994). Recommended practice for extreme wave analysis. *Journal of hydraulic Research*. Nov 1;32(6):803-14.
- Syed Abdullah S.M. (1992). The Coastal Zone in Malaysia: Processes, Issues and Management Plan. Malaysian National Conservation Strategy. p.36.
- Mackay, E. B., Challenor, P. G., & Bahaj, A. S. (2011). A comparison of estimators for the generalised Pareto distribution. *Ocean Engineering*, 10 38(11), 1338-1346.
- Zhang, J., & Stephens, M. A. (2009). A new and efficient estimation method for the Generalized Pareto Distribution. *Technometrics*, 51(3), 316-325.
- Dupuis, D. J., & Tsao, M. (1998). A hybrid estimator for generalized Pareto and extreme-value distributions. *Communications in Statistics- Theory and Methods*, 27(4), 925-941.
- 15 de Zea Bermudez, P., & Kotz, S. (2010). Parameter estimation of the generalized Pareto distribution?Part I. *Journal of Statistical Planning and Inference*, 140(6), 1353-1373.
- Davison, A. C., & Smith, R. L. (1990). Models for exceedances over high thresholds. *Journal of the Royal Statistical Society. Series B (Methodological)*, 393-442.
- Barnett, V., & Lewis, T. (1994). Outliers in statistical data.
- 20 Gringorten, I. I. (1963). A plotting rule for extreme probability paper. *Journal of Geophysical Research*, 68(3), 813-814.
- Petrauskas, C., & Aagaard, P. M. (1971). Extrapolation of historical storm data for estimating design-wave heights. *Society of Petroleum Engineers Journal*, 11(01), 23-37.
- Romeu, J. L. (2003). The Chi-Square: A large-sample goodness of fit test. *Start: Selected Topics in Assurance Related Technologies*, 10(4), 1-6.
- 25 Vivekanandan, N. (2015). Flood frequency analysis using method of moments and L-moments of probability distributions. *Cogent Engineering*, 2(1), 1018704.
- Corder G.W., Foreman D.I. (2009). Nonparametric statistics for non-statisticians: a step-by-step approach.
- Li, F., Bicknell, C., Lowry, R., & Li, Y. (2012). A comparison of extreme wave analysis methods with 1994-2010 offshore Perth dataset. *Coastal Engineering*, 69, 1-11.
- 30 Thompson, P., Cai, Y., Reeve, D., & Stander, J. (2009). Automated threshold selection methods for extreme wave analysis. *Coastal Engineering*, 56(10), 1013-1021.
- Scarrott, C., & MacDonald, A. (2012). A review of extreme value threshold estimation and uncertainty quantification. *REVSTAT-Statistical Journal*, 10(1), 33-60.