Weather and Climate Extremes 5-6 (2014) 40–47

Contents lists available at ScienceDirect

Weather and Climate Extremes

journal homepage: www.elsevier.com/locate/wace

Statistical analysis of extreme ocean waves in Galle, Sri Lanka

T. Thevasiyani [a,*], K. Perera [b]

[a] Postgraduate Institute of Science, University of Peradeniya, Sri Lanka
[b] Department of Engineering Mathematics, Faculty of Engineering, University of Peradeniya, Postgraduate Institute of Science, University of Peradeniya, Sri Lanka

ARTICLE INFO

Article history:
Received 4 January 2014
Received in revised form
4 July 2014
Accepted 28 July 2014
Available online 27 August 2014

Keywords:
Peak Over Threshold method
Mean residual life plot
Generalized Pareto Distribution
Return levels
Return period
Extreme value analysis

ABSTRACT

Study of extreme wave heights is vital for design of coastal structures. The purpose of this study is to estimate the wave heights for several return periods in order to use them in coastal constructions. Wave height measurements collected from Galle, Sri Lanka were analyzed. Analysis was separately performed for sea, swell and overall waves, season wise. Peak Over Threshold method (POT) was used for the sample selection. The possible range of threshold values were identified by using the Mean residual life plot and the specific threshold value was selected using the Generalized Pareto Distribution (GPD). Using the POT method, the GPD was fitted for the sampled data and the special type of GPD was identified statistically. Further diagnostic plots were obtained to ensure the validity of the distribution. Return levels were calculated for several return periods and the confidence intervals were constructed for the return levels.

Exponential distributions were the best fitted distributions for south-west (SW) monsoon and October–November (ON) season for sea wave heights, while the Beta distribution was the best fitted distribution for swell wave heights. Pareto distribution fits well with the overall wave heights for south-west monsoon and overall wave heights were fitted well with Beta distribution for October–November (ON) and March–April (MA) seasons. Analysis was omitted for March–April (MA) season and December–February (DF) season for sea and swell waves, while December–February (DF) season was omitted for overall waves as they did not have significant extreme values. It was found that, the return levels for sea wave heights were comparatively higher than the swell wave heights for all the seasons. When designing the coastal constructions, the return levels of extreme wave heights in south-west monsoon should be considered than the other seasons, and also comparatively the return levels of sea waves have significant impact than the swell waves. Moreover return levels of overall wave heights also have to be considered in designing.

1. Introduction

Often, engineering construction design of ships, boats, ports, seawall and harbors are determined by the behavior of the ocean waves. As a consequence of human settlement along the coastline and of economic activities across the ocean, extreme events generated within the ocean have to be carefully considered. Design structure of a construction will be designed with the probability of failure for a given or expected life time of the structure. Some natural events such as, rainfall, flood and occurrence of storm events in sea states are not following bell-shaped normal distributions. In such a situation extreme events were placed in the tail of the distribution. Those extreme points were identified as outliers in preliminary statistical analysis. But an outlier has to be checked carefully and proper method of analysis has to be carried out.

Sri Lanka is an island encircled by Indian Ocean and this country is an important place of the international ship path. Particularly, Colombo and Galle are considered as, central import and export commercial harbors. There was a directional wave buoy instrument installed at Galle harbor for the purpose of collecting the wave data in 1989 February by Coast conservation Project. With this data Scheffer et al. (1994) perform the EVA and obtained the return levels for SW monsoon of sea, swell and overall wave heights using the Weibull distribution. In this paper Generalized Pareto Distribution (GPD) using Peak Over Threshold(POT) method was used to estimate the distribution of the wave heights and using the identified distribution, return levels were calculated.

Unusually larger waves are called extreme waves. Even though the occurrences of the extreme waves are rare, it will make severe damages on the coastal and off-shore structures. It is very important to consider the extreme wave heights during the design process of the marine structures. Occurrences of extreme wave heights plays an important role in the construction of coastal lines and offshore structures (Soares and Scotto, 2004;

* Corresponding author. Tel.: +94 0772255188.
E-mail address: t.shiyani@yahoo.com (T. Thevasiyani).

http://dx.doi.org/10.1016/j.wace.2014.07.003

Stansell, 2004). Soares and Scotto (2001) applied several parametric models such as Weibull and f_3 to fit the long-term distribution of significant wave heights. The results indicated that the accuracy of prediction depends on the behavior of the extreme wave height values. Covariate effects are often ignored in practice because of complexity of extreme value modeling with covariates (Jonathan et al., 2008).

GPD is the distribution for threshold excesses and the shape parameter (ξ) is determining the tail behavior of the distribution (Jonathan et al., 2008; Stansell, 2004). Rayleigh distribution underestimate the occurrence of extreme crest and over-estimate the trough heights (Stansell, 2005). Buoy measurements are the trustworthy data source and can be used in the analysis. (Panchanga et al., 1998). However, coverage in data collection is limited by the failures of the instrument and damage caused by climate change.

Another vital aspect is the seasonality constraints in climate data. It was suggested that the year should be split-up into months or other sensible period by Carter and Challenor (1981). When dealing with coastal management, analysis of flooding risk and the design of off-shore structures depend on the seasonal or monthly characteristic in the return value calculation (Minguez et al., 2010). When we omit the homogeneity and carry out the analysis, it will give the unreliable results. This homogeneity can be ensured by separation into carefully selected seasons (Monsoonal analysis) (Minguez et al., 2010; Mendeza et al., 2008). Distribution of extreme wave heights differs with the season in which it occurs (Mackay et al., 2010; Soares and Scotto, 2001). Seasonality and duration were contributed to accurate estimation (Mendeza et al., 2008; Mackay et al., 2010).

Maximum individual wave heights are important in constructing offshore structure designs (CapitGo and Burrows, 1995; Alves et al., 2003). Peak Over Threshold (POT) method for wave heights was used and it was fitted to GPD (Mackay et al., 2010). Annual Maxima (AM) approach solves immediately some problems of the initial distribution method, however main difficulty of this method is, dealing with EVA with lack of enough data for distribution fitting (Soares and Scotto, 2004). The AM method is not appropriate, because less number of annual maxima data points will not provide good estimates (Soares and Scotto, 2001).

Various methods can be used for the threshold selection. Mean residual life plot was suggested by Stansell (2004) to select the specific cut-off value. Based on statistical theory Maximum Likelihood Estimator (MLE) is considered as the best parameter estimation method. Estimated parameters are unbiased and having relatively small variance (CapitGo and Burrows, 1995). There is a unique test of goodness-of-fit for GPD called "Bootstrap method for GPD". Villasenor-Alva and Gonzalez-Estrada (2009) performed this test using the data obtained from Mexico City's ozone level. Ultimate step of the analysis was the estimation of the return values with the given probability of occurrence, using the best fitted extreme value distribution (CapitGo and Burrows, 1995).

2. Data

Data collected from the CCD-GTZ Coast conservation project was used for the analysis. A directional and roll buoy WAVEC, manufactured by DATAWELL B.V., Netherlands was used to collect the wave measurements from deep sea water. It was installed 8 km south off Galle harbor at about 70 m water depth. This is the average water depth of the relatively narrow continental shelf. Even though 70 m is not considered as deep water, buoy was installed there by considering the mooring requirement and financial and practical operational constraints. In SW monsoon data coverage rate was higher than the other seasons. The wave buoy was deployed in February 1989, and data were collected every three hours for 5 and ½ years. Generally, in SW monsoon season, wave heights are higher than the other seasons in Galle, as it is located in south-west part of Sri Lanka. Unit of measurement of wave height is in meters.

Three types of waves have been separately analyzed:

- Sea waves: the waves generated under the influence of the wind within the wind field.
- Swell waves: the waves move out from the area of generation and are no longer subjected significantly to wind action.
- Overall waves: the combination of the sea and swell waves is called overall wave. Overall wave heights were calculated as follows:

$H_{overall} = \sqrt{H_{sea}{}^2 + H_{swell}{}^2}$, where H denotes the height of the waves.

In this study, sea, swell and overall waves were analyzed season wise. Seasons are given below:

- December–February (DF) season.
- March–April (MA) season.
- May–September (SW) monsoon season.
- October–November (ON) season.

3. Methods

According to the data set samples were selected above a threshold value u and POT method was used to fit the GPD distribution. The Mean Residual Life plot (MRL plot) and fitted GPD over a range of threshold values could be used for selecting the specific threshold value. Appropriate specific threshold value will be chosen by observing the stability of the GPD plot. Conditional distribution of excesses of a specific threshold u is determined by the probability density function, which is called conditional tail distribution. This conditional tail distribution can be approximated using GPD, and the distribution function is given below.

$$G(y, \sigma_u, \xi) = 1 - \left(1 + \xi\frac{y}{\sigma}\right)^{-1/\xi}$$

where σ is the scale parameter in the range of $\sigma > 0$ and ξ is the shape parameter in the range of $-\infty < \xi < \infty$

Shape parameter of ξ in GPD is same as shape parameter of Generalized Extreme Value (GEV) distribution:

$$\left(1 + \xi\frac{y}{\sigma}\right)^{-1/\xi} > 0$$

$E(y) = \frac{\sigma}{1-\xi}$, where $\xi < 1$.

Further using the sign of the shape parameter ξ, GPD will be divided into three different types of distributions.

Case 1: $\xi = 0$: Light tail distribution. Take the limit as $\xi \to 0$ then, $G(y, \sigma_u, \xi)$ follows **Exponential distribution** with mean σ with the distribution function given below:

$$G(y, \sigma_u, 0) = 1 - \exp\left(-\frac{y}{\sigma}\right)$$

Case 2: $\xi > 0$: Long tail **Pareto distribution**, where $1 - G(y)$ decays at the same rate as $y^{-1/\xi}$ for large y. And the distribution function defined as follows:
$G(y) = 1 - cy^{-\alpha}$, where $\xi = (1/\alpha)$.
Case 3: $\xi < 0$: This distribution has a finite upper end point at $-(\sigma/\xi)$. It is called **Beta distribution**.