

Influence of tropical cyclones on the 100-year return period wave height - a study based on 39-year long ERA5 reanalysis data

Muhammed Naseef T^{1,2}, Sanil Kumar V^{1*}

¹Ocean Engineering, CSIR-National Institute of Oceanography
(Council of Scientific & Industrial Research), Dona Paula, Goa - 403 004 India

²Research Scholar, Marine Science Department, Bharathidasan University,
Tiruchirappalli 620023 Tamil Nadu, India

*Corresponding author: Tel: 0091 832 2450 327 Fax: 0091 832 2450 602

Email: sanil@nio.org

Abstract

Various engineering activities in a coastal and offshore environment require detailed assessment about the probable extreme wave characteristics in its lifetime to prevent human as well as property loss. The accurate estimation of wave height with a 100-year return period not only contribute to the level of protection but also the scale of investments. In this study, significant wave height return levels for a northwestern Bay of Bengal location are estimated using extreme value models; Generalized Extreme Value (GEV) and Generalized Pareto Distribution (GPD). Both block maxima and Initial Distribution methods are used in the GEV and GPD incorporated with Peak Over Threshold (POT) series. The study is based on the European Centre for Medium-Range Weather Forecasts ERA5 reanalysis hourly data at 0.25° X 0.25° spatial resolution from 1979 to 2017. The GPD estimates show that the return levels do not follow prevailing seasonality over the location since the return levels greatly depend upon the extreme events and most of the extreme events are during the northeast monsoon season (October-January). The POT with October-January data gives 9.08 m as 100-year return level, and corresponding Indian southwest monsoon (June-September) return level is only 5.41 m. When different block periods ranging from 15 to 39 years are considered, the 50-year return levels range from 5.12 to 8.67 m for GEV with annual maxima. The influence of tropical cyclones on return levels are examined considering different case studies, and the results show that the strength of a tropical cyclone alters the return levels more significantly than the number of tropical cyclones in the observation. The 100-year return level decreases by ~30% when wave height during the two cyclones; Phailin and Hudhud are removed from the dataset.

Keywords: Extreme Value Analysis, Surface waves, reanalysis, North Indian Ocean, return period

1. Introduction

Wind-generated surface gravity waves are the most prominent feature at the ocean surface, and they exert large forces on the ocean structures (Chakrabarti, 2005). The significant wave height (SWH), approximately which is the average of the highest one-third of the wave heights in a sea state is used to describe the wave climate variation at a location. The return level of SWH with a specific return period governs the design of the structure. Tropical cyclones (TCs) are the most substantial natural phenomena which generate high waves. High waves pose a threat to the integrity of ocean structures. The Bay of Bengal (BoB) is a TC-prone area in the North Indian Ocean (NIO), which accounts for ~7% of the global TCs. TCs in NIO tend to peak during May, October and November (Gupta et al., 2019). Most of the severe TCs of the BoB is during the northeast monsoon season (October-January) mainly in October and November. Extreme ocean waves in the BoB arise due to intense TC winds. An increase in the intensity and frequency of TC/depression is reported in the BoB during 1981-2010 (Balaguru et al., 2014). During 1998 to 2012, the annual mean SWH in the western BoB show a positive trend of 0.0103 m yr^{-1} (1.03% of the annual mean per year) (Sanil Kumar et al., 2018). BoB has strong seasonality and in general, the annual cycle is divided into three seasons, pre-monsoon (February-May), southwest monsoon (June–September), northeast monsoon (October–January). The region is in the limelight due to the increased proposals related to industrial and offshore exploration activities. The rim countries in the region such as India, Thailand, Myanmar and Bangladesh face significant loss every year, not only human life but also property, due to the TC disturbances over the BoB. Therefore, predicting the most probable return level of SWH is a crucial factor in preventing damage from extreme calamities. The accurate prediction of extreme environmental parameters such as wave height with desired return level corresponding to the lifetime of marine structure or activity makes common necessity for various coastal and offshore activities (Takbash et al., 2019). Uncertainties associated with extreme value estimates makes the difficult task in designing and constructions of marine structures and these setups depend on a comprehensive and valid statistical analysis of historical datasets and ultimately give the required return level estimates (Meucci et al., 2018).

The main goal of Extreme Value Analysis (EVA) is to evaluate the probability distribution of an environmental parameter from its record of observed samples (Goda, 1988). The EVA has become common practice in disciplines such as hydrology or ocean science and atmospheric science. The scientific community aims to quantify the return level with the help of EVA. An N-year return level is a level in which it may exceed once in every N year on average (Meucci et al., 2018). Usually, extreme values are outliers or outside of the available observations so that an extrapolation of the

variable well beyond its range is needed (Samayam et al., 2017). As observational data spanning 100 years are seldom available, the return levels are generally approximated by fitting well-known distributions such as Generalized Extreme Value (GEV) distribution or Generalized Pareto distribution (GPD). The GEV family uses block maxima series such as Annual Maxima (AM), Monthly Maxima (MM) and Seasonal Maxima (SM) datasets and GPD uses Peak Over Thresholds (POT) series (Breivik et al., 2013).

Various researchers around the world have carried out the EVA of SWH on the global level as well as basin level relaying various datasets (e.g.: Katz et al., 2002; Alves et al., 2003; Soares and Scotto, 2004; Mendez et al., 2008; Vinoth et al., 2011). Teena et al. (2012) and Samayam et al. (2017) carried out return level estimation in NIO region using numerical modelling and reanalysis datasets. Naseef and Kumar (2017) estimated return levels at a location off the central west coast of India using wave rider buoy measurements. Other than these few studies, return levels in NIO, especially BoB, is less explored by the scientific community. The duration of data usually used in climatology and return level estimation is ~30 years. Hence, it is essential to know the difference in the value of the return level estimate based on shorter duration data.

Gopalpur located along the northwestern part of the BoB is influenced by TC (Amrutha et al., 2014). Maximum SWH of 7.3 m and maximum wave height of 13.5 m is measured at 50 m water depth off Gopalpur during TC Phailin in 2013 (Amrutha et al., 2014). Anjali Nair et al. (2018) observed that a single TC influenced the annual maximum SWH off Gopalpur, but the annual average SWH value is almost same (~ 1 m) in years with TC (in 2014) and without TC (in 2015). The present study analyses return levels estimated by both GEV and GPD at Gopalpur. One of the objectives of this article is to compare the estimated return periods of individual wave heights with and without the tropical cyclone. Another objective is to quantify the variation in return level for different data lengths. The article is structured as follows. In the next section, the details of the data used and the probability distribution model is introduced. The results are discussed in section 3, along with different case studies, including and excluding the TCs and the influence of the length of the data series on the return level estimation. Finally, the conclusions are presented in Section 4.

2. Data and Methods

2.1 Data

The SWH data from the ERA5 global atmospheric reanalysis (Hersbach and Dee, 2016) produced by the European Centre for Medium-Range Weather Forecasts (ECMWF) is used in the study. The ERA5 is the fifth generation of ECMWF reanalyses of the global climate having a spatial resolution

of 31 km globally and the temporal resolution of one hour (Hersbach and Dee, 2016). Naseef and Kumar (2019) reported that SWH and maximum wave height (H_{\max}) of ERA5 show a good agreement with the measured buoy data in the coastal (bias ~ 0.29 m) and deep waters (bias ~ 0.18 m) of the BoB. The hourly SWH data for the period from 1979 to 2017 at a location off Gopalpur, India, located at 19.25° N and 85.50° E (Figure 1) is used in the study. The average depth at the chosen ERA5 grid is 497 m, hence not influenced by the shallow water effects. The whole observations are further categorised into seasonal series as pre-monsoon, southwest monsoon and northeast monsoon for seasonal return level estimation.

2.2 Methods

The return level estimation of SWH is carried out using commonly used extreme value distributions named as GEV and GPD. Jenkinson (1955) proposed the three parameter GEV model by coupling three kinds of typical extreme value distributions such as Gumbel, Frechet and Weibull distributions. The return levels are estimated by inverting its Cumulative Distribution Function (CDF) as:

$$X_T = \begin{cases} \beta + \frac{\alpha}{k} \left\{ 1 - \left[-\ln \left(1 - \frac{1}{T} \right) \right]^k \right\}, & k \neq 0 \\ \beta - \alpha \ln \left[-\ln \left(1 - \frac{1}{T} \right) \right], & k = 0 \end{cases} \quad (1)$$

Where α is the scale parameter, β is the location parameter and k is the shape parameter. For $k=0$, the GEV is Gumbel distribution; For $k > 0$, it is Frechet distribution and for $k < 0$ it is Weibull distribution. GEV models are for return level estimation by resampling the data into block maxima series such as MM, SM and AM and taking all observations, i.e., Initial Distribution Method (IDM). The advantage of using the GEV distribution with AM is that few decisions are required during the calculation of the distribution parameters and quantiles.

The main drawback of the GEV block maxima method is that it considers only one peak within each block of the observation period so that it ignores other extreme values that would have significant importance in the EVA. Moreover, the block maxima approach requires a long period of observation for good results. An alternative sampling approach known as POT is used in the study to overcome the inconvenience as mentioned earlier (Jonathan and Ewans, 2013). POT approach considers all observations greater than a significantly high threshold. GPD model incorporates POT observations so that GPD is capable of extracting the maxima in each epoch above a given threshold. The return level estimation by GPD model is carried out using the equation (2):

$$X_T = \begin{cases} k + \frac{\alpha}{k} [1 - (\lambda T)^{-k}], & k \neq 0 \\ k + \alpha \ln(\lambda T), & k = 0 \end{cases} \quad (2)$$

Where λ is, the crossing rate of the threshold u , given as Nu/N . The Nu is the total number of exceedances over the selected threshold, u and N are the total numbers of years of record, and α and k are the fitted GPD parameters to data exceedances $x-u$. In the present study, return levels by GPD model are estimated for all seasons in addition to full-year data series. The parameter for both GEV and GPD are estimated following Probability Weighted Moment (PWM) method and parameters are estimated for each block maxima series and POT samples separately.

2.3 Threshold Selection

The threshold selection for POT resampling is a major step in the GPD analysis. Extreme value analysis requires the observations under consideration should be independent and identical. The independent characteristics of the POT observation expect a suitable combination among threshold value and minimum separation time between the events. That is, if the threshold is high enough, the separation time can be reduced without compromising independence and vice-versa (Palutikof et al., 1999). Furthermore, the threshold must be high enough for the exceedance to converge to GPD as well as the POT should have adequate extreme events left to reach a reliable GPD estimate (Abild et al., 1992). In the present study, the threshold is selected in the following way. Firstly, we chose a tentative threshold for de-clustering the storm events with a minimum separation time. Then we selected several possible thresholds (or a range of a threshold values) with the help of Simple Mean Exceedance (SME) plot and parameter stability plots which are drawn by fitting the data series for the threshold range from tentative threshold to its maxima. The final threshold was obtained by comparing the goodness of fit test result and reliable 100-year return levels. A typical SME plot and parameter stability plot are shown in Figure 2 (a) and (b). Where solid lines are corresponding to shape parameter, k , and dotted is modified scale parameter, α^* . The goodness of fit test plot and return level regression plot are shown in Figure 3b in which the solid line is corresponding to the number of exceedances yearly on average, λ . Kolmogorov–Smirnov (KS) and Cramér–von Mises (CM) are the two goodness of fit tests applied in the study. The symbols in Figure 3 corresponds to the thresholds, where the fit tests are failed.

Return levels corresponding to different return periods are calculated by following both GEV and GP distribution models. The parameters for the distribution such as shape, scale and location are obtained by PWM method following Katz et al. (2002) and Teena et al. (2012) recommendations for wave height observations covering the short period. In the GEV method, both IDM and block

maxima (AM, SM, MM) approaches are considered. Different prevailing seasons in the region are considered for GPD method, and hence seasonal return levels are quantified. Selection of a suitable threshold to obtain POT series is a challenging task because the selected threshold should be high enough for POT series to be independent observation and at the same time, threshold should be small enough to have an adequate number of data points left in the POT series to have an approximate distribution within GPD family. In this study, the thresholds are selected by the help of SME plots, parameter stability plots and goodness of fit test results. In addition to that, both 90th and 95th percentiles also considered as thresholds for comparison.

Figures 2(a) and 2(b) are the typical SME and parameter stability plots used to find a range of thresholds for each POT series. The mean value of each series is taken as a tentative threshold with minimum 48 hours separation between each peak resulted in more than 30 and 10 exceedance observations corresponding to full-year POT and seasonal POT respectively. Then a range of thresholds is selected using SME plot (Figure 2a) by choosing lowest value above which the graph is an approximately straight line and parameter stability plot (Figure 2b) where shape parameters, k , and modified scale parameter, α^* are an almost straight line. A final threshold for each series are selected by KS and CM fit test result (Figure 3a and 3b) in such a way that thresholds have maximum p-value and λ . In the typical plots which correspond to southwest monsoon season, 1.7 to 3 m are selected as the range of thresholds from SME plot and the final threshold of 1.94 m is obtained by comparing with the goodness of fit tests. KS test gives 0.847 and CM gives 0.844 as p-values for the chosen threshold. The crossing rate of the threshold, λ is 17.21 for the selected threshold.

3. Results and Discussion

The maximum value of SWH in the dataset is 9.06 m, which is generated during the extremely severe cyclonic storm Phailin in 2013 followed by 5.50 m during another extremely severe cyclonic storm Hudhud during 2014. The hourly SWH varied from 0.32 to 9.06 m with an average value of 1.31 m. The average seasonal SWH is highest (~1.67 m) during the southwest monsoon and lowest (0.91 m) during the northeast monsoon, whereas the maximum SWH is during northeast monsoon followed by southwest monsoon with 5.00 m. The annual mean SWH varied from 1.165 m in 1982 to 1.356 m in 1995 with an annual average value of 1.28 m and a standard deviation of 0.05 m (Figure 4a). The annual maximum SWH varied from 2.83 to 9.06 m from 1979 to 2017 with an average value of 4.0 m (Figure 4b). Monthly average SWH is maximum in June (~1.83 m), followed by July (~1.77 m) and August (~1.69 m), whereas the monthly average SWH is minimum in January (~0.72 m).

Estimated shape parameters for all of the GEV models are negative, indicating Weibull distributions. The shape parameters for the GPD model shows both pre-monsoon and northeast monsoon season distributions correspond to Weibull distribution and the other two series are Gumbel distributions. The return level estimated using GEV are shown in Figure 5, and that of GPD is in Figure 6. The result indicates that the return level corresponding to the 100-year return period ranges from 5.28 m to 7.88 m with GEV model. IDM underestimate the 100-year return level by ~41% compared to the highest observation in the whole observation period. None of the AM, SM and MM projected more than 8.0 m as 100-year return level. AM series underestimates ~13% compared to the highest observation, which is corresponding to extremely severe cyclonic storm Phailin passed through the study region on 12 October 2013 (Amrutha et al., 2014). The 200-year return level based on GEV model using AM series is 9.02 m, whereas the SWH during the TC Phailin is 9.06 m and it indicates that the return period of the TC Phailin is 205 years. Second Highest storm in the location was another extremely severe cyclonic storm Hudhud passed during October 2014 which generated a maximum wave height of 5.50 m at the study location. The GPD analysis incorporated with seasonality result 100-year return level ranges from 5.41 to 9.08 m using threshold POT. But threshold as the 90th and 95th percentile of observation gives 7.57 m and 7.82 m respectively. It is to be noted that both AM and 95th percentile 100-year return level are almost comparable. The return levels do not follow the seasonal characteristics of prevailing wave climate, highest SWH during southwest monsoon season and relatively calm during the pre-monsoon season, but return levels are higher during the northeast monsoon season due to the occurrence of most of the rare severe extreme events, for example, Phailin and Hudhud, during northeast monsoon season.

3.1 Case studies

A total of fifteen events consist of extremely severe cyclonic storms and deep depressions happened during the study period are considered to study the influence of cyclonic disturbances on the return level estimates. The events considered in the analysis passed within 500 km vicinity of the study location. The general information regarding these extreme events is given in Table 1 and tracks of TC having category greater than 4 are shown in Figure 1. The GEV distribution model with AM and MM series are used for the case studies. We considered four scenarios in which 1) considering the highest two cyclones - Phailin and Hudhud in the data, 2) considering Phailin only and removing Hudhud cyclone from the data, 3) considering Hudhud only and removing Phailin cyclone from the data, and 4) considering all extreme events related to SWH higher than 4 m. Each scenario has three cases in such a way that Case 0 is the actual estimate, Case 1 in which the observation corresponding to highest SWH is replaced with the annual average of the series (i.e., removing the cyclone data),

Case 2 in which lowest observation is replaced with equivalent SWH corresponds to highest observation (i.e., an additional cyclonic effect). Figure 7 shows the graphical representation of different cases involved in the case studies for scenario 4 considering both AM and MM GEV distributions analysis.

As mentioned earlier, the MM series gives 6.43 m and AM series estimates 7.88 m as 100-year return levels. MM series with different scenarios of case 0 shows 2.31% and 7.42% decrease in the return levels corresponding to Phailin and Hudhud respectively (Table 2). AM series with Case 0 and only considering Phailin cyclone shows a slight decrease of 3.18% and considering only Hudhud gives ~25% of decrease in the return levels. The maximum reduction in return levels for each series is different during Case 1. For MM series, the highest underestimation of ~26% occurred when the observations corresponding to all cyclones/depressions which generated SWH greater than 4.0 m is replaced with the average of the whole series and corresponding AM estimates indicate a 48.03% reduction. It is to noted that the MM results when the data points of both Phailin and Hudhud cyclones (Scenario 1) are replaced with average of the series (Case 1) is exactly same as the Phailin data is removed (Scenario 3) from the data and Hudhud data is replaced with average SWH of the year 2014. The reduction in return level for the case of AM series is almost identical with ~30% except scenario 4. Case 2 shows how return levels will behave when an additional cyclonic effect corresponding to the severest cyclone is introduced into the datasets. Generally, the return level should increase when an equal cyclonic effect corresponding to the severest cyclone in the data is added into the observation. But interestingly, the case study that do not include Phailin cyclone in the dataset (Scenario 3) made an underestimation of ~5.18% and ~21.41% for MM and AM series respectively. So it can be inferred that the intensity of the TC will make more significant difference in the return levels than the number of TCs. The maximum increase in the return levels (~24.9%) occurred when AM series with both Phailin and Hudhud cyclones are considered (scenario 1). Scenario 4 in which all cyclonic disturbances with SWH > 4 m made higher variations for both case 1 and case 2 except for case 2 where AM series shows a slight increase in return levels compared to corresponding other scenarios. Removing the effect of prominent extreme events from the data resulted in ~48% reduction in 100-year return levels while considering the AM series. At the same time, the MM series return level reduced only ~26%. The addition of new cyclonic effect equalling to the top extreme events resulted in an increase of 24.04% during MM analysis. But the AM series did not make a marginal difference (~13.02% only) from the actual AM estimates even the series consisted an extra fifteen extreme events out of 39-year data points.

3.2 Influence of the length of the data series

The GEV model with AM series with different blocks of data series is used to understand the behaviour of the length of the data series on the return level estimation. Figure 8 shows the return levels obtained when different block lengths (15 to 39 AM data points) are considered. The study shows that 50-year return levels range from 5.12 to 8.67 m during the different block lengths. The variation in 50-year estimates for data up to 2008 is not significant since it ranged from 5.12 to 5.35 m. The return level using first 25-years block (1979 to 2003) is less than that of the 20-year block (1979 to 1998). The first 35-years (1979-2013) AM series gives a return level of 6.91 m, which is higher than that (6.89 m) of a full 39 years (1979-2017) AM block (Figure 8). The standard deviations (SD) of SWH observations in different blocks higher than 20 years are correlated with return levels (Table 3). That is, the block 2000-2017 have a higher SD of 1.41 m corresponds to higher return levels (8.67 m), and the block 1979-2003 have lesser SD and lower return levels (5.19 m). Similarly, whenever SD is increased among the different blocks, the return levels also increased. The same happened in scenario 4 of the case study. Actual AM series has SD of 1.06 m and which increased to 1.16 m during Case 2 (doubled the cyclonic effect) that leads to an overestimation of ~ 13 %. The SD is lowered to 0.36 m during Case 1, making a considerable reduction of ~ 48 % in the return levels. The same characteristics are seen in the MM series also. Original MM series has an SD of 0.89 m and ~13% increase in SD during case 2 results in 24% overestimation, and 17% reduction in SD during case 1 made ~26% underestimation.

The increased frequency and intensity of extreme events after 2000, seven out of eleven extreme events having SWH higher than 4.5 m in the 39 years observation, especially Hudhud and Phailin, made the 50-year return level overshoot to 8.67 m for 2000-2017 block (Figure 8). The particular block has 1.41 m SD. The return levels using 1979-2012 AM data which do not include both Hudhud and Phailin having SD of 0.62 m do not exceed 6.5 m even for 740 years return period. Therefore, the different case studies and the block analysis suggest that the strength of the extreme events have more influence on the return level than the length of the observation period. The low return level obtained from 1979-2012 period is due to the annual maximum SWH is greater than 4.5 m only in six years and it is higher than 5 m only once during 1979 to 2012 period.

4. Conclusion

The GEV estimation with the block maxima series has higher return levels compared to IDM. The highest 100-year return levels are achieved by northeast monsoon series as ~ 9.08 m. The two extremely severe cyclonic storms- Phailin and Hudhud made a massive difference in the final

estimates. Any other series used in both GEV and GPD model did not cross 8.0 m as 100-year return levels. The prevailing wave climate in the study location possesses a strong seasonality with high seasonal average SWH during the southwest monsoon and low seasonal average SWH during the northeast monsoon and pre-monsoon period. But the 100-year return level is not following this characteristic feature, because prominent TCs are during the northeast monsoon season and hence resulted in higher return levels in the northeast monsoon. The 100-year return levels range from 4.26 m to 9.08 m for different POTs. Only the POT with northeast monsoon series estimated return level more than highest observation. The 95th percentile POT return level corresponding to 100-years is almost the same as that of AM series in GEV with 7.88 m.

Different case studies show how tropical cyclone affects return levels. The maximum increase in 100-year return level is obtained when additional cyclonic effects equalling to Phailin and Hudhud cyclones are introduced to the data. The case studies indicate that the strength of a tropical cyclone can alter the return levels more significantly than the number of tropical cyclones in the data series. The AM estimates with different block size gives insight into the influence of block size on the return levels. The return level is more correlated with the changes in the standard deviation happened while considering different block sizes. The 50-year return level for the data from 2000 to 2017 is 25.8% higher than that for the data from 1979 to 2017 due to the occurrence of cyclones in 2013 and 2014. The present study results apply to a location in the BoB. Since several TCs were formed in the BoB, a similar study should be carried out for the entire BoB.

Acknowledgements

This work would not have been possible without the ERA5 data provided by the ECMWF. ERA5 data is obtained from the ECMWF data server: <http://data.ecmwf.int/data>. The authors acknowledge the Council of Scientific and Industrial Research, New Delhi for funding the research work. We thank Dr. T.M. Balakrishnan Nair, Head, OSISG and Mr Arun Nherakkol, Scientist, INCOIS, Hyderabad and Mr Jai Singh, Technical Assistant, CSIR-NIO for help during the coastal data collection. The authors thank Chilka Development Authority, Odisha for providing the logistics during wave data collection. We thank the two reviewers and the Editor for the suggestions which improved the scientific content of this paper. This manuscript is a part of the Doctoral thesis of the first author registered with Bharathidasan University, Tiruchirappalli and is NIO contribution ----.

References

- Abild J, Andersen EY and Rosbjerg D (1992) The climate of extreme winds at the Great Belt, Denmark, *J. Wind. Eng. Ind. Aerod.*, 41, 521–532. doi:10.1016/0167-6105(92)90458-M.
- Alves, JHGM and Young IR (2003) On estimating extreme wave heights using combined Geosat, Topex/Poseidon and ERS-1 altimeter data, *Appl. Ocean. Res.*, 25, 167–186.
- Amrutha MM, Sanil Kumar V, Anoop TR, Balakrishnan Nair, TM, Arun Nherakkol and Jeyakumar C (2014) Waves off Gopalpur, northern Bay of Bengal during the Cyclone Phailin, *Annales Geophysicae*, 32, 1073-1083. doi:10.5194/angeo-32-1073-2014.
- Anjali Nair M, Sanil Kumar V and Amrutha MM (2018) Spectral wave characteristics in the nearshore waters of northwestern Bay of Bengal, *Pure and Applied Geophysics*, 175, 3111-3136. doi.org/10.1007/s00024-018-1838-5.
- Balaguru KS, Taraphdar LR Leung and GR Foltz (2014) Increase in the intensity of postmonsoon Bay of Bengal tropical cyclones, *Geophys. Res. Lett.*, 41, 3594–3601, doi:10.1002/2014GL060197.
- Breivik ØO J, Aarnes JR, Bidlot A Carrasco and Ø Sætra (2013) Wave extremes in the northeast Atlantic from ensemble forecasts. *J. Climate*, 26, 7525–7540, <https://doi.org/10.1175/JCLI-D-12-00738.1>.
- Chakrabarti and SK (2005) Handbook of Offshore Engineering, Vol-1, *Elsevier ocean engineering book series*, Elsevier London, UK: p 661, ISBN-13: 978-0080443812.
- Goda Y (1988) On the methodology of selecting design wave height. Proc. 21st Int. Conf. on Coastal Eng., Malaga, Spain, *American Society of Civil Engineers*, 899–913.
- Gupta S, Jain I, Johari P and Lal M (2019) Impact of Climate Change on Tropical Cyclones Frequency and Intensity on Indian Coasts. In: Rao P., Rao K., Kubo S. (eds) Proceedings of International Conference on Remote Sensing for Disaster Management. Springer Series in Geomechanics and Geoengineering. *Springer*, Cham, https://doi.org/10.1007/978-3-319-77276-9_32.
- Hersbach H and Dee D (2016) ERA5 reanalysis is in production. *ECMWF Newsletter* 147.7.
- Jenkinson AF (1955) The frequency distribution of the annual maximum (or minimum) values of meteorological elements. *Quarterly Journal of the Royal Meteorological Society*, 81(348), 158-171.
- Jonathan P and Ewans K (2013) Statistical modelling of extreme ocean environments for marine design: a review. *Ocean Eng.* 62, 91–109.
- Katz RW, Parlange MB and Naveau P (2002) Statistics of extremes in hydrology, *Adv. Water Resour.*, 25, 1287–1304, [https://doi.org/10.1016/s0309-1708\(02\)00056-8](https://doi.org/10.1016/s0309-1708(02)00056-8).
- Méndez, FJ, Menéndez M, Luceño A, Medina R and Graham NE (2008) Seasonality and duration in extreme value distributions of significant wave height, *Ocean Eng.*, 35, 131–138.
- Meucci A, Young IR and Breivik Ø (2018) Wind and Wave Extremes from Atmosphere and Wave Model Ensembles. *Journal of Climate*, 31(21), 8819-8842. doi:10.1175/jcli-d-18-0217.1.

Naseef TM and Kumar VS (2017) Variations in return value estimate of ocean surface waves—a study based on measured buoy data and ERA-Interim reanalysis data, *Nat Hazard Earth Sys*, 17.10: 1763.

Naseef TM and Kumar VS (2019) Climatology and trends of the Indian Ocean surface waves based on 39-year long ERA5 reanalysis data, *International Journal of Climatology* (under revision).

Palutikof JP, Brabson BB, Lister DH and Adcock ST (1999) A review of methods to calculate extreme wind speeds. *Meteorological applications*, 6(02), 119-132.

Samayam S, Laface V, Annamalaisamy SS, Arena F, Vallam S and Gavrilovich P (2017) Assessment of reliability of extreme wave height prediction models, *Nat. Hazards Earth Syst. Sci.*, 17, 409–421, <https://doi.org/10.5194/nhess-17-409-2017>.

Sanil Kumar V, Jossia Joseph, Amrutha MM, Jena BK, Sivakholundu KM and Dubhashi KK (2018) Seasonal and interannual changes of significant wave height in shelf seas around India during 1998–2012 based on wave hind cast, *Ocean Engineering* 151 (2018) 127–140.

Soares CG and Scotto M (2004) Application of the rlargest-order statistics for long-term prediction of significant wave height, *Coast Eng.*, 51, 387–394, <https://doi.org/10.1016/j.coastaleng.2004.04.003>.

Takbash A, Young IR., and Breivik Ø (2019) Global Wind Speed and Wave Height Extremes Derived from Long-Duration Satellite Records. *Journal of Climate*, 32(1), 109-126. doi:10.1175/jcli-d-18-0520.1.

Teena NV, Sanil Kumar V, Sudheesh K and Sajeev R (2012) Statistical analysis on extreme wave height, *Nat Hazards.*, 64, 223–236, <https://doi.org/10.1007/s11069-012-0229-y>.

Vinoth J and Young IR (2011) Global estimates of extreme wind speed and wave height. *Journal of Climate*, 24(6), 1647-1665. doi:10.1175/2010JCLI3680.1.

Table 1. The Characteristic features of tropical cyclones considered in the case studies (source: IMD annual cyclone review and RSMC reports)

No.	Cyclonic Storm Or Depression	Date, Time & place of genesis (Lat °N/Lon °E)	Date, Time (UTC) and place of landfall	Estimated maximum wind(kts), Date & Time	Maximum T. No attained	Max SWH observed (m) and the date
01	Cyclone Three over BoB (05-10 December 1981)	05 December 0600 UTC near 10.7/87.5	Near the Indian/Bangladesh border 21.5/87.9 1200UTC of 10 December	75 kts at 0600 UTC of 09 December 17.0/86.7	T 4.0	4.79 10-Dec-1981
02	Severe cyclonic storm- Bangladesh cyclone (21-30 November 1988)	21 November 1800 UTC near 5.6/99.6	Near Bangladesh-west Bengal border 21.8/89.0 1200 UTC of 29 November	110 kts at 1200 UTC of 29 November 21.8/89	T 5.5	4.31 28-Nov-1988
03	Tropical storm over BoB (26-28 July 1992)	26 July 0000 UTC near 20.0/88.5	Near Paradip 20.5/87.0 1800 UTC of 26 July	30 kts at 1800 UTC of 26 July 20.5/87.0	T 2.0	4.26 27-Jul-1992
04	Very severe cyclonic storm over BoB (India cyclone) (07-09 November 1995)	07 November 0000 UTC near 11.3/91.4	Near Andra Pradesh-Orissa border 17.6/83.5 0000 UTC of 09 November	70 kts at 1800 UTC of 08 November 16.5/84.2	T 4.0	4.73 09-Nov-1995
05	Extremely severe cyclonic storm over BoB (21-25 November 1995)	21 November 1200 UTC near 6.5/91.0	Near Cox's bazar (South eastern Bangladesh) 21.5/93.5 0900 UTC of 25 November	102 kts at 1200 UTC of 24 November 16.5/86.5	T 5.5	4.84 24-Nov-1995
06	Extremely severe cyclonic storm over Bob (15-17 October 1999)	15 October 0300 UTC near 13.5/92.5	Near Gopalpur 19.2/85.0 1800 UTC of 17 October	90 kts at 1800 UTC of 17 October 19.2/85.0	T 5.0	4.46 17-Oct-1999
07	Super cyclonic storm (Odisha cyclone) over BoB (25-31 October 1999)	25 October 0600 UTC near 12.0/98.5	Near Paradip 20.5/86.0 0600 UTC of 29 October	140 kts at 0600 UTC of 29 October 20.5/86.0	T 6.0	4.00 28-Oct-1999
08	Depression over BoB (12-13 June 2001)	12 June 0300 UTC near 19.5/87.0		25 kts at 0300 UTC of 12 June 19.5/87.0	T 1.5	4.07 13-Jun-2001
09	Cyclonic Storm Pyarr over	17 September	Kalingapatnam, Andhra	35 kts at 1500 UTC	T 2.5	4.56

	BoB (17-21 September 2005)	0300 UTC near 20.0/90.5	Pradesh 18.5/84.0 0300 UTC of 19 September	of 18 September 19.5/86.5		19- Sep-2005
10	Deep Depression over BoB (02-05 July 2006)	02 July 0000 UTC near 20.0/89.5	Between Paradip and Chanbali, Odisha 20.5/88.0 1200 UTC of 02 July	30 Kts at 1200 UTC of 02 July 20.5/88.0	T 2.0	4.56 03-July-2006
11	Deep Depression over BoB (02-05 August 2006)	02 August 0300 UTC near 20.5/87.5	Between Puri and Gopalpur, Odisha 19.5/85 0000 UTC of 03 July	30 kts at 0000 UTC of 03 July 19.5/85.0	T 2.0	4.76 04- Aug-2006
12	Deep Depression over BoB (28-30 June 2007)	28 June 0000 UTC near 18.5/87.0	Puri, Odisha 19.5/86 0300 UTC of 29 June	30 kts at 0000 UTC of 29 June 19.5/86.0	T 2.0	5.00 29 June 2007
13	Deep Depression over BoB (05-07 August 2007)	05 August 0000 UTC near 20.0/88.5	Between Chandbali and Paradip, Odisha 0130 UTC of 6 August 2007	30 kts at 0000 UTC of 06 August 20.5/87.0	T 2.0	4.92 06-Aug-2007
14	Extremely Severe Cyclonic Storm, "PHAILIN" over the BoB (08-12 October, 2013).	08 October, 0300 UTC near 12.0/96.0	Gopalpur 19.2/84.9 1700 UTC of 12 October 2013	115 kts at 0300 UTC of 11 October near 16.0/88.5	T 6.0	9.05 12-Oct-2013
15	Extremely Severe Cyclonic Storm "Hudhud" over the BoB (07-12 October 2014	07 October, 0300 UTC near 11.5/195.0	Visakhapatnam 17.7/83.3 between 0630 & 0730 UTC of 12 October 2014	100 kts at 1800 UTC of 11 October near 16.4/84.7 .	T 5.0	5.50 12-Oct-2014

Table 2. The 100-year return levels from different case studies. Percentage difference from Case 0 is in the bracket

Data set	Case 0 (m)	Case 1 (m)	Case 2 (m)
Scenario 1. Both Cyclones (Phailin and Hudhud)			
Monthly Maxima	6.43	5.79 (-9.97%)	7.07 (9.98 %)
Annual Maxima	7.88	5.46 (-30.70 %)	9.83 (24.91 %)
Scenario 2. Phailin			
Monthly Maxima	6.28 (-2.31%)	5.81 (-9.61 %)	6.27 (-2.53 %)
Annual Maxima	7.62 (-3.18%)	5.52 (-29.88 %)	9.49 (20.60 %)
Scenario 3. Hudhud			
Monthly Maxima	5.95 (-7.42%)	5.79 (-9.97 %)	6.09 (-5.18 %)
Annual Maxima	5.89 (-25.18%)	5.52 (-29.85 %)	6.18 (-21.41 %)
Scenario 4. All cyclones/depressions with SWH > 4.0 m			
Monthly Maxima	6.43	4.76 (-25.97 %)	7.98 (24.04 %)
Annual Maxima	7.88	4.09 (-48.03 %)	8.90 (13.02 %)

Table 3. Standard deviation and estimated return levels using annual maxima data series for different blocks

No.	Year Blocks	Number of Years	Standard Deviation (SD) m	Return Levels (m)	
				20-Years	50-Years
1	1979-1993	15	0.612	4.84	5.12
2	1979-1998	20	0.598	4.88	5.24
3	1979-2003	25	0.590	4.84	5.19
4	1979-2008	30	0.631	5.00	5.35
5	1979-2013	35	1.084	5.75	6.91
6	1979-2017	39	1.060	5.79	6.89
7	2000-2017	18	1.412	6.65	8.67

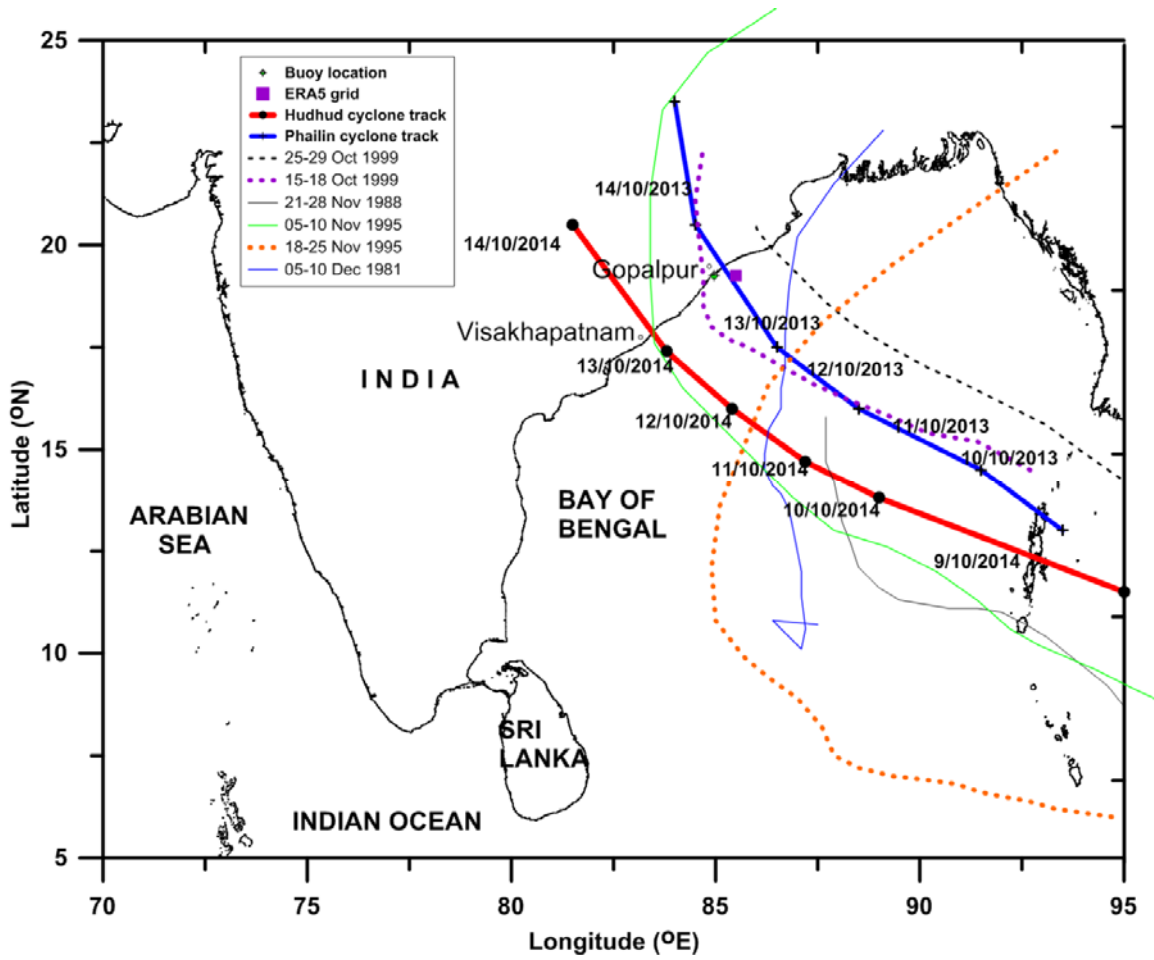


Figure 1. The map shows the study location near Gopalpur in north-western Bay of Bengal. The tracks of the tropical cyclones and the depression during the study period is also shown along with the location of the waverider buoy

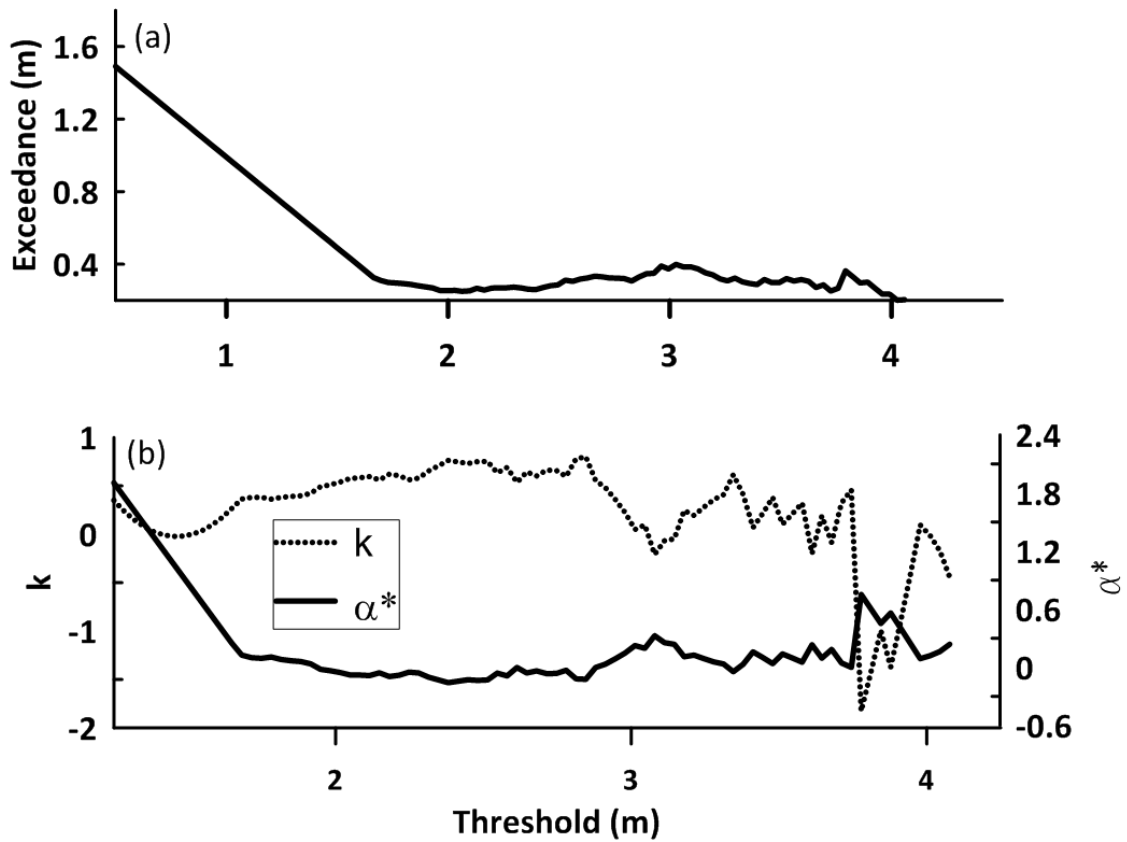


Figure 2. A typical (a) SME and (b) Parameter stability plots used to obtain suitable thresholds for Peak Over Threshold selection. This plot corresponding to the south-west monsoon season and the selected range of thresholds is 1.7 m to 3.0 m

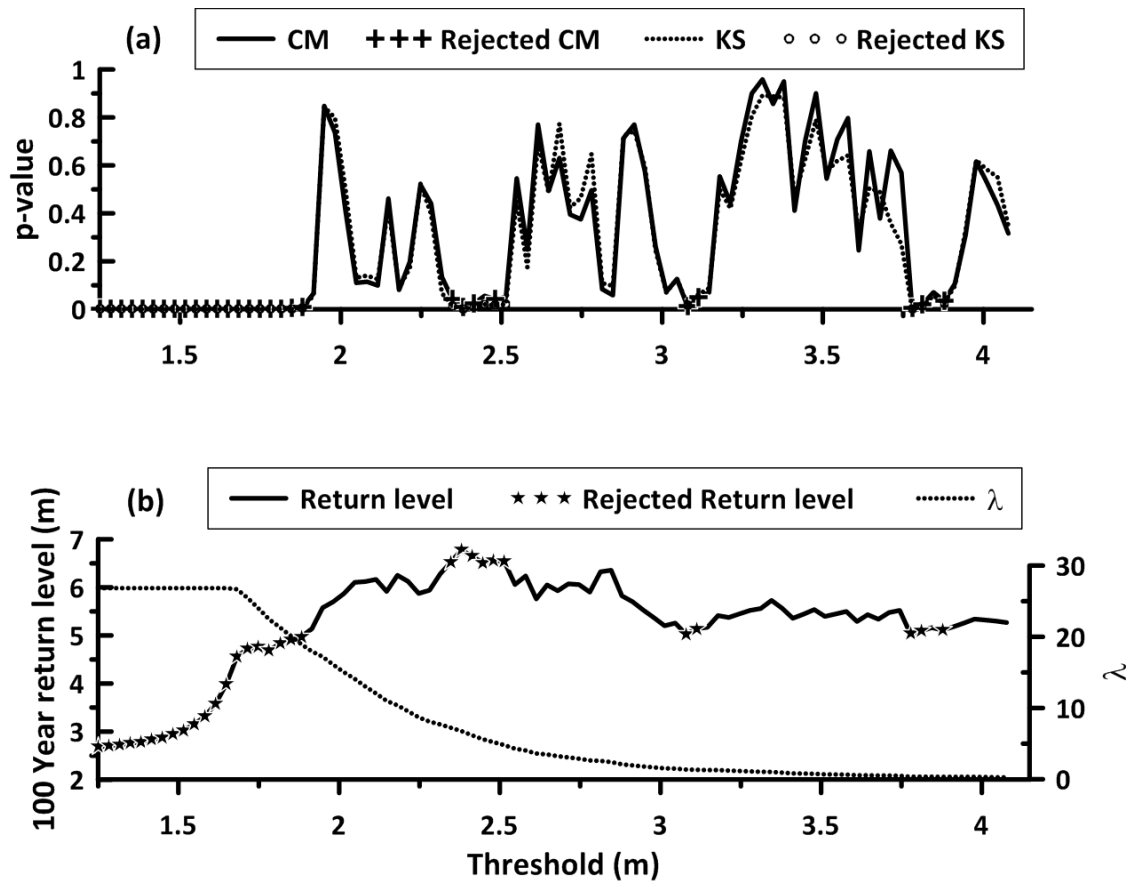


Figure 3. (a) A typical Goodness of fit test and (b) return level regression plots used to obtain suitable thresholds for POT series. This plot corresponding to southwest monsoon season and the threshold selected is 1.94 m

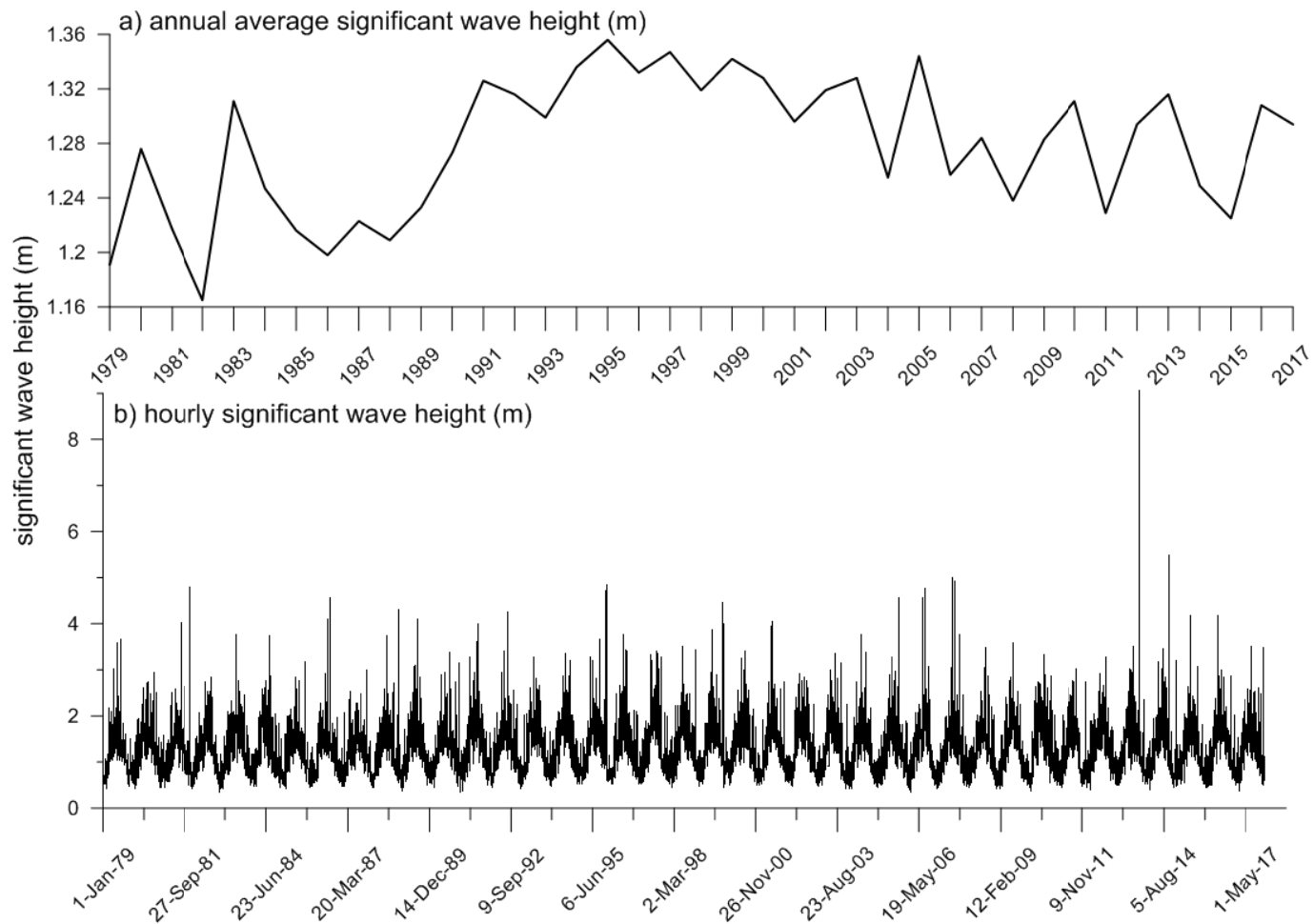


Figure 4. Time series plot of a) annual average significant wave height and b) hourly significant wave height from 1979 to 2017

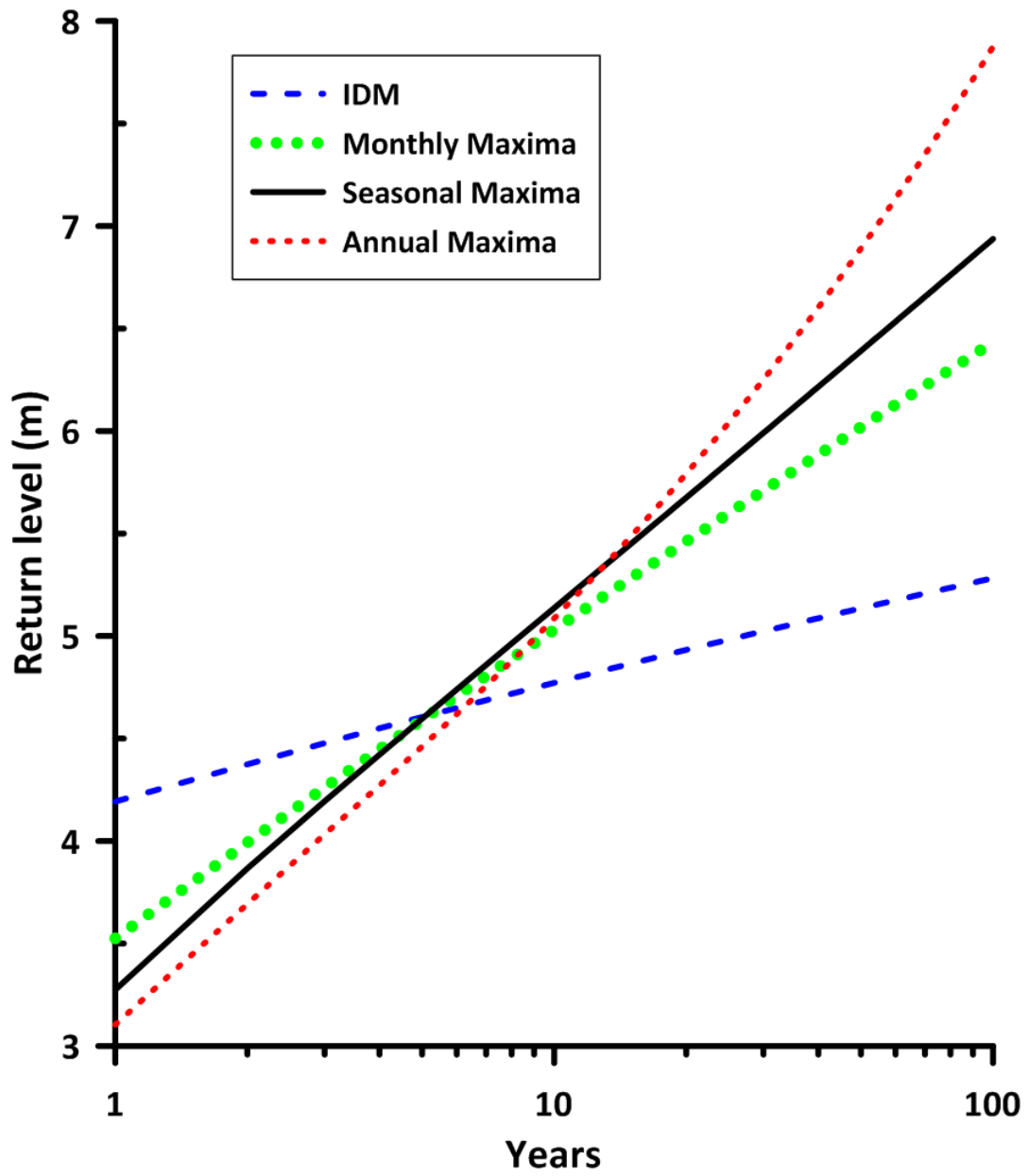


Figure 5. Return levels of significant wave heights by GEV model for different data series

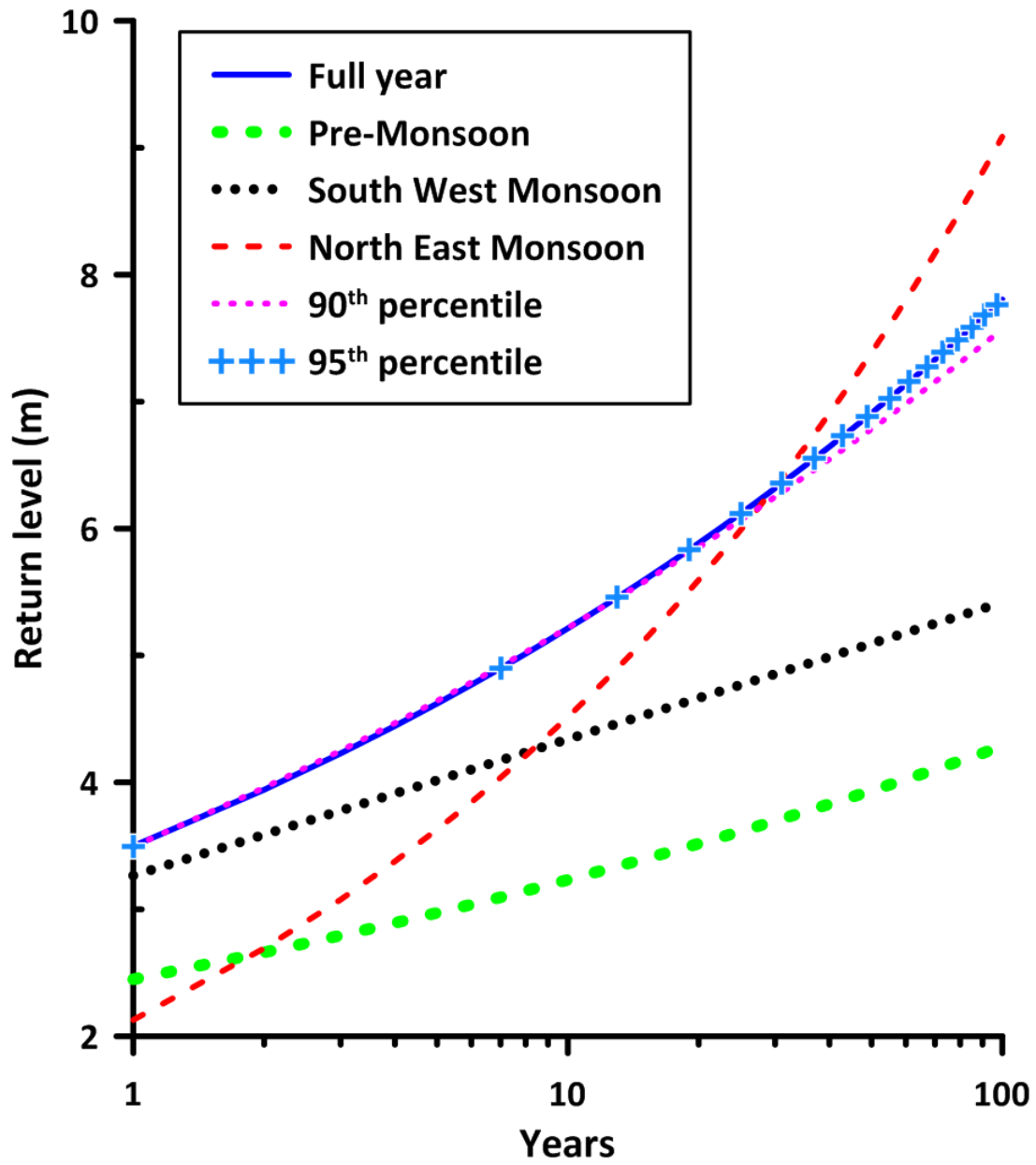


Figure 6. Return levels of significant wave heights by GPD model for different seasons and with POT at 90 and 95 percentiles

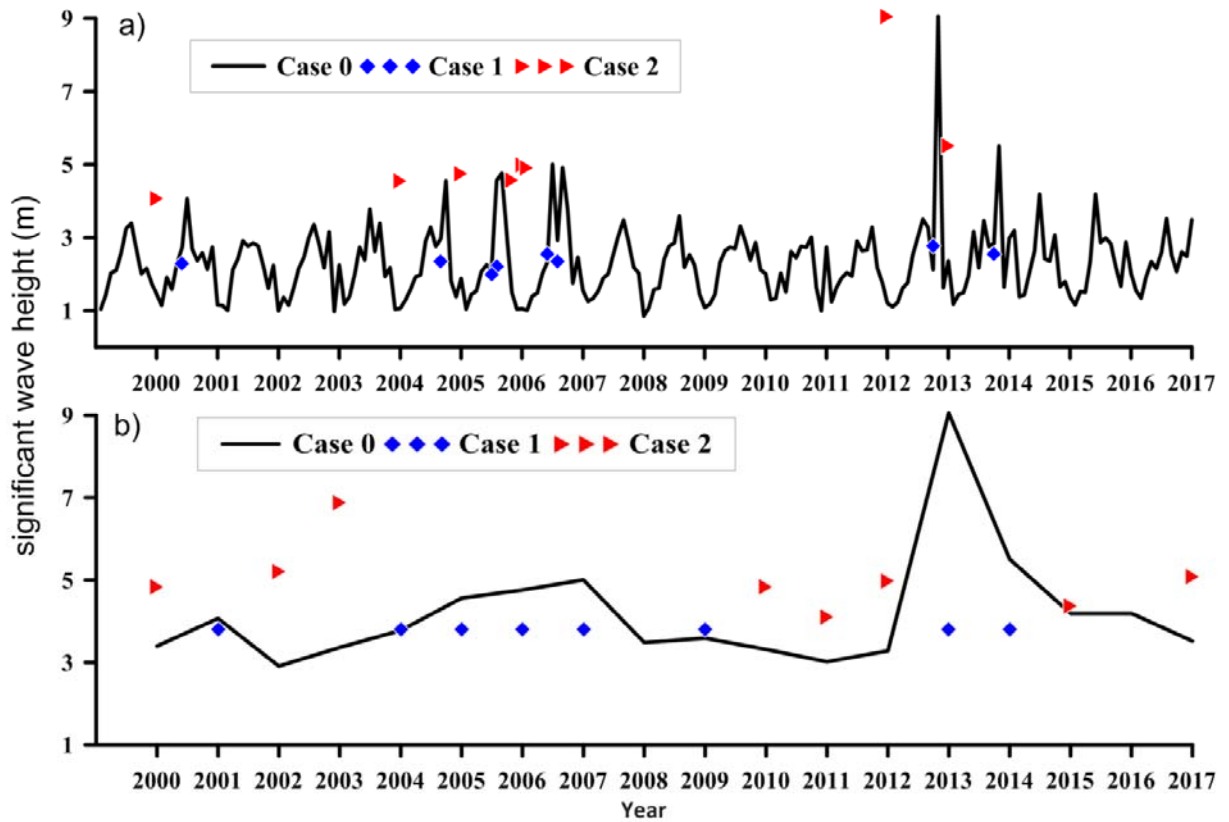


Figure 7. Time series from 2000 to 2017 representing different cases during scenario 4 in the case study. The solid line represents actual monthly maxima data, the solid circle indicates the data points where case 1 applied and the cross symbol corresponds to data points where case 2 is applied. (a) is for monthly maximum series and (b) is annual maxima series

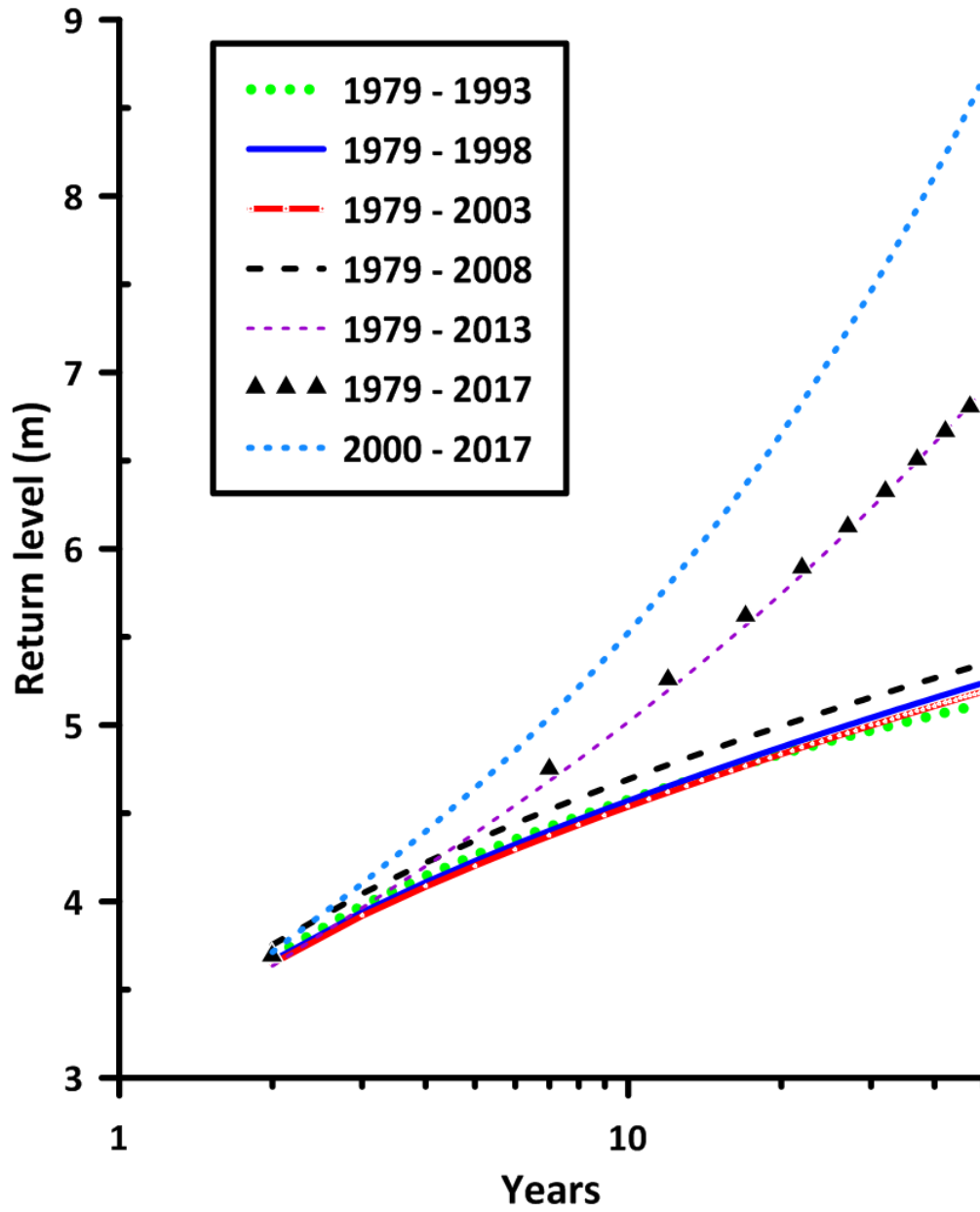


Figure 8: Return levels of different return periods obtained with annual maximum data series considering different blocks periods