Applied Ocean Research 113 (2021) 102750

Contents lists available at ScienceDirect

Applied Ocean Research

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

Spatiotemporal variations and extreme value analysis of significant wave height in the South China Sea based on 71-year long ERA5 wave reanalysis

Jichao Wang [a,*], Jincan Liu [a], Yue Wang [a], Zhihong Liao [b], Peidong Sun [a]

[a] College of Science, China University of Petroleum, Qingdao 266580, China
[b] National Meteorological Information Center, Beijing, 100081, China

ARTICLE INFO

Keywords:
Significant wave height
Long-term trend
Extreme value analysis
ERA5
The South China Sea

ABSTRACT

A thorough understanding of the long-term trends and extreme characteristics in the significant wave height (SWH) contributes greatly to coastal and offshore engineering activities and mitigation of marine disasters. In this study, the annual spatiotemporal variability of the SWH in the South China Sea (SCS) and the return periods for six locations are evaluated based on ERA5 wave reanalysis in a long time period (1950–2020). Long-term trends are estimated by using a popular non-parametric method, the Theil-Sen estimator, and then mapped to show the spatial variability of mean and extreme SWH. Basin-averaged analysis is also performed to investigate the general tendency of the mean SWH in the SCS, with an increasing rate of 0.11 cm/year. The well-known Mann-Kendall test is used to assess the significance of the trends. Significant positive trends in extreme SWH are mainly distributed in the eastern part of the central SCS around the Luzon Strait and the southwestern part of the SCS. In the extreme value analysis, by comparing 2 classical extreme value distribution models combined with 5 sampling methods, the Generalized Pareto Distribution incorporating the Peak Over Threshold (GPD-POT) method has turned out to be suitable for evaluating the return periods in the SCS based on 71-year SWH dataset. Extreme value analysis for different time lengths shows a correlation between the return level and the time span. Small and medium samples may lead to unstable parameter estimation and increased errors. The 100-year return values obtained by GPD-POT using the 71-year wave reanalysis are more credible at P1-P6 with 8.86, 11.79, 11.35, 10.52, 6.50 and 7.71 m, respectively. Both the estimated return periods and the number of extreme events quantified by the Method of Independent Storms (MIS) indicate that extreme events are closely related to the number of tropical cyclones. Seasonally, most extreme events in the SCS occur from June to December, with the summertime maximum SWH situated above the 15th degree of northern latitude and rapidly shifting to south in the fall.

1. Introduction

The South China Sea (SCS) with abundant mineral resources, oil and gas reserves, and fishery resources, plays a vital role in the Indian-Pacific economic circle. Analyzing the wave characteristics and extreme parameters especially the spatiotemporal variability and return periods of significant wave height (SWH) is essential to understand the wave climate in the SCS, which is of high practical value in the design of structures, mitigation of marine disasters and navigation.

Theil-Sen estimator (Sen, 1968; Theil, 1992), together with the Mann-Kendall test (Aydoğan and Ayat, 2018; Mann, 1945), is widely used for trend analysis of meteorological and oceanic problems. The probability distribution model and sampling method and are two key parts in the extreme value analysis (EVA). The investigation of the extreme behavior is fitted by models, such as Generalized Extreme Value (GEV) (Jenkinson, 1955) and Generalized Pareto Distribution (GPD) (Iii, 1975), which infer the distribution of the extreme values (Coelho et al., 2008). The fundamental sampling methods include the Initial Distribution Method (IDM) (Naseef and Kumar, 2020; Shamji et al., 2020), the Annual Maxima (AM) method (Li et al., 2016; Ruggiero et al., 2010), the Method of Independent Storm (MIS) (Cook, 1982; Devis--Morales et al., 2017) and the Peak Over Threshold (POT) (Goda et al., 2001) method.

A bulk of researchers have put numerous efforts in EVA and trend analysis in the SCS. Cao et al. (2018) picked four different locations to compare the applicability of GEV-AM and GPD-POT using the 20-year

* Corresponding author.
E-mail address: wangjc@upc.edu.cn (J. Wang).

https://doi.org/10.1016/j.apor.2021.102750
Received 18 March 2021; Received in revised form 3 June 2021; Accepted 4 June 2021
Available online 24 June 2021

(1995–2014) hindcast data obtained from the third-generation wave model Simulating WAves Nearshore (SWAN) (Booij et al., 1999). The performance of two sampling methods AM and POT over a longer period (1975–2014) was investigated by Shao et al. (2018) through SWAN. Wang et al. (2018) applied the GPD model to analyze the seasonality of extreme wave climate, and the SWH data from 1976 to 2014 was simulated by the wave numerical model WAVEWATCH III (Kim and Lee, 2018). Osinowo et al. (2016) focused on the spatial variability of the long-term trends in yearly/seasonal/monthly mean SWH and 99th percentile SWH, using 30-year hindcast data derived from the WAVEWATCH III. Based on ERA-Interim reanalysis (Dee et al., 2011) from 1979 to 2016 supplied by European centre for Medium-Range Weather Forecasts (ECMWF), Luo and Zhu (2019) explored the spatial variability of the trends in the annual and seasonal maximum SWH.

Although the previous studies have made some achievements, the length of dataset is at most 40 years. Nevertheless, a longer period of wave parameters allows for a more systematic and comprehensive of trend analysis. Generally, the stability of the parameters to be estimated is better and the reliability of the return values is stronger when longer time series are adopted in EVA. In addition, ECMWF's ERA-Interim wave reanalysis has been widely applied for extremes and trends studies (Agarwal et al., 2013; Brink et al., 2005; Portilla et al., 2013; Shanas and Kumar, 2014; Shanas and Sanil Kumar, 2014), whereas its latest generation ERA5 reanalysis (Hersbach et al., 2020) has rarely been used for relevant analyses in the SCS. Benefitting from the developments in model physics, core dynamics and data assimilation, ERA5 has numerous strengths compared to ERA-Interim, such as much higher spatial and temporal resolution, information on variation in quality over space and time, better global balance of precipitation and evaporation. One of the objectives of this study is to characterize extreme behavior of SWH by using a reliable EVA method at specific locations and assess the long-term (71 years) variation of wave climates in the whole SCS based on the highly enhanced ERA5 wave reanalysis. Another objective is to investigate the variation of return values for different data series lengths.

The paper is organized as follows. In the next section, study area, selected locations and ERA5 SWH dataset are highlighted. The methodologies of trend analysis and EVA are described in Section 3. In Section 4, a detailed discussion about the results is presented. Finally, the main conclusions are summarized in Section 5. Some materials are listed in appendix A.

2. Study area and materials

2.1. Study area and selected locations

Connected to the Pacific Ocean to the east through narrow straits or waterways, the SCS is a large semi-enclosed marginal sea with frequent tropical cyclones. As shown in **Fig. 1**, the SCS has a large variation in water depth (Ris et al., 1999; Sun et al., 2020), containing deep-sea areas and shallow water near the coastal areas. Mainly due to the influence of the Australasian monsoon system, the weather and climate of the SCS is strongly dominated by the southwest monsoon in summer and the northeast monsoon in winter (Chu et al., 2004). The present study focused on the region ranging from 9 to 24°N and 105–120°E in the SCS (as shown in **Fig. 1** with red rectangular box).

Six locations (P1-P6) denoted by red dots in **Fig. 1** are considered to perform the EVA. The location coordinates, water depths and corresponding specific areas in the SCS are presented in **Table 1**. It can be observed that the six locations are scattered far apart in the study region, and the water depths vary from shallow to deeper with great changes.

2.2. Dataset

The ERA5 wave reanalysis used in this study is the fifth generation of ECMWF, which embodies a thorough record of global atmospheric, land surface and ocean waves since 1950 (Hersbach et al., 2020). Currently data is available from 1950, split into Climate Data Store entry for 1950–1978 (preliminary back extension) and from 1979 onwards. In contrast with the ERA5 product from 1979 onwards, the preliminary data does suffer from tropical cyclones that are sometimes unrealistically intense. Although the ERA5 SWH has a spatial resolution of 0.5° with hourly output throughout, a higher resolution can be achieved by interpolation. In this paper, the ERA5 SWH data from 1950 to 2020 used for spatiotemporal trends and extreme value analysis are provided by ECMWF, with a spatial resolution of 0.125°

The ERA5 performance is evaluated and validated against in-situ SWH using statistical metrics, correlation coefficient (R), root mean squared error (RMSE) and bias. R is employed to measure the linear correlation between reanalyses and observations. The RMSE reflects the average deviation and the bias is the discrepancy between the two datasets. Good performance is indicated by RMSE and bias as close to 0 as possible while R close to 1. The three statistic indices are defined in the following equations

$$R = \frac{\sum_{i=1}^{N}\left[\left(C_i - \overline{C}\right)\left(O_i - \overline{O}\right)\right]}{\sqrt{\sum_{i=1}^{N}\left(C_i - \overline{C}\right)^2 \sum_{i=1}^{N}\left(O_i - \overline{O}\right)^2}} \quad (1)$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(C_i - O_i)^2} \quad (2)$$

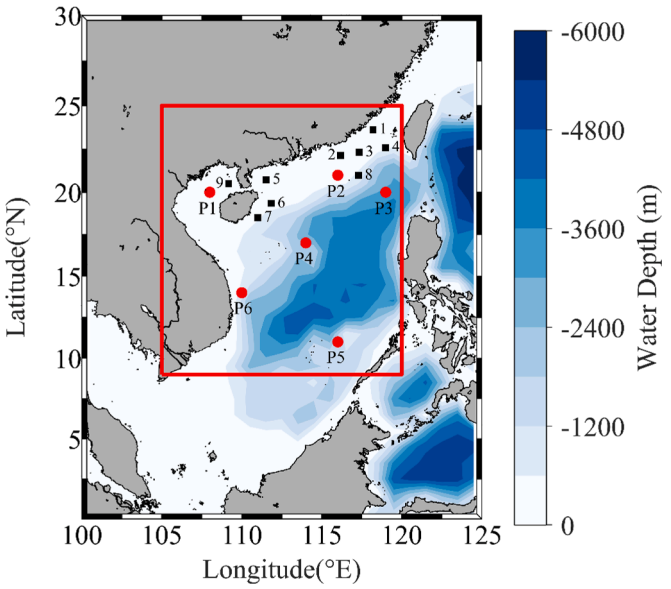

Fig. 1. The study area and bathymetry. The red circular dots marked with P1-P6 represent feature locations (the corresponding SWH series is from internationally available data sets provided by ECMWF). The black square dots present 9 buoy stations. The bathymetric data is sourced from ETOPO1.

Table 1
Statistics for the study location characteristics.

Location	Specific area	Lat. (°N)	Lon. (°E)	Water depth (m)
P1	Beibu Gulf	20	108	49
P2	Dongsha Islands	21	116	128
P3	Luzon Strait	20	119	3032
P4	Zhongsha Islands	17	114	3053
P5	Nansha Islands	11	116	2671
P6	Coast of Vietnam	14	110	1274

$$Bias = \frac{1}{N}\sum_{i=1}^{N}(C_i - O_i) \tag{3}$$

where N is the total number of data, C_i denotes the ERA5 reanalysis value, $\overline{C}$ indicates the mean value, O_i denotes the buoy observation and $\overline{O}$ represents the mean of the measurements.

The in-situ buoy measurements marked by black squares in Fig. 1, spanning a full year of 2020, are utilized to validate the accuracy of ERA5 SWH. The hourly archived observational data of nine buoys are provided by the China Meteorological Administration. The SWH of nearest grid point from ERA5 for each buoy is compared with the corresponding observations. Fig. 2 presents a scatter plot between buoy and ERA5 SWH of all nine locations, illustrating that they have a good agreement. The R is high up to 0.95, and the RMSE is within a reasonable limit (less than 10% of the maximum buoy SWH). The ERA5 slightly underestimate the SWH on average with the bias of −0.07 m. Detailed information for each buoy, including coordinates and validation indicators of SWH, is exhibited in Table 2.

3. Methodology

3.1. Trend analysis

There are two kinds of methods to achieve trend estimation, namely parametric and nonparametric methods. Parametric trend detection methods are applicable to independent and normally distributed datasets, while non-parametric method datasets are not restricted by normal distribution. In this study, two nonparametric methods, the line of best fit and Theil-Sen estimator, are employed to assess the trends (Aydoğan and Ayat, 2018; Sen, 1968; Theil, 1992). Moreover, the Theil-Sen method uses the Mann-Kendall Test to estimate the significance of the trends (Aydoğan and Ayat, 2018; Mann, 1945; Shamji et al., 2020).

3.2. Extreme value analysis

The primary objective of EVA is to provide an estimation of the N-year return value, i.e., the level that may be exceeded once every N year averagely. The GEV and GPD are two mainstream models adopted for fitting the sampling series in EVA, where the well-known Probability Weighted Moments (PWM) method is applied to estimate model parameters (Hosking et al., 1985; Hosking and Wallis, 1987; Katz et al., 2002; Teena et al., 2012).

The sampling methods used in this study are Initial Distribution Method (IDM), Block Maximum (BM) method, Peak Over Threshold (POT) method and Method of Independent Storm (MIS).

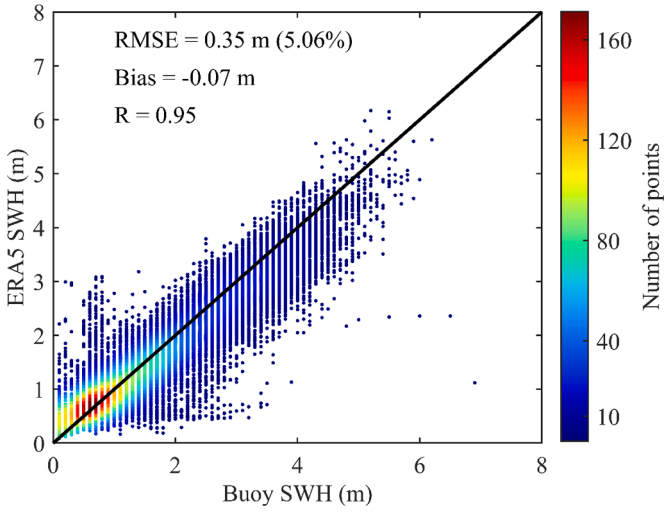

Fig. 2. The scatter plot between the ERA5 SWH and Buoy SWH.

Table 2
Statistics for buoy characteristics and validation of SWH.

Buoy ID	Lat. (°N)	Lon. (°E)	Data Number	R	RMSE (m)	Bias (m)
1	23.63	118.20	8302	0.93	0.55 (8.53%)	−0.29
2	22.15	116.17	3059	0.97	0.28 (5.30%)	−0.11
3	22.33	117.34	6156	0.97	0.32 (6.42%)	−0.15
4	22.60	119.00	8096	0.94	0.34 (6.36%)	0.04
5	20.73	111.53	3506	0.96	0.30 (4.38%)	−0.08
6	19.35	111.84	6251	0.98	0.24 (4.12%)	−0.06
7	18.51	111.00	6250	0.90	0.25 (5.53%)	−0.05
8	20.99	117.29	5466	0.96	0.38 (6.14%)	0.02
9	20.50	109.17	8186	0.88	0.25 (7.81%)	0.06
Total	–	–	55,272	0.95	0.35 (5.06%)	−0.07

The sample obtained by IDM (Naseef and Kumar, 2020; Shamji et al., 2020) includes the entire initial data. The BM (Devis-Morales et al., 2017; Ferreira and Haan, 2015) sample is established by first separating the observational periods into non-overlapping periods equally and then selecting the maximum value for each, such as Annual Maxima (AM) or Monthly Maxima (MM). POT method (Goda et al., 2001; Shao et al., 2018) takes into account the extreme values of a sample that exceeds a given threshold. The basic idea of MIS (Cook, 1982; Devis-Morales et al., 2017; Harris, 1999; Palutikof et al., 1999) is to record the continuous SWH series that exceeds a certain threshold as an independent storm, and then restrict attention to the maximum value of each independent storm to form an extreme sample. The selection of threshold is introduced in Section 3.2.3, and the threshold values used in the MIS are the same selected in the POT.

3.2.1. GEV distribution

The cumulative distribution function (CDF) of GEV distribution for a given random variable is shown as

$$GEV(X;\alpha;\beta;k) = \begin{cases} \exp\left\{-\left[1+k\left(\frac{X-\beta}{\alpha}\right)\right]^{-\frac{1}{k}}\right\}, & k \neq 0 \\ \exp\left\{-exp\left[\left(\frac{-(X-\beta)}{\alpha}\right)\right]\right\}, & k = 0 \end{cases} \tag{4}$$

where α, β, and k correspond to the parameters of the scale, location, and shape of the distribution, respectively. Specifically, the scale parameter is restricted to be greater than 0. Three typical types of extreme value distributions distinguished by shape parameters are coupled in GEV, including Gumbel (k=0), Frechet (k>0) and Weibull distributions (k<0) (Jenkinson, 1955).

The return level X_T of GEV is derived by inverting its CDF as shown

$$X_T = \begin{cases} \beta + \frac{\alpha}{k}\left\{1-\left[-\ln\left(1-\frac{1}{T}\right)\right]^{k}\right\}, & k \neq 0 \\ \beta - \alpha ln\left[-\ln\left(1-\frac{1}{T}\right)\right], & k = 0 \end{cases} \tag{5}$$

where T refers to the return period.

3.2.2. GPD distribution

The CDF for the GPD model is

$$GPD(X;\alpha;\beta;k) = \begin{cases} 1-\exp\left\{-\left[1+k\left(\frac{X-\beta}{\alpha}\right)\right]^{-\frac{1}{k}}\right\}, & k\neq 0 \\ \exp\left\{-exp\left[\left(\frac{-(X-\beta)}{\alpha}\right)\right]\right\}, & k=0 \end{cases} \quad (6)$$

where α, β, and k represent the parameters of the scale, location, and shape of the distribution, respectively.

Then the returned value X_T estimation based on the GPD distribution is carried out using the Eq. (7):

$$X_T = \begin{cases} k+\frac{\alpha}{k}\left[1-(\lambda T)^{-k}\right], & k\neq 0 \\ k+\alpha ln(\lambda T), & k=0 \end{cases} \quad (7)$$

where $\lambda = \frac{N_u}{N}$. N_u is the total amounts of exceedance above the certain threshold u, and N denotes the total number of years on record.

3.2.3. Threshold selection in the POT

The appropriate selections of threshold and minimum time span play a vital role in the frequency and exceedance estimates of EVA results (Beguería, 2005; Luceño et al., 2006; Méndez et al., 2006). In this paper, the dispersion index method is employed to determine the threshold (Shamji et al., 2020), where dispersion index refers to the ratio between the variance and the expectation of the number of peaks. An acceptable peak separation should give a dispersion index near one. The thresholds corresponding to P1-P6 displayed in Fig. 3 are 4.11, 5.18, 5.60, 5.12, 3.56 and 3.94 m, respectively. A minimum time span of 5 days is adopted in the POT.

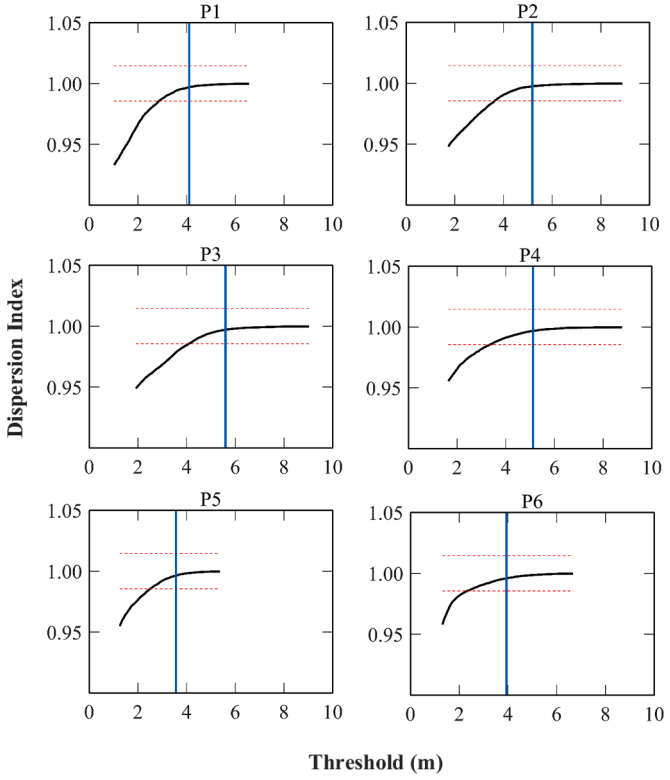

Fig. 3. Threshold selection at P1-P6. Solid blue lines indicate threshold selected by the dispersion index. The red dotted lines represent the upper and lower limits of 95% confidence level.

4. Results and discussion

4.1. Annual trends analysis

Basin-averaged analysis implies averaging the values of each node spatially, and it is usually carried out when performing trend analysis for the entire basin. Based on the basin-averaged annual mean SWH time series, the best-fit line and Theil-Sen estimator methods are employed to evaluate the temporal variability of the mean sea states in the whole SCS. In Fig. 4, the long-term trend of the basin-averaged annual mean SWH is shown along with the 95% confidence interval of slope. The basin-averaged trend for the entire SCS is positive, with rates of 0.11 and 0.10 cm/year derived from the Theil-Sen estimation method and the best-fit line method, respectively. Although the best-fit line method is a crude estimator to assess the trend, its result is similar to that obtained from Theil-Sen estimator (Aydoğan and Ayat, 2018; Vanem and Walker, 2013). The corresponding Mann-Kendall test for Theil-Sen estimation has a value of 2.819 and the confidence level is 99.5%.

The spatial patterns of the long-term trends in annual mean SWH derived from the Theil-Sen estimation and the best-fit line method are illustrated in Fig. 5. The maps obtained by the two methods are similar. Based on the Theil-Sen method, the annual mean SWH exhibits a positive trend in most of the regions, especially in the Luzon Strait and the southwestern part of the SCS, while the negative trend is mainly found in the Beibu Gulf. The slope variation of the annual average trend ranges from −0.1 to 0.3 cm/year, which is narrower than the −0.4 to 1 cm/year of Osinowo et al. (2016).

The confidence level of trend is estimated by the Mann-Kendall significance test, and Fig. 6 displays the spatial pattern of the trends in annual mean SWH with confidence levels greater than 75%. Positive trends are significant in most of the areas, except for a part of the northern SCS. The Beibu Gulf is an important area with considerable negative trends, especially the nearshore region, with a confidence level of 99.9%. For positive trends, the majority of the SCS has confidence levels greater than 95%.

The spatial variability analysis of trends and their significances are conducted for annual 99th percentile SWH (Fig. 7) and annual maximum SWH (Fig. 8) as well. The significant positive trends are concentrated in the southwest of SCS and around Luzon strait in the east of the central SCS, while the significant negative area is mainly in the northern corner of the SCS. For the above significant regions, the trend of 99th percentile SWH coincides well with Osinowo et al. (2016). The trend of annual maximum SWH around the Luzon strait agrees with Luo and Zhu (2019). Compared with the significant trend map of annual mean SWH in Fig. 6, the maps of extreme conditions occupy a much smaller area of significant positive trends, and the significant negative area shifts from Beibu Gulf to the northern corner of the SCS.

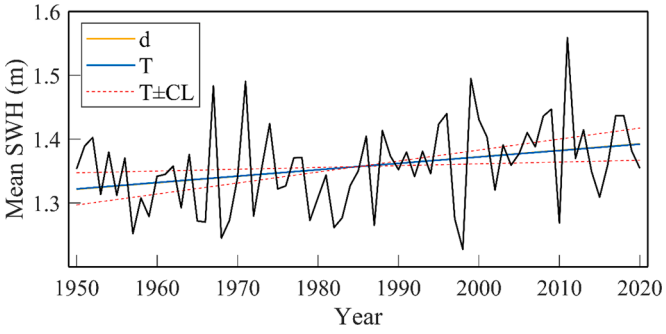

Fig. 4. Trends in basin-averaged annual mean SWH. Solid yellow line (d) and blue line (T) refer to the best-fit line method and Theil-Sen estimator, respectively. The red dotted line.

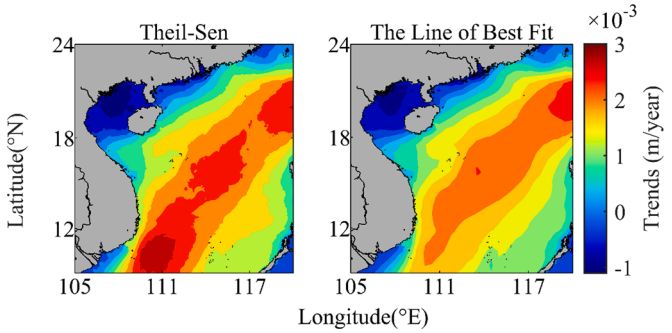

Fig. 5. Trends for annual mean SWH (m/year) based on Theil-Sen estimator (left) and best-fit line method (right).

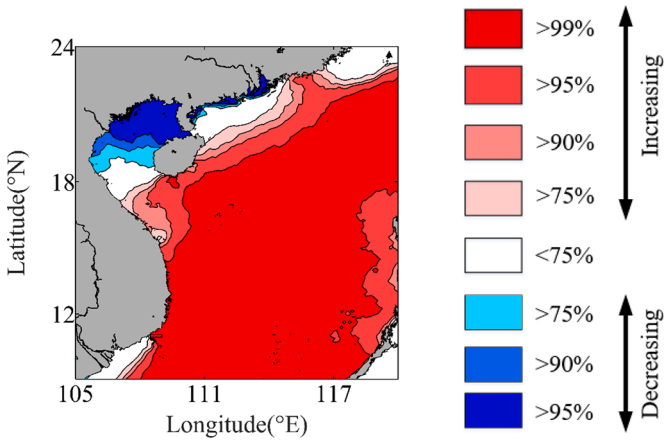

Fig. 6. The significance of the long-term trends in annual mean SWH based on Mann-Kendall test for different confidence levels.

4.2. Monthly spatial pattern of maximum SWH

The maximum SWH (MSWH) is identified per month throughout the period in Fig. 9, which could visualize seasonal variations for the most extreme events in the SCS. MSWH has a clear seasonal variability, generally 1–7 m in most parts of the SCS in winter (December-February; DJF) and spring (March-May; MAM), 5–11 m in large parts during summer (June-August; JJA) and autumn (September-November; SON). In particular, values of 15 m or higher occur in July, with a minimum value of 0.83 m for MSWH. The spatial pattern of MSWH in this domain is consistent well with the trajectories of most tropical cyclones (TCs). TCs activity in the SCS is largely concentrated in the JJA and SON. Location-wise, TCs are mainly located north of 15°N in JJA, while in SON, TCs recede rapidly to the south. The MSWH over 8 m in JJA and SON evidences the pathway of most storms in the SCS.

4.3. Return periods and qualification of extreme events

Considering the variability of ocean parameters over time, extreme samples can be selected within a probabilistic framework when designing and operating offshore structures. The length of the ERA5 SWH data (71 years) is sufficient for prediction in marine applications where 100-year return periods are typically calculated.

The fitted model has a pivotal impact on the results, so it is necessary to ensure the model is well-fitted (Coles, 2001). The fitness of the sampling data to the extreme value distribution model is first measured by the statistical parameter RMSE. The error series is formed by the residuals between the empirical cumulative distribution of the sampling data and the theoretical model cumulative distribution. Table 3 lists the RMSE results for different EVA methods from P1-P6, where the RMSE is

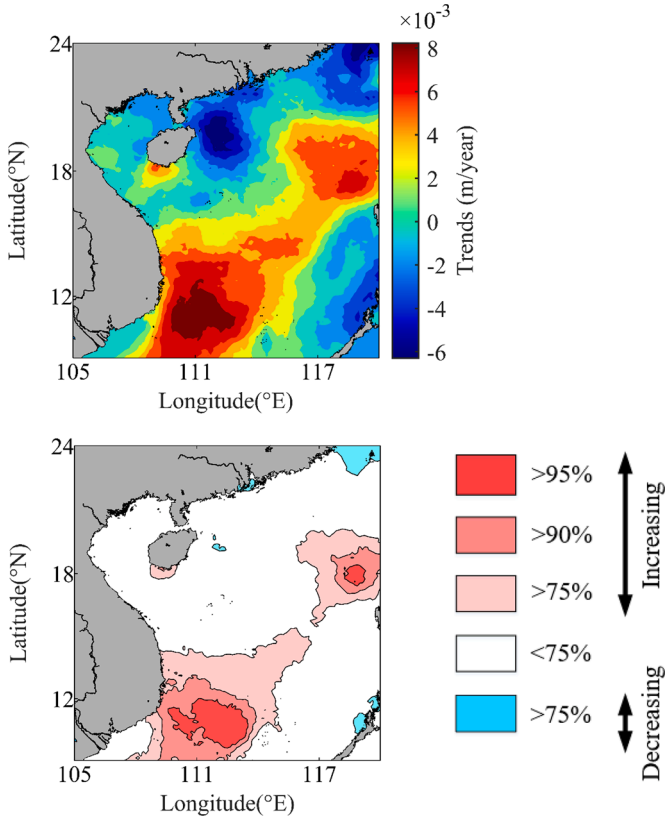

Fig. 7. Long-term trends for annual 99th percentile SWH (m/year) based on the Theil-Sen method (top) and corresponding Mann-Kendall test for different confidence levels (bottom).

less than 0.5 m for all locations, indicating that the sampling data have a good approximation to the distribution model.

However, the fit of the distribution tails in EVA deserves more attention. So, in this case, we further analyze the feasibility by drawing quantile-quantile plot (QQ-plot), which can present a comparison between the empirical and the fitted distribution function. The closer the scatters are to the diagonal line in the plot, the better the model fits. Fig. 10 shows the QQ-plots of different EVA methods at each position. In GEV-IDM and GPD-MIS, the scatter points with large values are above the diagonal and have a distinct distance from the line. In the QQ-plots of GEV-MM, they are almost below the diagonal, which is unacceptable because it may lead to an underestimation of extreme values in the prediction. Although the scatters of both GEV-AM and GPD-POT are in good agreement with the diagonal, GPD-POT is a more reliable method for reasonable extrapolation in the SCS as long as the thresholds are selected appropriately (Shao et al., 2018). In conclusion, the GPD-POT has turned out to be the most reliable method in the estimation of extreme values in the SCS. The same results are also seen in a shorter data series from 1950 to 1980 (Fig. 11).

Fig. 12 presents the return periods of SWH from 1950 to 2020 for different EVA methods at each selected position. The return values increase rapidly in the first 50 years and then the growth slows down apparently, which shows the asymptotic form of the distribution models. The return values vary considerably when different EVA methods are used. GEV-IDM generates the lowest estimates of return values. The estimates of GEV-MM are much lower than the other three methods (except P6), which demonstrates a good consistence with the previous analysis of QQ-plots. Approximate results can be found in GEV-AM, GPD-POT and GPD-MIS. It is worth noting that at P6, the estimated return levels are all lower than the max SWH of years on record. Selecting the 100-year return value obtained by GPD-POT as the design standard, the results of P1-P6 are 8.86, 11.79, 11.35, 10.52, 6.50 and

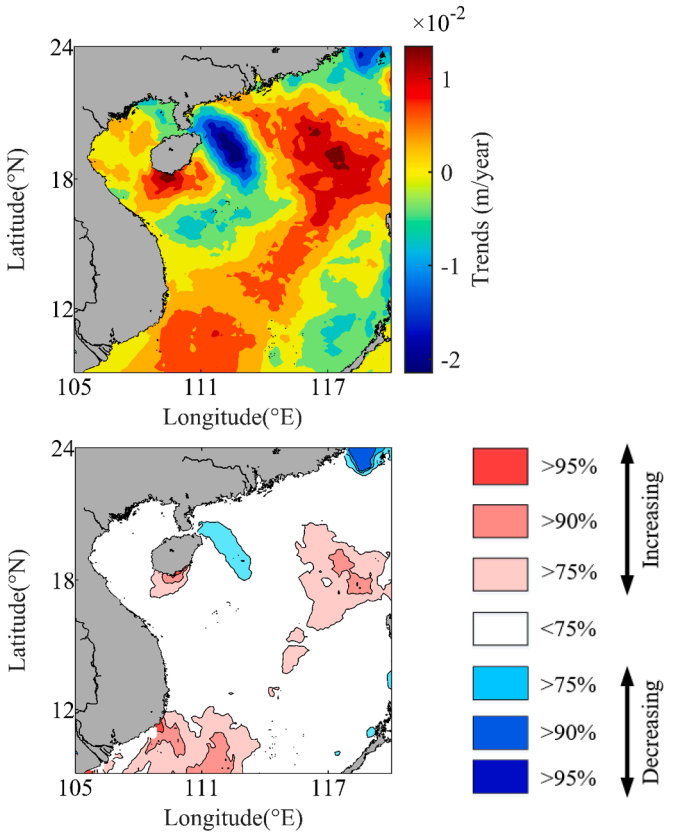

Fig. 8. Long-term trends for annual maximum SWH (m/year) based on the Theil-Sen method (top) and corresponding Mann-Kendall test for different confidence levels (bottom).

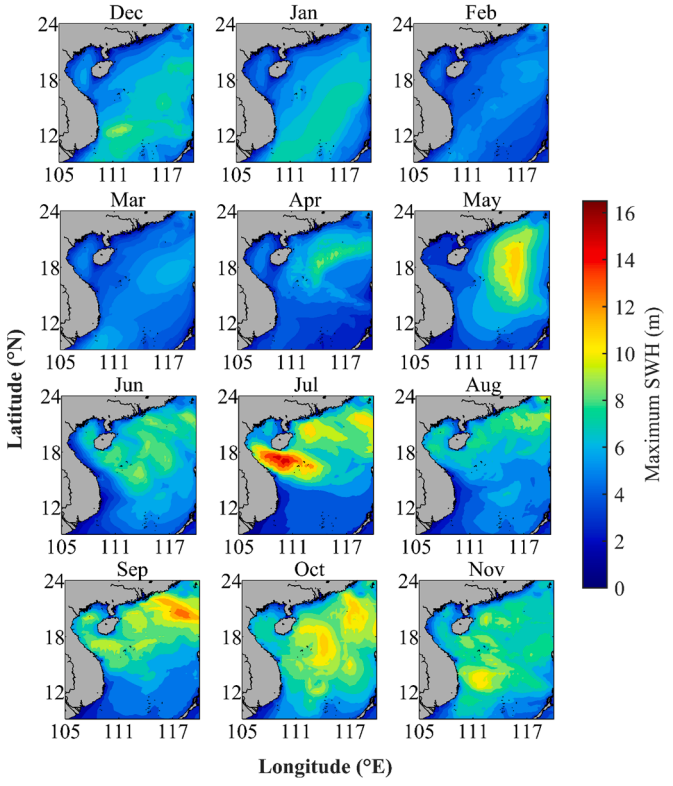

Fig. 9. Monthly spatial pattern of MSWH (m) from 1950 to 2020 in the SCS.

Table 3
The RMSE (m) for different EVA methods at P1-P6.

Method	P1	P2	P3	P4	P5	P6
GEV-IDM	0.054	0.134	0.152	0.152	0.070	0.229
GEV-AM	0.104	0.200	0.199	0.199	0.060	0.161
GEV-MM	0.065	0.165	0.192	0.192	0.071	0.247
GPD-POT	0.075	0.266	0.134	0.134	0.098	0.140
GPD-MIS	0.144	0.463	0.152	0.152	0.171	0.128

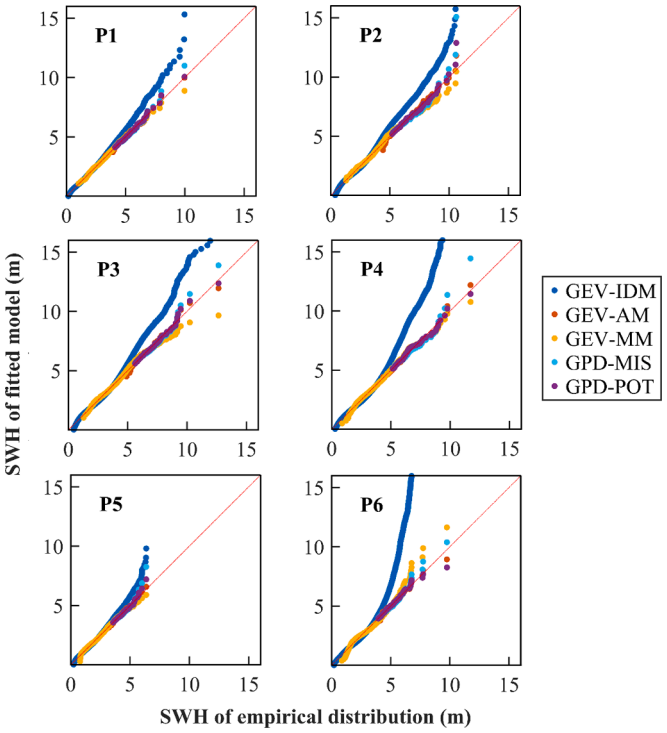

Fig. 10. The QQ-plots for different EVA methods at P1-P6 from 1950 to 2020. The blue, tangerine, yellow, purple, and cyan scatters represent the GEV-IDM, GEV-AM, GEV-MM, GPD-POT, and GPD-MIS methods, respectively. The red dotted lines are referred to the diagonals.

7.71 m, respectively.

4.3.1. *Influence of the length of the data series*

The GPD-POT method with different blocks of data series at each position is applied to explore the effect of the length of the data series on the return value estimation. Table 4 shows the standard deviations (SD) and 100-year return levels of sampling observations when different block lengths are considered. The SD of SWH series in different blocks are correlated with 100-year return levels, which means that when the SD increases, the 100-year return level generally increases as well. There is also an association between block length and 100-year return values. Probably because with the increasing length of the data series, more information on the extremes is available and the overall value of the sampling extremum data increases.

4.3.2. *Qualification of extreme events*

The MIS method is also imposed to quantify the extreme events and their duration for each point from 1950 to 2020, whose temporal variations are analyzed in annual cycle. For comparison, the threshold for all points is set as the maximum threshold 5.60 m. Table 5 lists the total number of extreme wave events per month and its average duration at each position. The majority of extreme events occur during June to December, with few occurrences in the remaining months. The number of extreme events at the location of P1, P2 and P3 (almost at the same

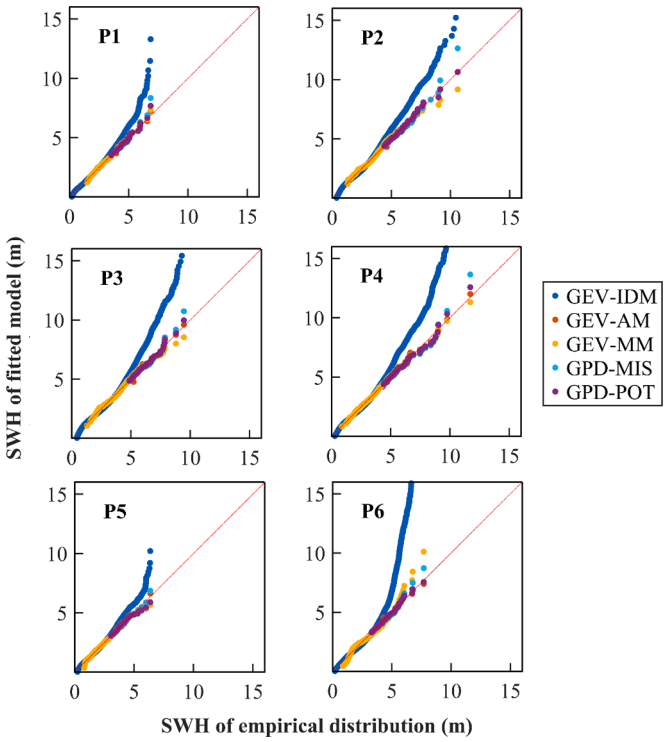

Fig. 11. The QQ-plots for different EVA methods at P1-P6 from 1950 to 1980. The blue, tangerine, yellow, purple, and cyan scatters represent the GEV-IDM, GEV-AM, GEV-MM, GPD-POT, and GPD-MIS methods, respectively. The red dotted lines are referred to the diagonals.

latitude) is 30, 100 and 170 respectively, the different values of which can be attributed to their geographical locations. P3 is near the Luzon Strait with many TCs passing through, while P1 is inside of the Beibu Gulf, where most TCs are blocked to some extent by Hainan Island during their westward journey. Given that P2, P4 and P5 are at almost the same longitude, the possible reason causing the disparity in the total number of extreme events among them is that TCs rarely pass over the southern SCS. The amount of TCs can also be used to explain the differences among the estimated return values at different location. For P4, although there are few extreme events in May, the duration of each is long, up to 33 h.

5. Conclusions

Adequate knowledge of the spatiotemporal variations and the return periods for significant wave height (SWH) has considerable influence on

Table 4
The estimated 100-year return level and corresponding standard deviation (in the parentheses) by using GPD-POT method for different time blocks at P1-P6 (Both units: m).

Year blocks	1950-1970	1950-1980	1950-1990	1950-2000	1950-2010	1950-2020
P1	6.05 (0.677)	7.05 (0.696)	7.20 (0.768)	8.59 (0.882)	8.85 (0.975)	8.86 (0.948)
P2	8.90 (1.139)	10.06 (1.178)	9.84 (1.112)	10.03 (1.136)	11.72 (1.260)	11.79 (1.300)
P3	8.79 (0.976)	9.46 (0.970)	10.67 (1.082)	10.55 (1.065)	10.82 (1.086)	11.35 (1.197)
P4	8.86 (0.980)	9.63 (1.033)	9.27 (1.082)	10.23 (1.191)	10.41 (1.148)	10.52 (1.192)
P5	4.97 (0.526)	5.59 (0.589)	5.66 (0.577)	6.31 (0.604)	6.44 (0.613)	6.50 (0.622)
P6	6.74 (0.860)	7.22 (0.900)	7.41 (0.869)	7.41 (0.839)	7.94 (0.914)	7.71 (0.896)

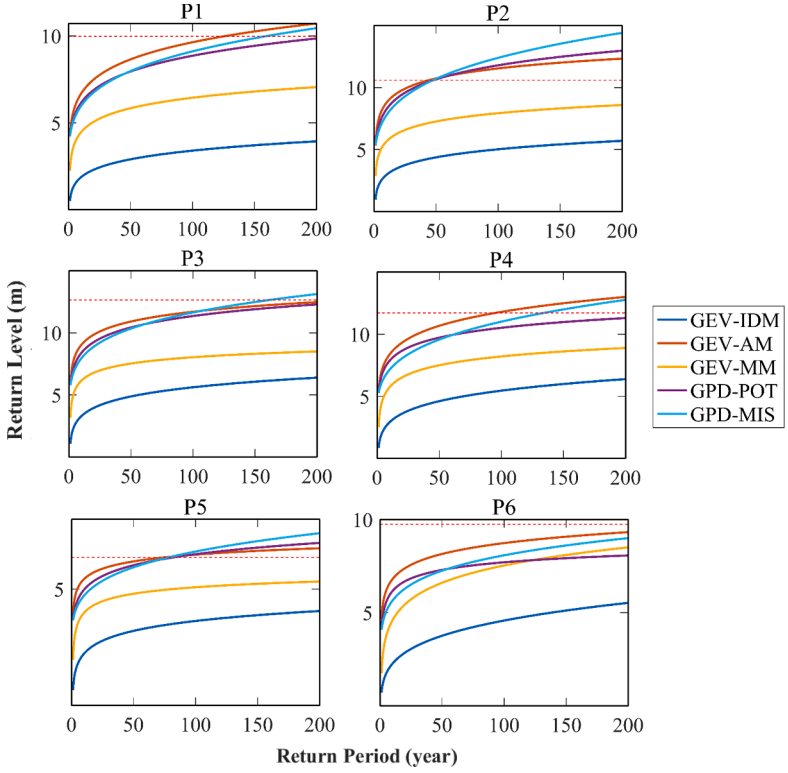

Fig. 12. The estimated return level from 1950 to 2020 for different EVA methods at P1-P6 positions. The blue, tangerine, yellow, purple, and cyan lines represent the GEV-IDM, GEV-AM, GEV-MM, GPD-POT, and GPD-MIS methods, respectively. The red dotted lines are referred to the max SWH of years on record.

Table 5
Total number of extreme wave events given per month during the analyzed time period and its average duration (in the parentheses) at P1-P6 (Unit: h).

Month	P1	P2	P3	P4	P5	P6	Total
Jan	0	0	4(11.75)	2(17.00)	1(20.00)	2(11.50)	9
Feb	0	0	0	0	0	0	0
Mar	0	0	1(7.00)	0	0	0	1
Apr	0	2(17.50)	3(8.67)	1(17.00)	0	0	6
May	1(12.00)	4(14.00)	2(22.00)	2(33.00)	1(16.00)	1(15.00)	11
Jun	2(15.00)	9(6.33)	9(8.89)	5(17.00)	1(2.00)	0	26
Jul	4(5.25)	10(16.30)	14(21.14)	8(8.88)	0	0	36
Aug	4(14.50)	12(16.17)	14(12.21)	6(9.33)	0	0	36
Sep	7(11.14)	21(16.52)	19(15.58)	9(20.56)	0	2(7.00)	58
Oct	9(8.44)	27(17.52)	32(23.03)	31(19.58)	1(17.00)	8(20.88)	108
Nov	3(15.00)	13(19.92)	44(18.95)	27(16.56)	2(9.00)	16(11.06)	105
Dec	0	2(9.50)	31(15.16)	23(19.26)	4(11.00)	10(11.70)	70
Total	30	100	173	114	10	39	466

the engineering applications and marine activities. In this study, the trends analysis and extreme value analysis (EVA) for SWH by using the classical methods based on ERA5 wave reanalysis are performed in the South China Sea (SCS) from 1950 to 2020.

The basin-averaged annual mean SWH evaluated by the Theil-Sen estimator reveals an increasing trend with the rate of 0.11 cm/year. Long-term annual trend analysis shows that trend tendency and significance vary in the whole SCS, spatially. The annual mean SWH is found to be increasing in the large areas of the SCS, except for the Beibu Gulf, which has a negative trend. The significant positive trends of annual extreme SWH are mainly distributed in two areas, around Luzon strait in the east of the central SCS and the southwest of SCS. Compared against mean sea states, significant positive trends cover smaller areas and negative trends shift to the north corner of SCS for severe sea states.

By comparing different well-known EVA methods, the Generalized Pareto Distribution incorporating the Peak Over Threshold (GPD-POT) method is suitable for 71-year data to estimate the return levels in the SCS, and the 100-year return values corresponding to the feature locations P1-P6 are 8.86, 11.79, 11.35, 10.52, 6.50 and 7.71 m respectively. The return level is correlated with the changes in the standard deviation and the time span of the sampling data while considering different block sizes. For the 71-year data, the results of the extreme value analysis are more credible due to the vast amount of extreme value information available and the good stability of the parameter estimates. The disparities in return levels and total number of extreme events are mainly related to the tropical cyclone events. The spatial patterns of the maximum SWH have a distinct seasonal variation, in which the higher values are primarily concentrated in summer and autumn. In the future, a further consideration of the wind is required in the SCS to give a more comprehensive understanding on the trends and extreme conditions.

CRediT authorship contribution statement

Jichao Wang: Conceptualization, Funding acquisition, Investigation, Methodology, Project administration, Supervision, Formal analysis. **Jincan Liu:** Software, Validation, Visualization, Writing - original draft, Writing - review & editing. **Yue Wang:** Writing - original draft, Writing - review & editing. **Zhihong Liao:** Data curation, Resources, Writing - review & editing. **Peidong Sun:** .

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This study is supported by the National Key Research and Development Program of China (2018YFC1406200 and 2017YFC1405600), the Shandong Provincial Natural Science Foundation (ZR2020MD060) and the Fundamental Research Funds for the Central Universities (19CX05003A-5). Useful comments of anonymous reviewers on draft version of this paper are highly appreciated.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.apor.2021.102750.

References

Agarwal, A., Venugopal, V., Harrison, G.P., 2013. The assessment of extreme wave analysis methods applied to potential marine energy sites using numerical model data. Renew. Sustain. Energy Rev. 27, 244–257. https://doi.org/10.1016/j.rser.2013.06.049.

Aydoğan, B., Ayat, B., 2018. Spatial variability of long-term trends of significant wave heights in the Black Sea. Appl. Ocean Res. 79, 20–35. https://doi.org/10.1016/j.apor.2018.07.001.

Beguería, S., 2005. Uncertainties in partial duration series modelling of extremes related to the choice of the threshold value. J. Hydrol. 303, 215–230. https://doi.org/10.1016/j.jhydrol.2004.07.015.

Booij, N., Ris, R.C., Holthuijsen, L.H., 1999. A third-generation wave model for coastal regions: 1. Model description and validation. J. Geophys. Res. Oceans 104, 7649–7666. https://doi.org/10.1029/98JC02622.

van Brink, H.W., van den, Können, G.P., Opsteegh, J.D., Oldenborgh, G.J., Burgers, G., 2005. Estimating return periods of extreme events from ECMWF seasonal forecast ensembles. Int. J. Climatol. 25, 1345–1354. https://doi.org/10.1002/joc.1155.

Cao, H.J., Yi, F., Feng, W.B., 2018. Estimate of the extreme wave height in the South China Sea using GPD method. IOP Conf. Ser. Earth Environ. Sci. 189, 052035 https://doi.org/10.1088/1755-1315/189/5/052035.

Chu, P.C., Qi, Y., Chen, Y., Shi, P., Mao, Q., 2004. South China Sea wind-wave characteristics. Part I: validation of wavewatch-III Using TOPEX/Poseidon data. J. Atmospheric Ocean. Technol. 21, 1718–1733. https://doi.org/10.1175/JTECH1661.1.

Coelho, Ferro, T, C., Stephenson, D., Steinskog, D., 2008. Methods for exploring spatial and temporal variability of extreme events in climate data. J. Clim. 21, 2072. https://doi.org/10.1175/2007JCLI1781.1.

Coles, S., 2001. An Introduction to Statistical Modeling of Extreme Values. Springer Series in Statistics. Springer, London.

Cook, N.J., 1982. Towards better estimation of extreme winds. J. Wind Eng. Ind. Aerodyn. 9, 295–323. https://doi.org/10.1016/0167-6105(82)90021-6.

Dee, D.P., Uppala, S.M., Simmons, A.J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M.A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A.C.M., Berg, L.van de, Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A.J., Haimberger, L., Healy, S.B., Hersbach, H., Hólm, E.V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A.P., Monge-Sanz, B.M., Morcrette, J.-J., Park, B.-K., Peubey, C., Rosnay, P.de, Tavolato, C., Thépaut, J.-N., Vitart, F., 2011. The ERA-Interim reanalysis: configuration and performance of the data assimilation system. Q. J. R. Meteorol. Soc. 137, 553–597. https://doi.org/10.1002/qj.828.

Devis-Morales, A., Montoya-Sánchez, R.A., Bernal, G., Osorio, A.F., 2017. Assessment of extreme wind and waves in the Colombian Caribbean Sea for offshore applications. Appl. Ocean Res. 69, 10–26. https://doi.org/10.1016/j.apor.2017.09.012.

Ferreira, A., Haan, L.de, 2015. On the block maxima method in extreme value theory: PWM estimators. Ann. Stat. 43, 276–298. https://doi.org/10.1214/14-AOS1280.

Goda, Y., Konagaya, O., Takeshita, N., Hitomi, H., Nagai, T., 2001. Population distribution of extreme wave heights estimated through regional analysis. In: Coastal Engineering 2000. Presented at the 27th International Conference on Coastal Engineering (ICCE), American Society of Civil Engineers. Sydney, Australia, pp. 1078–1091. https://doi.org/10.1061/40549(276)83.

Harris, R.I., 1999. Improvements to the 'method of independent storms'. J. Wind Eng. Ind. Aerodyn. 80, 1–30. https://doi.org/10.1016/S0167-6105(98)00123-8.

de Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., Chiara, G. D., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R.J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., Thépaut, J.-.N., 2020. The ERA5 global reanalysis. Q. J. R. Meteorol. Soc. 146, 1999–2049. https://doi.org/10.1002/qj.3803.

Hosking, J.R.M., Wallis, J.R., 1987. Parameter and quantile estimation for the generalized pareto distribution. Technometrics 29, 339–349. https://doi.org/10.1080/00401706.1987.10488243.

Hosking, J.R.M., Wallis, J.R., Wood, E.F., 1985. Estimation of the generalized extreme-value distribution by the method of probability-weighted moments. Technometrics 27, 251–261. https://doi.org/10.1080/00401706.1985.10488049.

Iii, J.P., 1975. Statistical inference using extreme order statistics. Ann. Stat. 3, 119–131. https://doi.org/10.1214/aos/1176343003.

Jenkinson, A.F., 1955. The frequency distribution of the annual maximum (or minimum) values of meteorological elements. Q. J. R. Meteorol. Soc. 81, 158–171. https://doi.org/10.1002/qj.49708134804.

Katz, R.W., Parlange, M.B., Naveau, P., 2002. Statistics of extremes in hydrology. Adv. Water Resour. 25, 1287–1304. https://doi.org/10.1016/S0309-1708(02)00056-8.

Kim, T.R., Lee, J.H., 2018. Comparison of high wave hindcasts during typhoon bolaven (1215) using SWAN and WAVEWATCH III Model. J. Coast. Res. 1096–1100. https://doi.org/10.2112/SI85-220.1.

Li, J., Chen, Y., Pan, S., Pan, Y., Fang, J., Sowa, D.M.A., 2016. Estimation of mean and extreme waves in the East China Seas. Appl. Ocean Res. 56, 35–47. https://doi.org/10.1016/j.apor.2016.01.005.

Luceño, A., Menéndez, M., Méndez, F.J., 2006. The effect of temporal dependence on the estimation of the frequency of extreme ocean climate events. Proc. R. Soc. Math. Phys. Eng. Sci. 462, 1683–1697. https://doi.org/10.1098/rspa.2005.1652.

Luo, Y., Zhu, L.S., 2019. Investigation of trends in extreme significant wave heights in the South China Sea. Aquat. Ecosyst. Health Manag. 22, 53–64. https://doi.org/10.1080/14634988.2018.1467194.

Mann, H.B., 1945. Nonparametric tests against trend. Econometrica 13, 245–259. https://doi.org/10.2307/1907187.

Méndez, F.J., Menéndez, M., Luceño, A., Losada, I.J., 2006. Estimation of the long-term variability of extreme significant wave height using a time-dependent Peak Over Threshold (POT) model. J. Geophys. Res. Oceans 111. https://doi.org/10.1029/2005JC003344.

Naseef, T.M., Kumar, V.S., 2020. Influence of tropical cyclones on the 100-year return period wave height—a study based on 39-year long ERA5 reanalysis data. Int. J. Climatol. 40, 2106–2116. https://doi.org/10.1002/joc.6321.

Osinowo, A., Lin, X., Zhao, D., Wang, Z., 2016. Long-term variability of extreme significant wave height in the South China Sea. Adv. Meteorol. 2016, 1–21. https://doi.org/10.1155/2016/2419353.

Palutikof, J.P., Brabson, B.B., Lister, D.H., Adcock, S.T., 1999. A review of methods to calculate extreme wind speeds. Meteorol. Appl. 6, 119–132. https://doi.org/10.1017/S1350482799001103.

Portilla, J., Sosa, J., Cavaleri, L., 2013. Wave energy resources: wave climate and exploitation. Renew. Energy 57, 594–605. https://doi.org/10.1016/j.renene.2013.02.032.

Ris, R.C., Holthuijsen, L.H., Booij, N., 1999. A third-generation wave model for coastal regions: 2. Verification. J. Geophys. Res. Oceans 104, 7667–7681. https://doi.org/10.1029/1998JC900123.

Ruggiero, P., Komar, P.D., Allan, J.C., 2010. Increasing wave heights and extreme value projections: the wave climate of the U.S. Pacific Northwest. Coast. Eng. 57, 539–552. https://doi.org/10.1016/j.coastaleng.2009.12.005.

Sen, P.K., 1968. Estimates of the regression coefficient based on Kendall's Tau. J. Am. Stat. Assoc. 63, 1379–1389. https://doi.org/10.1080/01621459.1968.10480934.

Shamji, V.R., Aboobacker, V.M., Vineesh, T.C., 2020. Extreme value analysis of wave climate around Farasan Islands, southern Red Sea. Ocean Eng. 207, 107395 https://doi.org/10.1016/j.oceaneng.2020.107395.

Shanas, P.R., Kumar, V.S., 2014. Comparison of ERA-Interim waves with buoy data in the eastern Arabian Sea during high waves. IJMS 437. July 2014.

Shanas, P.R., Sanil Kumar, V., 2014. Temporal variations in the wind and wave climate at a location in the eastern Arabian Sea based on ERA-Interim reanalysis data. Nat. Hazards Earth Syst. Sci. 14, 1371–1381. https://doi.org/10.5194/nhess-14-1371-2014.

Shao, Z., Liang, B., Li, H., Lee, D., 2018. Study of sampling methods for assessment of extreme significant wave heights in the South China Sea. Ocean Eng. 168, 173–184. https://doi.org/10.1016/j.oceaneng.2018.09.015.

Sun, Z., Zhang, H., Xu, D., Liu, X., Ding, J., 2020. Assessment of wave power in the South China Sea based on 26-year high-resolution hindcast data. Energy 197, 117218. https://doi.org/10.1016/j.energy.2020.117218.

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

Theil, H., 1992. A rank-invariant method of linear and polynomial regression analysis (Eds.). In: Raj, B., Koerts, J. (Eds.), Henri Theil's Contributions to Economics and Econometrics: Econometric Theory and Methodology, Advanced Studies in Theoretical and Applied Econometrics. Springer, Netherlands, Dordrecht, pp. 345–381. https://doi.org/10.1007/978-94-011-2546-8_20.

Vanem, E., Walker, S.-.E., 2013. Identifying trends in the ocean wave climate by time series analyses of significant wave heightdata. Ocean Eng. 61, 148–160. https://doi.org/10.1016/j.oceaneng.2012.12.042.

Wang, Z., Li, S., Dong, S., Wu, K., Yu, H., Wang, L., Li, W., 2018. Extreme wave climate variability in South China Sea. Int. J. Appl. Earth Obs. Geoinf. 73, 586–594. https://doi.org/10.1016/j.jag.2018.04.009.