WAVE ANALYSIS FOR THE DEVELOPMENT OF THIEN NGA - HAI AU OIL FIELD

Pham Thi Thanh Nga[(1)], Le Quoc Huy[(1)], Pham Van Tien[(1)], Doan Thi Thu Ha[(1)], Tran Thanh Thuy [(1)], Nguyen Lam Anh[(2)], Bui Trong Han[(2)]

[(1)]The Viet Nam Institute of Meteorology Hydrology and Climate Change

[(2)]Science Research and Engineering Institute

Received: 2 August 2023; Accepted: 25 August 2023

Abstract: *This study utilizes the SWAN model to simulate wave data, subsequently comparing it with ERA5 and CMEMS wave reanalysis datasets. Our analysis demonstrates that the SWAN-simulated wave data are more consistent with observations compared to ERA5 and CMEMS reanalyzed wave data. The predominant wave directions are Northeast and Southwest, with occurrence frequencies of 43.87% and 24.24%, and maximum significant wave heights of 7.72 m and 3.28 m, respectively. Significant wave heights for 50 and 100 year return periods are 8.17 m and 8.53 m in the NE direction, 3.44 m and 3.56 m in the SW direction, 3.78 m and 3.84 m during the Southwest monsoon, 10.09 m and 10.45 m during the Northeast monsoon, and 12.8 m and 13.6 m during storms, respectively. These results are instrumental for the structural design of the Thien Nga - Hai Au oil field in the Southern region of the East Sea Viet Nam.*

Keywords: *Extreme wave analysis, SWAN model, offshore structures.*

1. Introduction

Beyond the conventional oil and gas sector, Viet Nam is making substantial investments in renewable energy, in alignment with global trends. As outlined in Resolution 55-NQ/TW dated 11th February 2020, which lays out the framework for Viet Nam's national energy development strategy by 2030 with a vision toward 2045, and in line with Viet Nam's commitment to achieving net-zero emissions by 2050, there's a strong emphasis on the development of the offshore renewable energy sector, set to take center stage until the middle of the 21st century. This strategic shift will entail the construction and operation of numerous offshore facilities for harnessing wind, wave, tide, and current energy. Unlike their inland counterparts, offshore structures are subject to the additional complexities of ocean hydrodynamic processes, making them a critical factor in the design and engineering considerations for these projects [1]. Ocean hydrodynamic factors and meteorological factors coalesce to exert a combined impact on ofshore structures, necessitating their simultaneous considerations in the design process. This integration is reliant on accurate metocean data, which, to assure the reliability of offshore facility design criteria, should be both comprehensive and dependable. Essential metocean factors encompass wind, storm, air temperature, rainfall, seawater levels, wave properties, ocean currents, sea temperature, seawater salinity, and density. Key attributes to be included in this dataset are mean, maximum, minimum, percentiles, persistence, and extreme values at various return periods, collectively forming a critical foundation for the precise and resilient design of offshore structures etc.

In this paper, a part of the wave analysis results at the development site of the Thien Nga - Hai Au oil field are presented (Figure 1, red point). The study area is located off the Southeast coast of Viet Nam, about 200 km from Bach Ho station, about 180 km from Con Dao, 400

Corresponding author: Le Quoc Huy
E-mail: huylq2@gmail.com

km from Phu Quy Island, and 500 km from the Large Truong Sa Island. The extreme wave statistical methods are applied to calculate statistical wave characteristics and extreme wave values at different return periods. Extreme wave analysis is the extrapolation of a known wave time-series to a longer duration [2]. These extreme values have a decisive influence on the design of offshore structures. The study uses wave data simulated by the SWAN model in 30 years from 1991 to 2020, which has higher spatial and temporal resolution than widely published data sources such as ERA5 and CMEMS.

2. Data and method

2.1. Data

The study uses wave simulation data based on the SWAN model (Simulation Wave Nearshore) for 30 years from 1991 to 2020. The simulation data is downscaled into 1 hour temporal and 7 km spatial resolution across the East Viet Nam Sea. The input wind data for the SWAN model is extracted from the results of the WRF model (Weather Research and Forecasting Model) at 1 hour temporal and 9 km spatial resolution. The WRF model uses ERA5 data at a resolution of 0.25 x 0.25 degrees (lat/long), 1 hour time step [3], [4] as input. Simulated data by WRF, SWAN models are evaluated and compared with observation data and ERA5, CMEMS reanalysis data at some meteorological and hydrographic stations in Viet Nam in the Southof the East Viet Nam Sea. The locations of the metocean stations used to evaluate the reanalysis data set are shown in Figure 1. The metocean data used to evaluate the reanalyis dataset including wind, temperature, humidity, sunshine duration and waves are presented in Table 1. In this paper, only the evaluation results of wave data are presented.

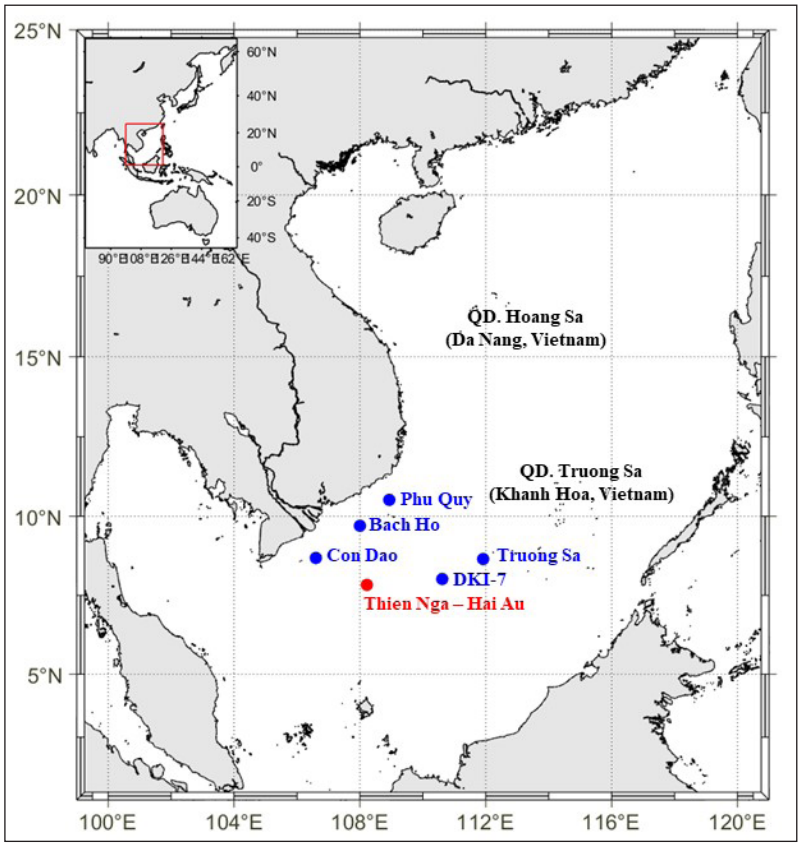

Figure 1. Metocean data domain, the location of the observation stations used to evaluate the reanalysis data (blue points) and the study site (red points)

Table 1. Factors used to evaluate the quality of the reanalysis data

Station	Wind	Temperature	Precipitation	Humidity	Sunshine duration	Wave
Con Dao	x	x	x	x	x	
Phu Quy	x	x	x	x	x	
DKI-7	x					
Bach Ho	x					x
Truong Sa	x					
ERA5	x	x				x
CMEMS						x

2.2. Method

2.2.1. Data evaluation

The wave simulation data is compared with the observations to evaluate the accuracy using different statistical criteria, thereby showing the similarity between the simulated wave data and the observed data [5]. The quantitative statistics were divided into three major categories: Standard regression, dimensionless, and error index. Standard regression statistics determine the strength of the linear relationship between simulated and measured data. Dimensionless techniques provide a relative model evaluation assessment, and error indices quantify the deviation in the units of the data of interest [7]. In this study, the accuracy of simulated wave data is evaluated based on Pearson correlation coefficient (r), coefficient of determination (R^2), Nash-Sutcliffe model efficiency coefficient (NSE) [8], Kling-Gupta efficiency index (KGE) [6], standard deviation rate RSR [9], [10] and percentage deviation (PBIAS) [11].

The coefficients r and R^2 are two of the standard regression statistics describing the degree of collinearity between simulated and observed data. The correlation coefficient r ranges from -1 to 1, if r = 1 or -1 then a perfect positive or negative linear relationship exists. Similarly, the coefficient R^2 describes the proportion of variance in the observed data explained by the model. R^2 ranges from 0 to 1, with higher values indicating less error variance, and typically values greater than 0.5 are considered acceptable [12], [13].

The NSE index is a dimensionless normalized statistic that determines the relative magnitude of the residual variance compared to the variance of the observed data. NSE indicates how well the plot of observed versus simulated data fits the 1:1 curve. NSE ranges from -∞ to 1, with NSE = 1 being the optimal value. Values between 0.0 and 1.0 are generally considered an acceptable efficiency level, whereas negative NSE values indicate unacceptable efficiency. Another dimensionless statistical indicator used in this study is the KGE, which is based on the decomposition of the NSE into its constituent components (correlation, standard deviation, mean deviation). KGE = 1 represents a perfect agreement between simulation and observation. Analogous to NSE = 0, KGE <0 indicates that the mean of the observations provides better estimates than the simulation [14], [15]. Positive KGE values are indicative of "good" model simulations, while negative KGE values are considered "bad".

Two of the error indices used in this study are RSR and PBIAS. The RSR standardizes the mean square error (RMSE) against the standard deviation of the observed data ($STDEV_{obs}$). The RSR varies from an optimal value of 0, which indicates zero RMSE or residual variation and therefore perfect model simulation, to a large positive value. The lower the RSR, the lower the RMSE, the better the simulation performance of the model. PBIAS index measures the average tendency of the simulated data to be larger or smaller than their observed counterparts [16]. The optimal value of PBIAS is 0, with low-magnitude values indicating accurate model

simulation. Positive values indicate model underestimation bias, and negative values indicate model overestimation bias [16].

The indicators r, R^2, NSE, KGE, RSR and PBIAS are calculated using equations (1), (2), (3), (4), (5) and (6), respectively:

$$r = \frac{n\sum_{i=1}^{n} S_i O_i - \sum_{i=1}^{n} S_i \sum_{i=1}^{n} O_i}{\sqrt{n}} \quad (1)$$

$$R^2 = \left[\frac{n\sum_{i=1}^{n} S_i O_i - \sum_{i=1}^{n} S_i \sum_{i=1}^{n} O_i}{\sqrt{n(\sum_{i=1}^{n} S_i^2) - (\sum_{i=1}^{n} S_i)^2}\sqrt{n(\sum_{i=1}^{n} O_i^2) - (\sum_{i=1}^{n} O_i)^2}}\right]^2 \quad (2)$$

$$NSE = 1 - \frac{\sum_{i=1}^{n}(O_i - S_i)^2}{\sum_{i=1}^{n}(O_i - \bar{O})^2} \quad (3)$$

$$KGE = 1 - \sqrt{(r-1)^2 + \left[\frac{\sqrt{\sum_{i=1}^{n}(S_i - \bar{S})^2}}{\sqrt{\sum_{i=1}^{n}(O_i - \bar{O})^2}} - 1\right]^2 + \left[\frac{\bar{S}}{\bar{O}} - 1\right]^2} \quad (4)$$

$$RSR = \frac{RMSE}{STDEV_{obs}} = \frac{\sqrt{\sum_{i=1}^{n}(S_i - O_i)^2}}{\sqrt{\sum_{i=1}^{n}(O_i - \bar{O})^2}} \quad (5)$$

$$PBIAS = \frac{\sum_{i=1}^{n}(S_i - O_i)}{\sum_{i=1}^{n} O_i} \times 100 \quad (6)$$

Where S_i and O_i are simulated and observed values, respectively, $\bar{O}$ is the average observed value, $\bar{S}$ is the average simulated value, and *n* is the number of values. The performance of the model is considered satisfactory if *r > 0.5*, *R^2 > 0.3, NSE > 0.5, RSR < 0.6, -25% < PBIAS < 25%* [17] and *KGE > 0.5* [15]. The recommended statistical criteria for evaluating the data are presented in Table 2.

Table 2. Statistical standards used to evaluate the wave reanalysis data

Evaluate	r	R^2	NSE	KGE	RSR	PBIAS
Very good	> 0.9	> 0.8	> 0.7	> 0.7	< 0.5	[-10 10]
Good	0.7 ÷ 0.9	0.5 ÷ 0.8	0.6 ÷ 0.7	0.6 ÷ 0.7	0.5 ÷ 0.6	[-15 -10) and (10 15]
Satisfactory	0.5 ÷ 0.7	0.3 ÷ 0.5	0.5 ÷ 0.6	0.5 ÷ 0.6	0.6 ÷ 0.7	[-25 -15) and (15 25]
Unsatisfactory	< 0.5	< 0.3	< 0.5	< 0.5	> 0.7	> 25 or < -25

2.2.2. Weibull distribution

Wave analysis for marine engineering design requires a clear understanding of waves based on sufficiently long and reliable data. The standard procedure in statistical analysis of extreme waves is [1, 18, 19]: (i) Selection of data for analysis; (ii) Selection of the best-fit distribution function for the wave data; (iii) Calculation of extreme wave values with different return periods based on the selected

distribution function; (iv) Calculation of the confidence interval. Various studies have been done on the method of estimating the return periods of the wave heights using one of the following probability distribution functions: Weibull, Gumbel, Exponential, General Pareto (GPD), and Log-normal distribution. In most cases, Weibull, Gumbel, Exponential, GPD and Log-normal distribution functions can give a good-fit of the wave data. In this study, we chose the Weibull distribution because it is more suitable for wave data in the study area as recommended in previous studies [1], [19], [20].

The probability density function of a three-parameter Weibull distribution is written as follows:

$$f(h) = \frac{\alpha}{\beta}\left(\frac{h-k}{\beta}\right)^{\propto-1} exp\left[-\left(\frac{h-k}{\beta}\right)^{\alpha}\right] \quad (7)$$

Where *h* is the wave height, *k* is the position parameter, *β* is the shape parameter and *α* is the scale parameter.

The cumulative frequency distribution function of the three-parameter Weibull distribution is written as follows:

$$F(h) = P(H \le h) = 1 - exp\left[-\left(\frac{h-k}{\beta}\right)^{\alpha}\right] (8)$$

The occurrence probability of *a* wave crest value *H* greater than or equal to *a* particular wave value *h* is:

$$P(H \ge h) = 1 - F(h) = exp\left[-\left(\frac{h-k}{\beta}\right)^{\alpha}\right] \quad (9)$$

For the Weibull distribution, extrapolation of the wave height h_{TR} to *a* defined return interval *TR* can be calculated based on the following equation:

$$h_{T_R} = k + \beta\left[ln(mT_R)^{1/\alpha}\right] \quad (10)$$

Where *m* is the number of wave data for 1 year used in the extreme wave analysis.

2.2.3. Wave persistence

Wave persistence represents the length of time for which a significant wave height limited by certain thresholds is likely to persist. In this paper, 6 hours, 12 hours, 24 hours, 36 hours, and 48 hours persistence for 1 m, 2 m, and 3 m significant wave height thresholds are examined.

3. Results and Discussion

3.1. Evaluation and selection of data sources

To compare with observed data (OBS), three sources of wave data are used in this study including SWAN data at spatial resolution of 7 km, temporal resolution of 1 hour, which is much more detailed than the ERA5 data at spatial resolution of 0.5 degrees (lat/long), temporal resolution of 1 hour, and the CMEMS wave data at spatial resolution of 0.2 degrees (lat/long), temporal resolution of 3 hours. The comparison results of significant wave heights in Figure 2 show high agreements of ERA5, SWAN and CMEMS with observed data at Bach Ho station. Overall, the SWAN wave simulation are better capturing the wave crests than the ERA5 and CMEMS, which are also shown in Figure 3.

The analysis results of the quality evaluation indicators of the wave data are highly positive, six evaluation indicators indicate good agreement for all three data sources SWAN, ERA5 and CMEMS as presented in Table 2, the difference in values of the statistical indicators of the three data sources are noncomparable (Figure 3). The values of statistical indicators of the three data sources ERA5, SWAN and CMEMS are as follows: r = 0.95 for all three sources; R^2 = 0.91, 0.90 and 0.90; NSEs are 0.90, 0.88 and 0.89; KGEs are 0.93, 0.9 and 0.94; the RSRs are 0.32, 0.35 and 0.33; the PBIAS = 5.58%, 3.64% and 2.54%, respectively.

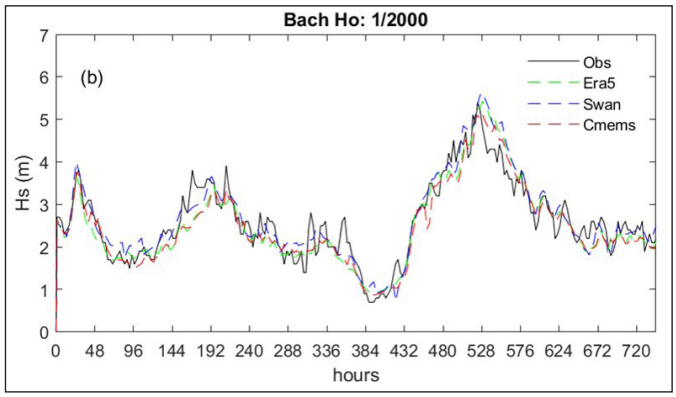

Figure 2. Comparison of significant wave height based on SWAN, ERA5 and CMEMS data with observed data at Bach Ho station in January 2000

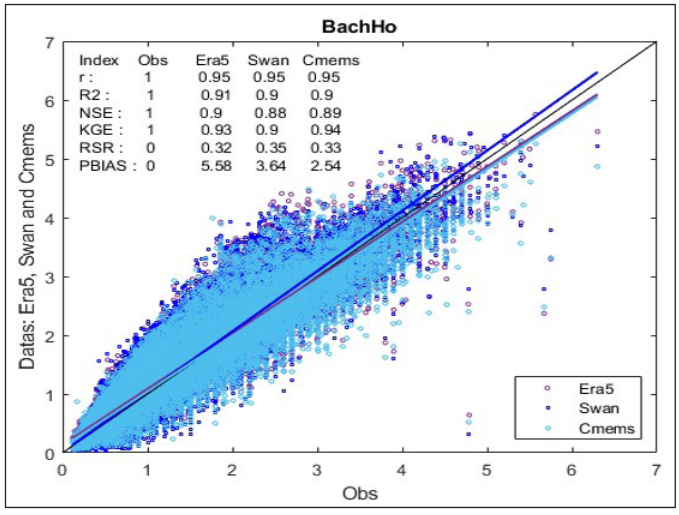

Figure 3. Correlation and quality evaluation indicators of SWAN, ERA5 and CMEMS data compared to the observed data at Bach Ho station

The SWAN wave data is not only equivalent in quality, but also has higher temporal and spatial resolution than the ERA5 and CMEMS data. Especially, compared to the other data sources, the SWAN wave data captures the observed wave crests better. Therefore, SWAN wave data was chosen to be used in this study.

3.1. Wave statistical characteristics

Tables 3 - 6 and Figure 4 - 6 present the characteristics and wave patterns in the study area during the period 1991 - 2020. Two main wave directions include: Northeast and Southwest. Northeast wave, accounting for 43.87% (up to 96.6% in January), is the main direction in the Northeast monsoon season from November to next March, the average wave heights of the months varies from 0.76 to 2.24 m, the maximum reaches 7.72 m; Southwest wave, accounting for 24.24% (up to 67.2% in July), is the main direction in the Southwest monsoon season from May to September. The average monthly wave heights varies from 0.48 to 0.98 m, the maximum is 3.28 m. Analysis of significant wave height percentiles shows that waves in May are the lowest with an estimate of 99 percent of waves having a height of less than 1.54 m and 95 percent of waves having a height of less than 1.11 m. In contrast, December is the month with the highest waves with an estimate of over 90 percent of significant waves greater than 1 m and over 50 percent of significant waves greater than 2 m (Table 3).

Table 3. Statistical characteristics of significant wave height in 30 years (1991 - 2020)

Month	Significant wave height [m]				Significant wave height at percentiles [m, %]							Main direction
	Min	Mean	Max	STD	99	95	90	50	10	5	1	
1	0.09	2.12	5.80	0.96	0.21	0.71	0.97	2.00	3.45	3.80	4.55	NE
2	0.08	1.68	5.48	1.01	0.14	0.33	0.51	1.54	3.12	3.54	4.38	NE
3	0.04	1.05	5.36	0.76	0.10	0.19	0.29	0.83	2.15	2.56	3.34	NE
4	0.04	0.60	3.39	0.46	0.09	0.13	0.16	0.45	1.24	1.51	2.13	E
5	0.02	0.48	2.76	0.32	0.09	0.14	0.18	0.38	0.91	1.11	1.54	SW
6	0.04	0.68	2.60	0.39	0.13	0.18	0.23	0.62	1.22	1.40	1.89	SW
7	0.06	0.88	2.84	0.49	0.12	0.20	0.27	0.83	1.56	1.71	2.08	SW
8	0.05	0.98	2.95	0.51	0.14	0.23	0.33	0.91	1.66	1.84	2.27	SW
9	0.04	0.83	3.28	0.48	0.12	0.20	0.26	0.76	1.46	1.68	2.12	SW
10	0.04	0.76	4.35	0.53	0.10	0.17	0.23	0.62	1.43	1.79	2.61	NE
11	0.07	1.42	7.72	0.82	0.19	0.36	0.51	1.25	2.53	3.00	3.95	NE
12	0.14	2.24	6.99	1.06	0.34	0.80	1.04	2.08	3.73	4.28	5.21	NE
1991 - 2020	0.02	1.14	7.72	0.89	0.11	0.19	0.26	0.90	2.39	3.00	4.09	NE, SW

Table 4. Frequency distribution of significant wave corresponding to wave direction in 30 years (1991 - 2020)

Hs (m)	Direction								Total
	N	NE	E	SE	S	SW	W	NW	
(0 ÷ 0.5)	1.76	5.07	6.95	1.45	1.37	5.64	3.83	1.06	27.13
[0.5 ÷ 1)	1.16	8.35	2.51	0.02	0.17	9.78	5.11	0.27	27.39
[1 ÷ 1.5)	0.64	9.08	0.38	0.01	0.02	6.31	2.77	0.10	19.31
[1.5 ÷ 2)	0.40	7.43	0.03	<0.01	<0.01	2.19	1.01	0.03	11.09
[2 ÷ 2.5)	0.19	5.43	<0.01		<0.01	0.27	0.27	0.01	6.19
[2.5 ÷ 3)	0.15	3.65	<0.01			0.04	0.03	0.01	3.89
[3 ÷ 3.5)	0.09	2.35	<0.01				0.01		2.45
[3.5 ÷ 4)	0.03	1.35					<0.01		1.39
[4 ÷ 4.5)	0.01	0.67							0.68
[4.5 ÷ 5)	<0.01	0.28					<0.01	<0.01	0.29
[5 ÷ 5.5)	<0.01	0.14							0.14
[5.5 ÷ 6)	<0.01	0.04					<0.01	<0.01	0.04
≥ 6		0.01					<0.01	<0.01	0.02
Total	**4.43**	**43.87**	**9.88**	**1.48**	**1.57**	**24.24**	**13.04**	**1.49**	**100**

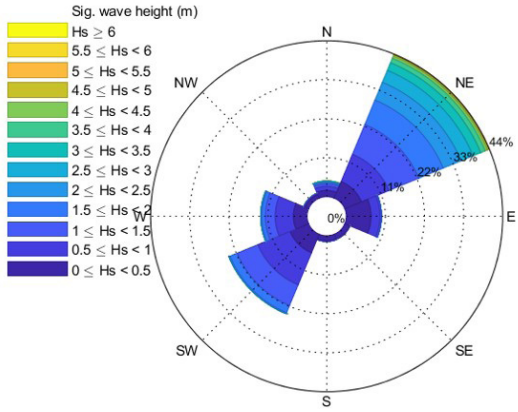

Figure 4. Pattern and direction of significant wave in the 30 years (1991 - 2020)

Table 5. Frequency distribution of significant waves corresponding to wave direction in January of 30 years (1991 - 2020)

Hs (m)	Direction								Total
	N	NE	E	N	NE	SW	N	NE	
(0 ÷ 0.5)	0.26	1.51	0.68	0.17	0.11	0.17	0.04	0.02	2.97
[0.5 ÷ 1)	0.13	6.78	0.70	0.01	0.01	0.06			7.70
[1 ÷ 1.5)	0.14	17.93	0.25	0.01	0.01			<0.01	18.34
[1.5 ÷ 2)	0.11	20.88	0.06						21.05
[2 ÷ 2.5)	0.18	17.12	0.04						17.34
[2.5 ÷ 3)	0.09	13.38	0.03						13.49
[3 ÷ 3.5)	0.08	9.74	0.02						9.83
[3.5 ÷ 4)	0.02	6.19							6.21
[4 ÷ 4.5)		1.99							1.99
[4.5 ÷ 5)		0.55							0.55
[5 ÷ 5.5)		0.41							0.41
[5.5 ÷ 6)		0.11							0.11
≥ 6									
Total	1.00	**96.59**	1.78	0.19	0.13	0.23	0.04	0.02	100

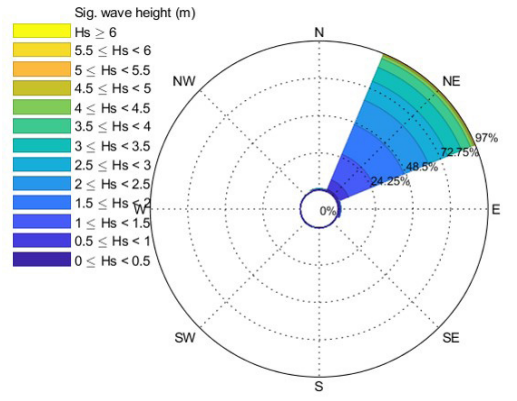

Figure 5. Pattern and direction of significant wave in 30 years (1991 - 2020)

Table 6. Frequency distribution of significant waves corresponding to wave direction in July of 30 years (1991 - 2020)

Hs (m)	Direction								Total
	N	NE	E	N	NE	SW	N	NE	
(0 ÷ 0.5)	1.05	2.67	3.07	1.33	1.40	11.39	5.07	1.31	27.29
[0.5 ÷ 1)	0.01	0.04	0.23		0.54	24.03	8.55	0.09	33.49
[1 ÷ 1.5)						21.87	4.99		26.86
[1.5 ÷ 2)						8.82	2.02	<0.01	10.84
[2 ÷ 2.5)						0.84	0.39	0.05	1.27
[2.5 ÷ 3)						0.22		0.02	0.24
[3 ÷ 3.5)									
[3.5 ÷ 4)									
[4 ÷ 4.5)									
[4.5 ÷ 5)									
[5 ÷ 5.5)									
[5.5 ÷ 6)									
≥ 6									
Total	1.06	2.71	3.30	1.33	1.94	**67.17**	21.01	1.48	100

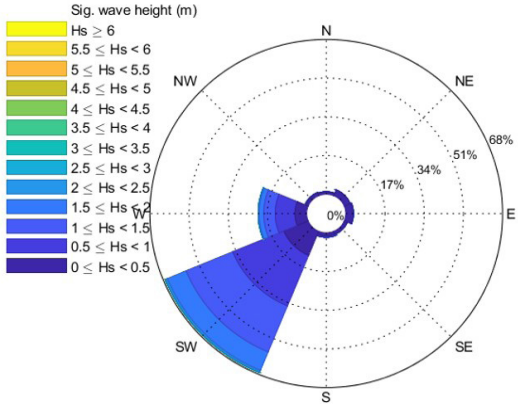

Figure 6. Pattern and direction of significant wave in 30 years (1991 - 2020)

Statistics of wave persistence including the threshold of significant wave height and the length of persistence time for 12-month wave data (Table 7) show that May is the month when the sea is calmest and most favorable for marine activities around the study area, followed by April, June and October, whereas December is the roughest month of the year. Specifically, in May, significant wave heights under 1 m persisting for 6 hours account for 92.0%, persisting for 12 hours account for 91.4%, for 24 hours account for 90.1%, for 36 hours account for 88.7 % and for 48 hours account for 88.0%; significant wave heights under 2m persisting for 6 hours account for 99.8%, for 12 hours account for 99.8%, for 24 hours account for 99.7%, for 36 hours account for 99.5% and for 48 hours account for 99.6%; especially, significant wave heights under 3m persisting for 6 hours, 12 hours, 24 hours, 36 hours and 48 hours account for 100% in all four months of May, June, July and August. In December, significant wave heights under 1 m persisting for 6 hours account for 8.4%, for 12

hours account for 7.8%, for 24 hours account for 6.5%, for 36 hours account for 5.0% and for 48 hours account for 4.7%; significant wave heights under 2 m persisting for 6 hours account for 45.5%, for 12 hours account for 43.5%, for 24 hours account for 40.2%, for 36 hours account for 37.0% and for 48 hours account for 34.0%; significant wave heights under 3 m persisting for 6 hours account for 77,5%, for 12 hours account for 76.3%, for 24 hours account for 74.0%, for 36 hours account for 71.7% and for 48 hours account for 68.9%.

Table 7. Statistics of wave persistence below threshold by the time (%)

Month	Significant wave height under 1.0 m					Significantwave height under 2.0 m					Significant wave height under 3.0 m				
	6 h	12 h	24 h	36 h	48 h	6 h	12 h	24 h	36 h	48 h	6 h	12 h	24 h	36 h	48 h
1	10.1	9.6	8.3	6.5	6.0	48.5	46.7	44.0	41.0	39.4	79.9	78.6	76.3	74.5	72.5
2	30.0	29.3	27.5	25.1	23.1	65.4	64.0	61.7	59.4	57.6	87.5	86.8	85.0	83.7	82.8
3	56.8	55.3	53.1	50.0	49.0	86.6	85.7	84.1	82.7	82.2	97.7	97.5	96.9	96.5	95.7
4	82.0	80.9	79.1	77.0	75.6	98.5	98.3	97.9	98.0	97.3	99.9	99.9	99.8	99.8	99.6
5	92.0	91.4	90.1	88.7	88.0	99.8	99.8	99.7	99.5	99.6	100	100	100	100	100
6	77.8	75.9	72.9	69.5	67.8	99.2	99.0	98.7	98.5	98.2	100	100	100	100	100
7	58.8	56.6	52.9	50.7	48.0	98.2	97.8	97.0	96.3	95.9	100	100	100	100	100
8	53.4	51.7	48.1	46.9	42.6	96.7	96.1	95.1	94.5	93.6	100	100	100	100	100
9	64.0	62.3	59.2	56.3	53.8	98.0	97.7	96.9	96.8	95.6	99.8	99.8	99.8	99.7	99.6
10	71.9	70.3	67.3	64.8	63.0	96.3	95.9	95.2	94.5	94.0	99.5	99.4	99.3	99.0	98.9
11	33.4	31.7	28.9	25.3	24.7	77.7	76.3	74.0	72.0	69.8	94.5	93.9	93.0	92.2	91.1
12	8.4	7.8	6.5	5.0	4.7	45.5	43.5	40.2	37.0	34.0	77.5	76.3	74.0	71.7	68.9

3.2. Extremes of significant wave height

Table 8 presents the estimation results of significant wave height for all eight major directions (N, NE, E, SE, S, SW, W, and NW) for 1, 5, 10, 25, 50 and 100 year return periods. The significant wave heights show a very large difference between the directions, which can be divided into three groups: Group 1 includes the NE and N directions; group 2 includes the SW, W and NW directions; group 3 is the smallest group that includes the SE, E and S directions. The maximum value in the NE direction is 6 times larger than the minimum value in the SE direction, and two times larger than the remaining directions. Significant wave height values at 50 and 100 year return periods are 7.93 m and 8.4 m in the N direction; 8.17 m and 8.53 m in the NE direction; 2.13 m and 2.22 m in the E direction; 1.17 m and 1.22 m in the SE direction; 1.84 m and 1.93 m in the S direction; 3.44 m and 3.56 m in the SW direction; 3.56 m and 3.7 m in the W direction; and 2.91 m and 3.06 m in the NW direction, respectively.

Table 8. Extreme significant wave height corresponding to wave direction

Wave direction	Return period (year)					
	1	5	10	25	50	100
N	5.14	6.31	6.81	7.45	7.93	8.40
NE	5.93	6.91	7.30	7.80	8.17	8.53
E	1.58	1.82	1.92	2.04	2.13	2.22
SE	0.86	1.00	1.05	1.12	1.17	1.22

Wave direction	Return period (year)					
	1	5	10	25	50	100
S	1.31	1.54	1.63	1.75	1.84	1.93
SW	2.65	3.00	3.13	3.31	3.44	3.56
W	2.68	3.06	3.22	3.42	3.56	3.70
NW	2.00	2.38	2.54	2.75	2.91	3.06

When considering separately under the conditions of Northeast monsoon, Southwest monsoon and storm, the significant wave heights at 50 and 100 year return periods are 3.78 m and 3.84 m during the Southwest monsoon; 10.09 m and 10.45 m during the Northeast monsoon; 12.8 m and 13.6 m during storms, respectively (Table 9).

Table 9. Extreme significant wave height during monsoons and storms for different return periods

Return periods (year)	During monsoon		During storms
	Southwest monsoon (Southwest wave)	Northeast monsoon (Northeast wave)	
50	3.78	10.09	12.8
100	3.84	10.45	13.6

4. Conclusion

This paper focuses on evaluating the quality of wave data simulated by the SWAN model in the 30 year period 1991 - 2020. This data is used to calculate statistical characteristics and for the development and exploitation of the Thien Nga - Hai Au oil field off the Southeast coast of Viet Nam. Conclusions are given based on the results of wave analysis as follows:

With reference to the data recorded at the Bach Ho station, it is evident that the wave data simulated by the SWAN model proves more favorable for conducting extreme wave analyses compared to the reanalyzed data from ERA5 and CMEMS. This preference arises from several factors, including higher spatial and temporal resolutions, as well as a more accurate representation of wave crests closely aligning with the observed data.

Waves in the Thien Nga - Hai Au oil field site have two main directions: The Northeast wave accounts for 43.87%, the average monthly wave height is 0.76 - 2.24 m, the maximum is 7.72 m; the Southwest wave accounts for 24.24%, average monthly wave height is 0.48 - 0.98 m, maximum is 3.28 m.

The maximum wave heights for the return period of 50 and 100 years in the NE direction are 8.17 m and 8.53 m; in the SW direction are 3.44 m and 3.56 m; during the Southwest monsoon period are 3.78 m and 3.84 m; during the Northeast monsoon period are 10.09 m and 10.45 m; during storms are 12.8 m and 13.6 m.

Acknowledgment: *The completion of this article was made possible through the utilization of results obtained from the task titled "Metocean and Environment Data for the Development of Thien Nga - Hai Au Field, Block 12/11, Offshore Viet Nam".*

Disclaimer: *The authors declare that this article is the work of the authors, has not been published, and has not been copied from previous studies; there is no conflict of interest in the author group.*

References

1. Yoshimi Goda (2000), *Random Seas and Design of Maritime Structures (2nd Edition). Advanced Series on Ocean Engineering* - Volume 15.
2. Hassan Salah et al (2020), *"The use of exJtreme wave analysis for the design of offshore Structures",*

International Research Journal of Engineering and Technology, Volume: 07 Issue: 09, Sep. 2020. www.irjet.net, e-ISSN: 2395-0056, p-ISSN: 2395-0072.

3. Hersbach, H. et al (2020), *"The ERA5 global reanalysis", Q. J. R. Meteorol. Soc*., 146, 1999-2049. [Google Scholar] [CrossRef]
4. Bell, B. et al (2021), *"The ERA5 global reanalysis: Preliminary extension to 1950". Q. J. R. Meteorol. Soc.* 147, 4186-4227. [Google Scholar] [CrossRef]
5. Eric P. Chassignet et al (2007), *"The HYCOM (HYbrid Coordinate Ocean Mod-el) data assimilative system", Journal of Marine Systems*, Volume 65, Issues 1-4, March 2007, 60-83.
6. Gupta, H. V., et al (2009), *"Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling," Journal of Hydrology*, Volume 377, Issues 1-2, 20 October, 80-91. Accessed: 00, 2020. [Online]. Available: https://hdl.handle.net/11511/42626.
7. Legates, D. R., and McCabe, G. J. (1999), *"Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation", Water Resources Res,* January 1999. 35(1): 233-241.
8. Nash, J.E. and J.V. Sutcliffe (1970), *"River flow forecasting through conceptual models: Part 1. A discus-sion of principle", Hydrology,* 10(3): 282-290. CrossRef.
9. Singh, J. et al (2005), *"Hydrological modeling of the Iroquois River Watershed using HSPF and SWAT*", *J. Am. Water Resour Assoc* 41, 343-360.
10. Golmar Golmohammadi et al (2014), *"Evaluating Three Hydrological Distributed Watershed Models: MIKE-SHE, APEX, SWAT", Hydrology*, 1(1), 20-39; https://doi.org/10.3390/hydrology 1010020.
11. Sorooshian, S., Duan, Q. & Gupta, V. K (1993), *"Calibration of rainfall-runoff models: Application of global optimization to the Sacramento Soil Moisture Accounting Model", Water Resources Research.* Volume 29, Issue4, April 1993. Pages 1185-1194.
12. Santhi, C et al (2001), *"Validation of the SWAT model on alarge river basin with point and nonpoint sources", J. Ameri-canWater Resources Assoc*. 37(5): 1169-1188.
13. Van Liew, M. W. et al (2003), *"Hydrologic simulation on agricultural watersheds: Choosing between two models", Trans. ASAE 46(6)*: 1539-1551.
14. Koskinen, M., et al (2017), *"Resto-ration of nutrient-rich forestry-drained peatlands poses a risk for high exports of dissolved organic carbon, nitrogen, and phosphorus", Sci. Total Environ*., 586 (February), 858-869, doi:10.1016/j.scitotenv.2017.02.065.
15. Knoben, W. J. M. et al (2019), *"Technical note: Inherent benchmark or not? Comparing Nash-Sutcliffe and Kling-Gupta efficiency scores", Hydrology and Earth System Sciences*, 2019, 23(10), 4323-4331. https://doi.org/10.5194/hess-2019-327.
16. Gupta, H. V. et al (1999), *"Status of automatic calibration for hydrologic mod-els: Comparison with multilevel expert calibration", J. Hydrologic Eng*, 1999. 4(2): 135-143.
17. Moriasi, D.N (2007), *"Model Evaluation Guidelines For Systematic Quantification Of Accuracy in Watershed Simulations", Trans. ASABE*, 50, 885-900.
18. Martin Mathiesen et al (1994), *"Recommended practice for extreme wave analysis", Journal of Hydraulic Research,* Vol. 32, 1994, No. 6.
19. World Meteorological Organization, *Guide to Wave Analysis and Forecasting*, WMO no. 702, Geneva, Switzerland, 2nd Edition, 1998.
20. Smith, R.L. (1984), *"Threshold Methods for Sample Extremes. In Statistical Extremes and Applications", Springer: Dordrecht, the Netherlands,* pp. 621-638.