

Article

Global Wind and Wave Climate Based on Two Reanalysis Databases: ECMWF ERA5 and NCEP CFSR

Christos Stefanakos [†] 

SINTEF Ocean, Marine Modelling and Analysis Group, N-7465 Trondheim, Norway;
christos.stefanakos@sintef.no

[†] Current address: SINTEF Ocean, Postboks 4762 Torgard, N-7465 Trondheim, Norway.

Abstract: In the present work, the global wind and wave climate is studied on the basis of two well-known reanalysis products, namely ERA5 and CFSR-W (WW3 hereafter). Several statistical features of the datasets are assessed, such as seasonal variability, quantiles of the probability distribution, monthly, annual and inter-annual variability, and several error metrics. The time span covers a period of 31 years (1979–2009), a fact that assures that most of the long-scale features are equally present in both datasets. The analysis performed is depicted both on a global and regional scale. The results are also assessed by means of a global satellite altimeter dataset.

Keywords: significant wave height; wind speed; wind climatologies; wave climatologies; ERA5; CFSR-W; seasonal analysis; probability analysis



Citation: Stefanakos, C. Global Wind and Wave Climate Based on Two Reanalysis Databases: ECMWF ERA5 and NCEP CFSR. *J. Mar. Sci. Eng.* **2021**, *9*, 990. <https://doi.org/10.3390/jmse9090990>

Academic Editor: Eugen Rusu

Received: 20 August 2021

Accepted: 8 September 2021

Published: 11 September 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The study of ocean wave climate is of great importance for a large number of applications, including among others: ocean climate change studies, design of ships and offshore/coastal structures, planning of sea operations, wind and wave energy conversion.

There are various sources of wind and wave data, namely (i) visual observations [1], (ii) satellite altimeters [2,3], (iii) in situ buoy measurements, and (iv) numerical models [4,5]. Each one of them have its own advantages and disadvantages. The great advantage of numerical models is their wide coverage on a high time and space resolution in a global scale, making it possible to produce long-term wind and wave climatologies without gaps.

The quality of wave data is heavily dependent on the quality of the wind forcing, improvements on the physical modelling and assimilation techniques among others. At least two meteorological centers (National Centers for Environmental Prediction, NCEP, and European Centre for Medium-Range Weather Forecasts, ECMWF) work constantly for the last few decades towards the improvement of their forecast models. For this, apart from the results of their operational versions, they perform reanalysis studies from time to time to offer homogenized wave products.

(Stopa [6], Table 1) summarizes most of the well known wind datasets used throughout the years to generate global wave hindcasts. These wind products led to several global reanalyses of wave datasets over the last few decades. Among them, it is worth noting ERA-15, ERA-40, ERA-Interim from ECMWF [7,8], NCEP/NCAR, and CSFR-W from NCEP [8–10]. Other climatologies include HIPOCAS [11], GOW1 and 2 [12,13].

The emergence of these reanalysis databases has been followed by intercomparison studies against each other and/or against other sources of data such as buoy and satellite measurements. For example, Caires et al. [14] have compared six different reanalysis datasets (ERA-40 and NCEP/NCAR among them) against NOAA buoy and TOPEX/Poseidon altimeter datasets. Semedo et al. [15] compared ERA-40 against visual observations and satellite data. Stopa in a series of works has analyzed CSFR-W wave data and their relation to climate indices [16], against ERA-Interim and a number of altimeter data [17], using 12 different wind fields as input and against two different satellite

datasets [6]. Finally, [18] have compared the results of HIPOCAS climatology against CSFR-W and ERA-Interim.

In the present work, a global wind and wave climate analysis is performed by inter-comparing two reanalysis databases: (a) the CFSR-W [8,10] called for simplicity WW3, and (b) the newly released ERA5 by ECMWF through Copernicus Climate Data Store [19]. The former has been extensively studied by Stopa and coauthors [16,17], whereas the latter is studied here for the first time. Preliminary results, including only results for waves, have also been presented in [20].

In the next section, a brief description of the two datasets is given, as well as a description of the analysis procedure followed. Then, numerical results are presented and commented on, and finally, conclusions are drawn.

2. Methodology

2.1. Data Used

In the present study, three datasets were used for the intercomparison, covering the entire globe. The two of them were generated by means of third-generation numerical models, namely WAVEWATCH III [21] and WAM [4], and the third one consisted of satellite altimeter measurements merged from several satellite missions. The model data were in regular gridded netcdf format ($361 \text{ lats} \times 720 \text{ lons} = 259,920 \text{ datapoints}$), while the satellite data were stored in tracks.

The first dataset, called WW3 hereafter for simplicity, is a reanalysis generated by means of WAVEWATCH III model [10]. It consists of fields of significant wave height and wind speed covering the entire globe for the period 1979–2009 in 3-h intervals ($31 \text{ years} \times 2920 \text{ 3-h} = 90,520 \text{ time instances}$). For a more detailed description of the data and the model setup, one can see here: <https://polar.ncep.noaa.gov/waves/hindcasts/nopp-phase2.php> (last accessed: 10 September 2021).

The second dataset, ERA5, is a reanalysis generated by means of WAM model and has recently been released [19]. It also consists of fields of significant wave height and wind speed. Although the data cover the period 1979-present in hourly intervals, only the period 1979–2009 in 3-hourly intervals was taken into account to facilitate the comparison with WW3. For a more detailed description, see: <https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era5> (last accessed: 10 September 2021).

In addition, and for comparison purposes, satellite altimeter data from the archive of IFREMER were used. The archive contains data from nine altimeter missions, namely ERS-1,2, ENVISAT, TOPEX/Poseidon, Jason-1,2, GEOSAT-FO, Cryosat-2, SARAL, covering the period 1992–2016. Since some of the altimeters (ERS-1,2, ENVISAT, Jason-1, Cryosat-2, SARAL) have already been assimilated in ERA5 [19], they were excluded from the calculations. More detailed information about the missions, as well as about the validation against buoy data and the induced corrections can be found in [22].

Since satellite data have a different time-space data structure, and in order to make possible a comparison between them and the two model datasets, mean monthly values were calculated from all sources for 13 different subregions; see Figure 1. These discrete non-overlapping subregions have been defined by [23], such that the wave conditions within each of them to be qualitatively similar [24].

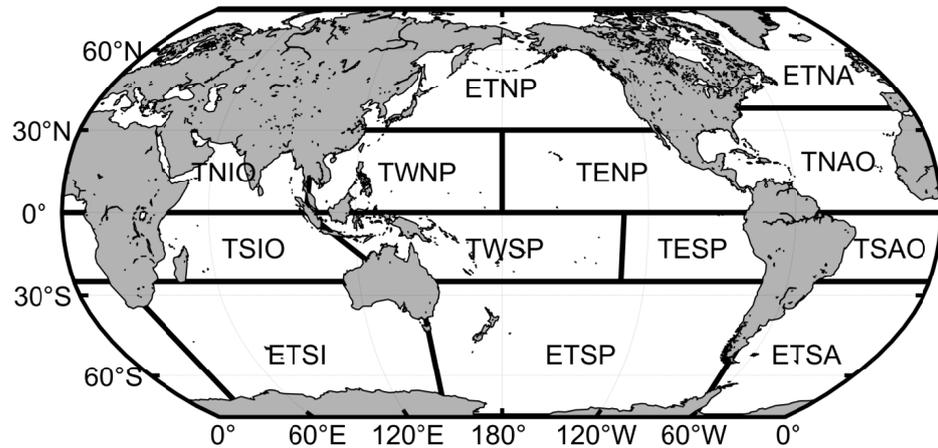


Figure 1. Selected subregions of the oceans: Extratropical South Indian (ETSI), Extratropical South Pacific (ETSP), Extratropical South Atlantic (ETSA), Tropical South Indian Ocean (TSIO), Tropical Western South Pacific (TWSP), Tropical Eastern South Pacific (TESP), Tropical South Atlantic Ocean (TSAO), Tropical North Indian Ocean (TNIO), Tropical Western North Pacific (TWNP), Tropical Eastern North Pacific (TENP), Tropical North Atlantic Ocean (TNAO), Extratropical North Pacific (ETNP), Extratropical North Atlantic (ETNA).

2.2. Statistical Analysis Procedure

The data were available as fields in the form

$$\{X(t_i, \phi_j, \lambda_k), \quad i = 1, \dots, I, \quad j = 1, \dots, J, \quad k = 1, \dots, K\}, \tag{1}$$

where i runs over the time instances, and j, k over the latitudes and longitudes, respectively.

In the sequel, two kind of analysis were performed: (a) field analysis, showing results for the entire field, and (b) averaged analysis, showing results averaged over all (or a subset of) datapoints. In this way, various statistics related to different aspects of the datasets were depicted.

First, the time index was reparametrized according to Buys–Ballot triple index [25] in order to properly treat variability at different time scales. Hence, the following triple index (y, m, n) was used. The first component y is the yearly index. The second $m = \{1, 2, \dots, 12\}$ is the monthly index. The third $n = \{1, 2, \dots, N_m\}$ represents the time within a month, with N_m being the number of 3-hourly observations within the m -th month.

According to the three-index notation, the field $X(t_i, \phi_j, \lambda_k)$ was reindexed as follows:

$$\{X(y, m, n, \cdot, \cdot), \quad y = 1, \dots, \quad m = 1, \dots, 12, \quad n = 1, \dots, N_m\}. \tag{2}$$

The three indices (y, m, n) represent three different time scales, making it possible to explicitly define statistics with respect to each one of them, separately.

2.2.1. Seasonal Analysis

First, the fields of monthly values of mean value and standard deviation were formed

$$\mu_3(y, m, \cdot, \cdot) = \frac{1}{N_m} \sum_{n=1}^{N_m} X(y, m, n, \cdot, \cdot), \tag{3}$$

$$\sigma_3(y, m, \cdot, \cdot) = \sqrt{\frac{1}{N_m} \sum_{n=1}^{N_m} [X(y, m, n, \cdot, \cdot) - \mu_3(y, m, \cdot, \cdot)]^2}. \tag{4}$$

Then, the mean monthly values were obtained by averaging Equations (3) and (4) over the years Y :

$$\tilde{\mu}_3(m, \cdot, \cdot) = \frac{1}{Y} \sum_{y=1}^Y \mu_3(y, m, \cdot, \cdot), \tag{5}$$

$$\tilde{\sigma}_3(m, \cdot, \cdot) = \frac{1}{Y} \sum_{y=1}^Y \sigma_3(y, m, \cdot, \cdot), \quad m = 1, 2, \dots, 12. \tag{6}$$

These parameters are also known as seasonal mean value and seasonal standard deviation, depicting the seasonal patterns of the data, and they have been used in a nonstationary time series modelling suitable for metocean and maritime parameters; see [25–27].

If, further, one averages over all ϕ_j 's and λ_k 's, an averaged picture for the mean monthly values in Equations (5) and (6) is obtained

$$\tilde{\mu}_3(m) = \frac{1}{JK} \sum_j \sum_k \tilde{\mu}_3(y, m, \cdot, \cdot), \tag{7}$$

$$\tilde{\sigma}_3(m) = \frac{1}{JK} \sum_j \sum_k \tilde{\sigma}_3(m, \cdot, \cdot). \tag{8}$$

Combining Equations (5) and (6), or, equivalently, Equations (7) and (8), one can calculate the coefficient of variation

$$\tilde{c}\tilde{v}_3(m, \cdot, \cdot) = \frac{\tilde{\sigma}_3(m, \cdot, \cdot)}{\tilde{\mu}_3(m, \cdot, \cdot)}, \tag{9}$$

depicting the Mean Monthly Variability (MMV) of the field (equiv. of the averaged data).

Similarly, one can calculate the yearly values of mean value and standard deviation

$$\mu_{32}(y, \cdot, \cdot) = \frac{1}{N_y} \sum_{m=1}^M \sum_{n=1}^{N_m} X(y, m, n, \cdot, \cdot), \tag{10}$$

$$\sigma_{32}(y, \cdot, \cdot) = \sqrt{\frac{1}{N_y} \sum_{m=1}^M \sum_{n=1}^{N_m} [X(y, m, n, \cdot, \cdot) - \mu_{32}(y, \cdot, \cdot)]^2}, \tag{11}$$

where

$$N_y = \sum_{m=1}^M N_m. \tag{12}$$

Then, the Mean Annual Variability (MAV) is obtained as

$$\tilde{c}\tilde{v}_{32}(\cdot, \cdot) = \frac{\tilde{\sigma}_{32}(\cdot, \cdot)}{\tilde{\mu}_{32}(\cdot, \cdot)}, \tag{13}$$

where $\tilde{\mu}_{32}(\cdot, \cdot)$ and $\tilde{\sigma}_{32}(\cdot, \cdot)$ are calculated in a similar way as in Equations (5) and (6).

Further, the Inter-Annual Variability (IAV) is defined as

$$\hat{c}\hat{v}_{32}(\cdot, \cdot) = \frac{\hat{\sigma}_{32}(\cdot, \cdot)}{\hat{\mu}_{32}(\cdot, \cdot)}, \tag{14}$$

where

$$\hat{\mu}_{32}(\cdot, \cdot) = \frac{1}{Y} \sum_{y=1}^Y \mu_{32}(y, \cdot, \cdot), \tag{15}$$

$$\hat{\sigma}_{32}(\cdot, \cdot) = \sqrt{\frac{1}{Y} \sum_{y=1}^Y [\mu_{32}(y, \cdot, \cdot) - \hat{\mu}_{32}(\cdot, \cdot)]^2}, \tag{16}$$

are the mean value and standard deviation of $\mu_{32}(y, \cdot, \cdot)$; see Equation (10). By investigating these parameters, one can look into the year-to-year variability; see, e.g., [16].

2.2.2. Probability Analysis

If the probability function of month m and year y is denoted by

$$\mathcal{F}(x; y, m, \cdot, \cdot), \quad y = 1, 2, \dots, Y, \quad m = 1, 2, \dots, M, \tag{17}$$

then the quantiles of it are given by

$$x_p = \mathcal{F}^{-1}(p; y, m, \cdot, \cdot), \quad p \in [0, 1]. \tag{18}$$

In the present paper the quantiles of $p = 50\%$ (median), 90%, 95%, 99%, and 99.9% were calculated.

The overall probability function, and consequently the associated quantiles, could be obtained by adding the frequencies of each month m and year y . Similarly, in the averaged analysis, the probability was obtained by adding the frequencies of all points together.

2.2.3. Error Analysis

The following statistics were used as error metrics in order to assess the difference between the two reanalysis datasets: Bias, Root-Mean-Square (RMSE), Scatter Index (SI), and Pearson’s correlation coefficient (CorrCoeff).

Following the parametrization of the previous sections, the monthly values of these metrics were (1 stands for ERA5, and 2 for WW3 dataset):

$$\text{Bias}(y, m, \cdot, \cdot) = \overline{(X_1 - X_2)}, \tag{19}$$

where $\overline{X} \equiv \mu_3(y, m, \cdot, \cdot)$ is calculated using Equation (3).

$$\text{RMSE}(y, m, \cdot, \cdot) = \sqrt{\overline{(X_1 - X_2)^2}}, \tag{20}$$

$$\text{SI}(y, m, \cdot, \cdot) = \text{RMSE} / \overline{X_1}, \tag{21}$$

$$\text{CorrCoeff}(y, m, \cdot, \cdot) = \frac{\overline{X_1 X_2} - \overline{X_1} \overline{X_2}}{\sqrt{(\overline{X_1^2} - \overline{X_1}^2)(\overline{X_2^2} - \overline{X_2}^2)}}. \tag{22}$$

In addition, according to [17], the following error metrics were also computed

$$\text{NBias}(y, m, \cdot, \cdot) = \overline{(X_1 - X_2)} / \sqrt{\overline{X_1^2}}, \tag{23}$$

$$\text{NSTD}(y, m, \cdot, \cdot) = \text{std}(X_2) / \text{std}(X_1), \tag{24}$$

$$\text{CRMSE}(y, m, \cdot, \cdot) = \text{std}(X_1 - X_2) / \text{std}(X_1), \tag{25}$$

where $\text{std}(X) \equiv \sigma_3(y, m, \cdot, \cdot)$ is calculated using Equation (4).

The mean monthly and/or annual variability of these indicators can be studied by applying appropriate averaging as in Equations (5) and (6) and/or Equations (10) and (11).

3. Numerical Results

3.1. Seasonal Analysis

Following the analysis procedure presented in the previous sections, the mean monthly variability was assessed on the basis of the mean monthly values $\tilde{\mu}_3(m, \cdot, \cdot)$ and $\tilde{\sigma}_3(m, \cdot, \cdot)$, calculated by Equations (5) and (6). According to these, there was a zonal distribution of the values with a distinct different behaviour between the Northern and the Southern Hemisphere. Especially, around the equator there was a zone with the least variability in

all months. It seemed that the variability of mean values of ERA5 was a bit lesser than the corresponding of WW3 in both hemispheres.

In addition, the seasonal variability on a regional basis was studied by analysing separately the 13 subregions shown in Figure 1. In Figure 2, the mean monthly values $\tilde{\mu}_3(m)$ and $\tilde{\sigma}_3(m)$, calculated using all points included in each subregion, are depicted for 8 out of the 13 subregions: two showing the variability in the northern extratropical zone (ETNP, ETNA), four in the tropical zone (TENP, TNAO, TNIO, TSIO) and two in the southern extratropical zone (ETSP, ETSA).

Additionally, and for comparison purposes, the mean monthly values of satellite data are also plotted. Although the satellite measurements did not have the structure of a time series, but rather of a time-space series, it was possible to calculate equivalent mean monthly values $\tilde{\mu}_3(m)$ and $\tilde{\sigma}_3(m)$ by considering all points included in each subregion [25]. In addition, this was expected to reduce the error due to undersampling of altimeters, as suggested by [28]. As expected, the general picture showed more pronounced variability in the extratropical areas (ETNA, North Atlantic; ETNP, North Pacific; ETSP, South Pacific; ETSA, South Atlantic), and a lesser one in the tropical zone (TWNP, TNAO, TESP, TSAO). Additionally, in TNIO and TSIO the regional phenomenon of monsoon was observed.

Overall, there was better agreement between ERA5 and WW3, rather than with the satellite data; especially, in the extratropical subregions (ETNP, ETNA, ETSP, ETSA). In the tropical zone, satellite data were also in agreement with ERA5 and WW3 (TWNP, TNAO, TESP, TSAO). Generally, small deviations between satellite and model data could be attributed to the fact that the former were estimated using 18 years (1992–2009), while the latter were estimated using 31 years (1979–2009).

Further, the mean annual variability was investigated in terms of $\tilde{c}\tilde{v}_{32}(\cdot, \cdot)$; see Figure 3. The agreement between ERA5 and WW3 datasets was good. Especially for the waves, more pronounced variability was shown by both models in the North Pacific and North Atlantic Ocean.

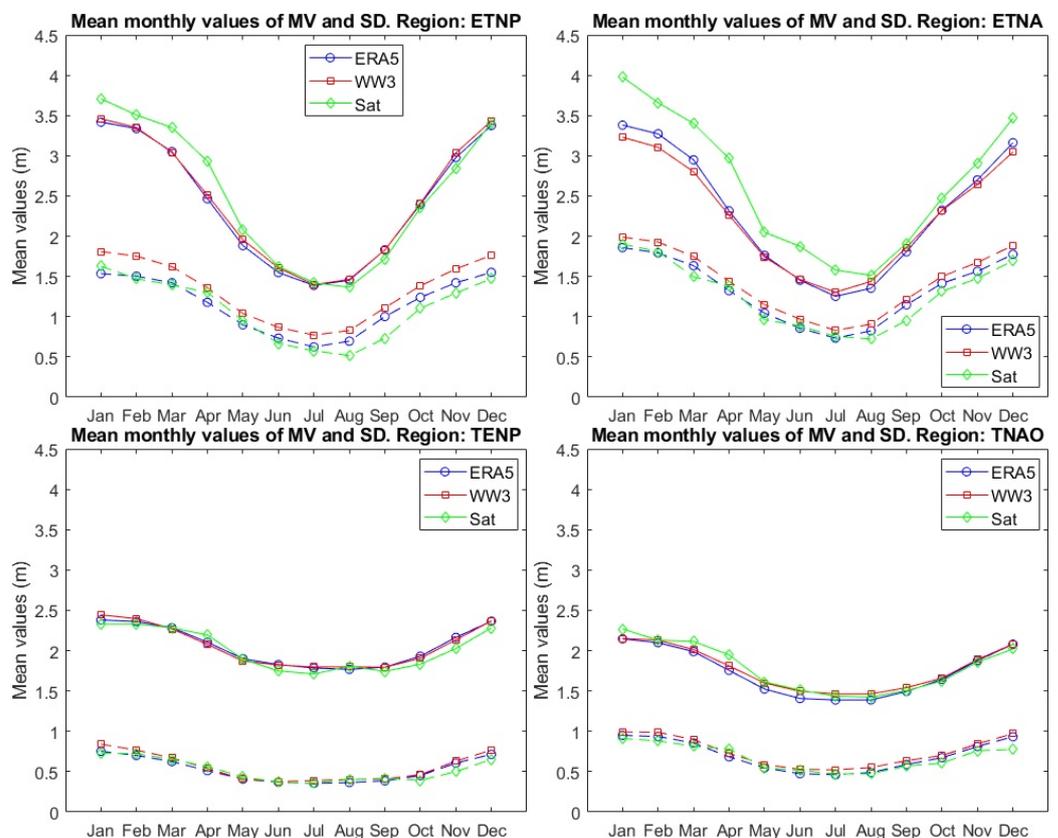


Figure 2. Cont.

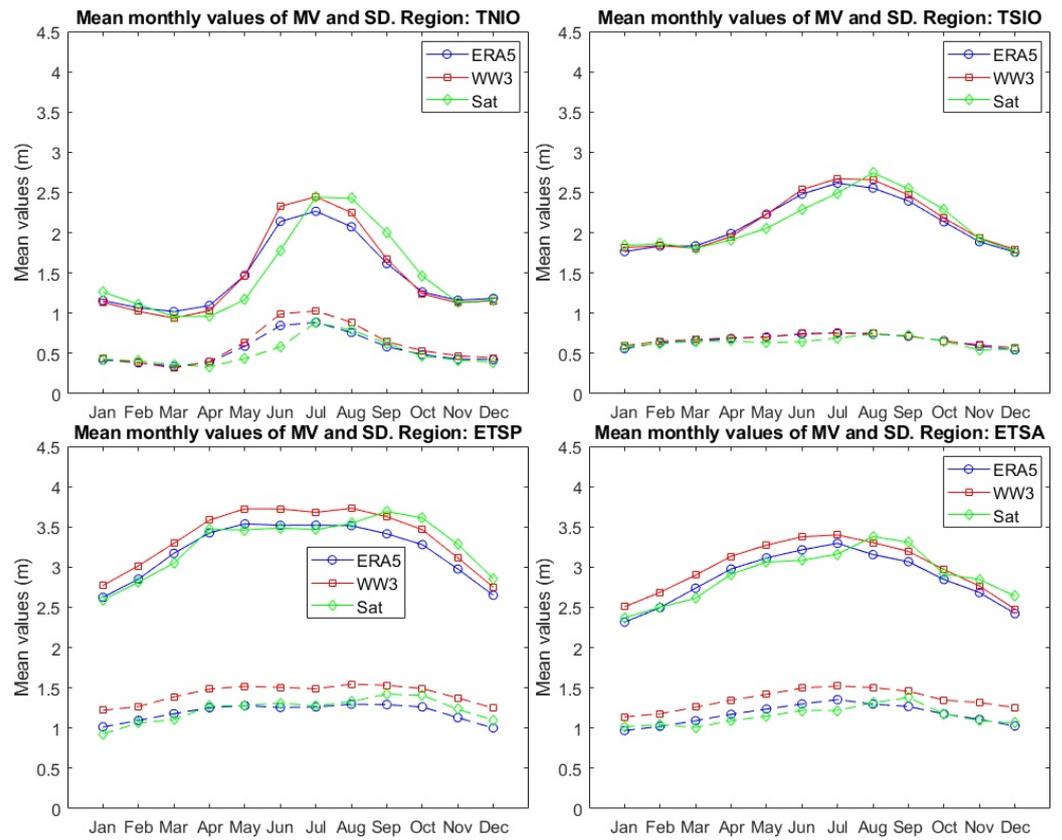


Figure 2. Mean monthly values for the northern extratropical zone (ETNP, ETNA), the tropical zone (TENP, TNAO, TNIO, TSIO) and the southern extratropical zone (ETSP, ETSA); see also Figure 1 (continuous: mean value, dashed: standard deviation).

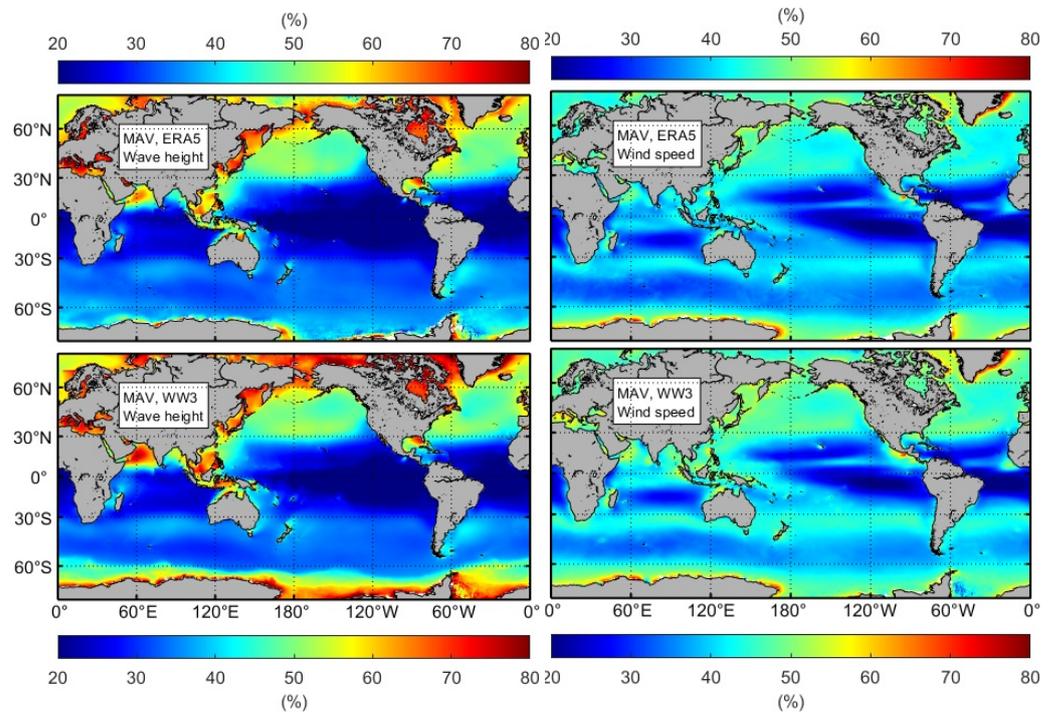


Figure 3. Mean annual variability (in %) (top: ERA5, bottom: WW3, left: significant wave height, right: wind speed).

Finally, the inter-annual variability was studied on the basis of $\hat{c}\hat{v}_{32}(\cdot, \cdot)$; see Equation (14). In Figure 4, the coefficient of variation is plotted for both datasets with WW3 exhibiting greater variability in most of the areas (for waves), and especially in the Southern Hemisphere (for wind). It is worth noted that, the present WW3 results were in accordance with findings in [16].

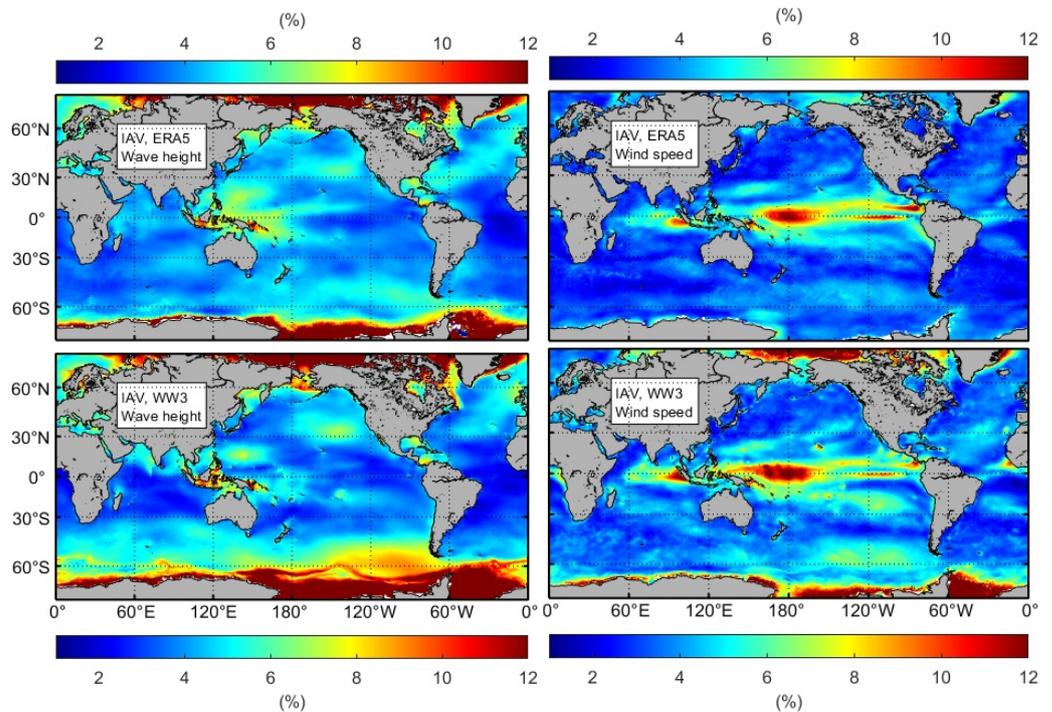


Figure 4. Inter-annual variability (in %) (top: ERA5, bottom: WW3, left: significant wave height, right: wind speed).

3.2. Probability Analysis

In this part of the analysis, the monthly empirical cdf’s of ERA5 and WW3 were calculated, and then, the quantiles of 50% (median), 90%, 95%, 99%, and 99.9% were chosen to described the behaviour of the empirical distribution. Due to the fact that the distribution was not symmetric, the median described the mean behaviour of the distribution better than the mean value. The other four were used for the description of the tail.

In Figure 5 the overall (averaged over all points) mean monthly values of the quantiles were plotted, giving a bird eye’s view of the monthly variability of the distribution. It seems that WW3 distribution had higher values in all quantiles. In addition, ERA5 exhibited greater (month-to-month) variability in the 99.9% quantile. On the average, the differences in each quantile between the two datasets were as follows: (a) significant wave height: $x_{50\%}$: -0.98% , $x_{90\%}$: -9.05% , $x_{95\%}$: -9.21% , $x_{99\%}$: -6.53% , $x_{99.9\%}$: -1.18% , (b) wind speed: $x_{50\%}$: -2.80% , $x_{90\%}$: -5.73% , $x_{95\%}$: -6.78% , $x_{99\%}$: -5.93% , $x_{99.9\%}$: -4.41% , where minus means that ERA5 had lower values than WW3. Further, it should be noted that some behaviors like, e.g., the low variability of the 99% quantile of wind speed might be attributed to the large spatial differences of this quantile, as we will see further below in Figures 6 and 7, which resulted in evening out these differences now where all datapoints were taken into account for the calculation.

In Figures 6 and 7, the mean annual fields of 50%, 90%, 95%, 99% and 99.9% quantiles of significant wave height and wind speed are depicted for ERA5 and WW3. In general, there better than the mean value again a zonal distribution of the quantiles, and ERA5 seemed to exhibit less variability than WW3 (especially the wave height). Additionally, quantile-values in the Southern Hemisphere seemed to be higher than the ones in the Northern counterpart, except from the two (resp. three) last quantiles for waves (resp. wind). The results were in accordance with [29,30], based on altimetry data, and [16], based on WW3 data.

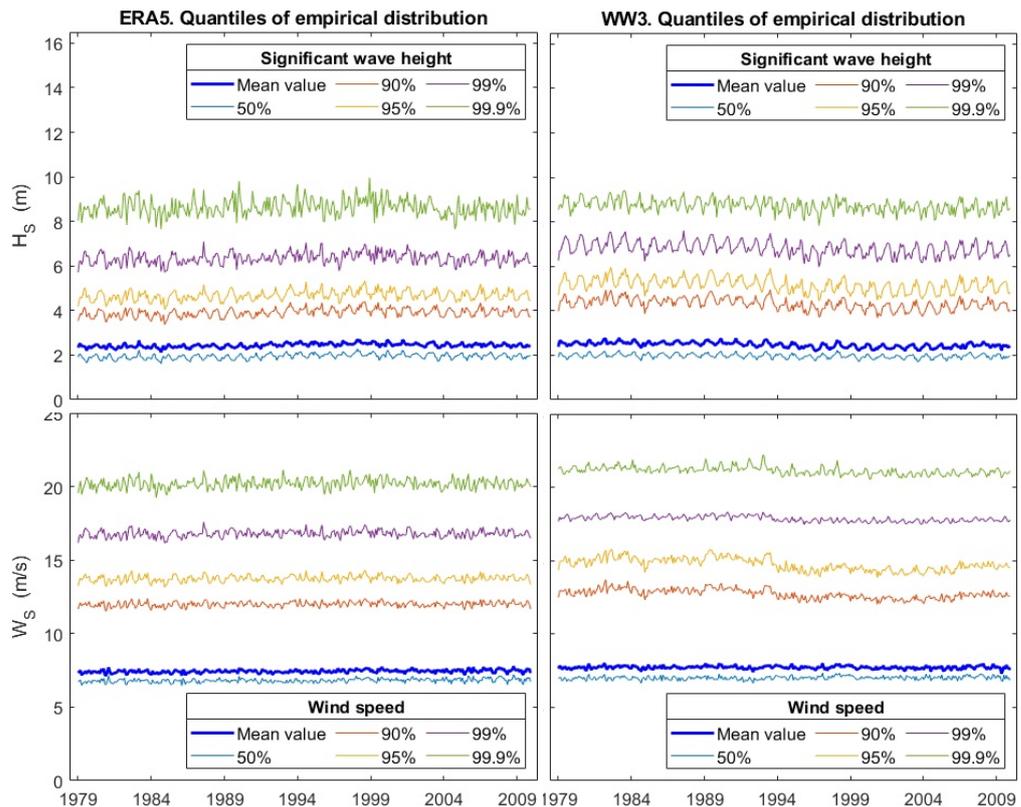


Figure 5. Monthly quantiles 50% (median), 90%, 95%, 99%, and 99.9% taking into account all datapoints (left: ERA5, right: WW3, top: significant wave height, bottom: wind speed).

3.3. Error Analysis

Here an assessment of the differences between the two reanalysis datasets was performed on the basis of the error metrics defined in Section 2.2.3. It is reminded that, in the definition of these metrics, 1 stands for ERA5, and 2 for WW3 dataset. Thus, a “minus” sign in the results means lower ERA5 values with respect to WW3 ones.

In Figure 8, the overall (averaged over all points) mean monthly variability of the error measures is given, namely bias, RMSE, SI, CorrCoeff, and NBias, NSTD, CRMSE (left panels: significant wave height, right panels: wind speed). In addition, in Tables 1 and 2, the monthly values for the plots are given.

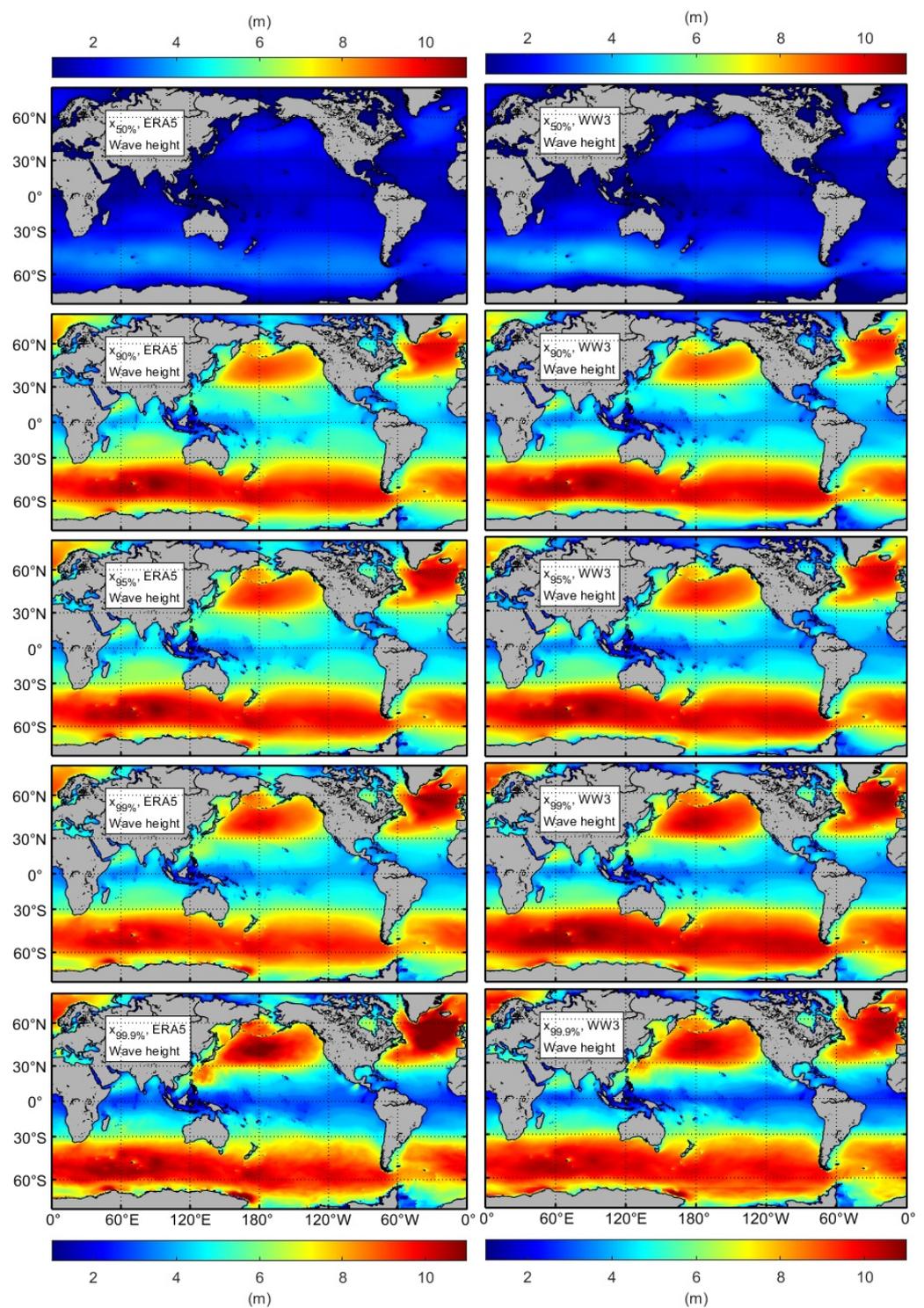


Figure 6. Mean annual fields of quantiles 50%, 90%, 95%, 99% and 99.9% of significant wave height (left: ERA5, right: WW3).

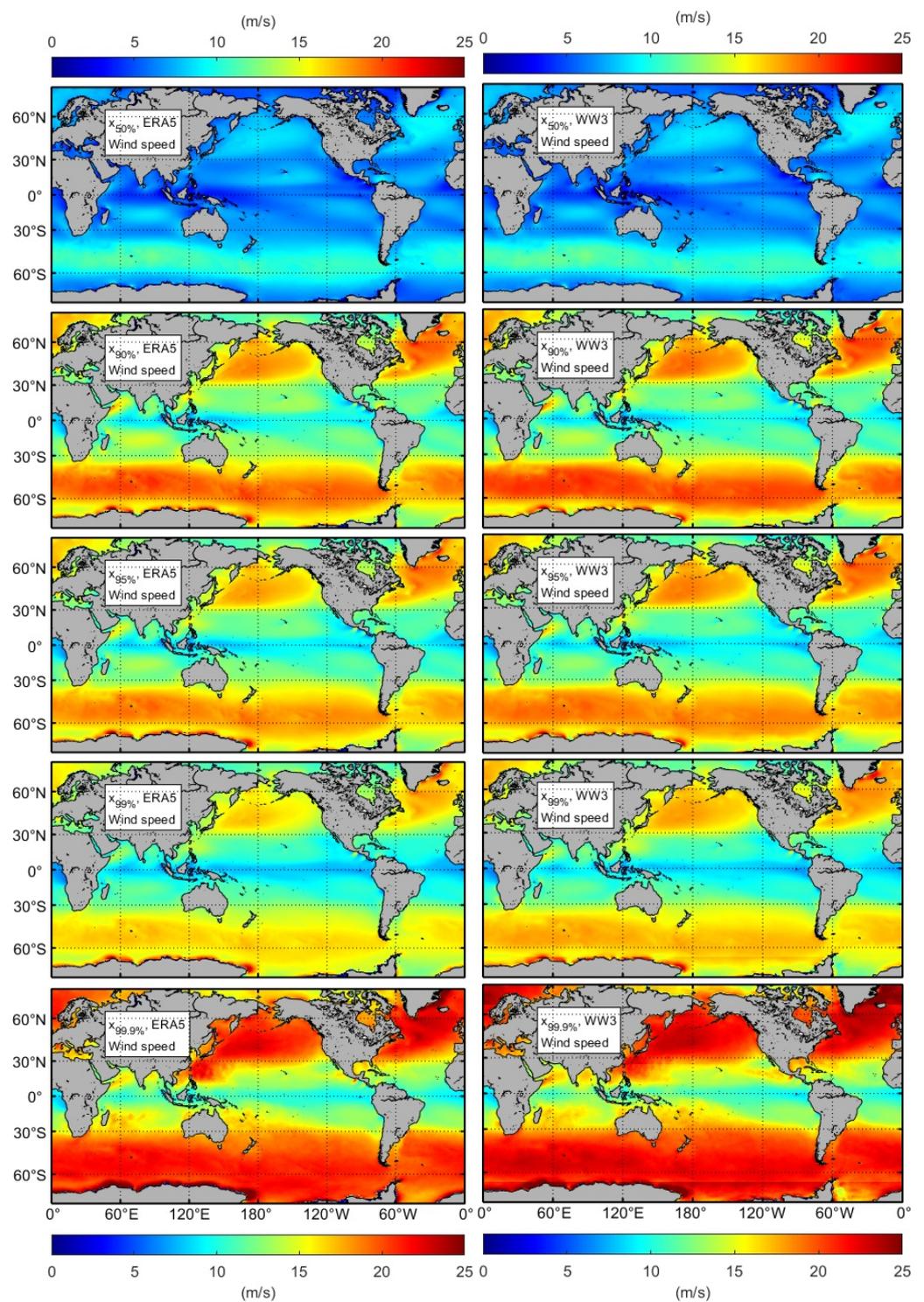


Figure 7. Mean annual fields of quantiles 50%, 90%, 95%, 99% and 99.9% of wind speed (left: ERA5, right: WW3).

Similarly, in Figure 9, the overall (averaged over all points) mean annual variability of the same quantities is also plotted (left panels: significant wave height, right panels: wind speed), and in Table 2, the monthly values are given.

Most of them seemed to have no significant variability at all. One might see a decrease in mean annual bias after 1994, which meant better agreement between the two reanalyses after that year. However, this fact was not considered statistically significant, since the values were relatively low. The same held true if one observed the normalized Bias NBias which showed a more stable behavior throughout the years. It is worth noticing that CorrCoeff was near 1 (0.9 for wave and 0.8 for wind), which means that the two datasets were in good correlation.

Table 1. Mean monthly variability of error measures between the two reanalyses (significant wave height).

Month	Bias	RMSE	SI	CorrCoeff	NBias	NSTD	CRMSE
January	−0.12	0.38	0.17	0.88	−0.04	1.11	0.49
February	−0.11	0.37	0.15	0.89	−0.03	1.08	0.47
March	−0.10	0.36	0.15	0.90	−0.02	1.09	0.46
April	−0.11	0.35	0.15	0.90	−0.03	1.12	0.45
May	−0.13	0.35	0.15	0.90	−0.03	1.14	0.46
June	−0.15	0.36	0.16	0.88	−0.05	1.17	0.49
July	−0.13	0.40	0.18	0.87	−0.05	1.18	0.52
August	−0.16	0.36	0.17	0.89	−0.05	1.19	0.50
September	−0.12	0.35	0.16	0.89	−0.03	1.13	0.47
October	−0.09	0.34	0.15	0.89	−0.02	1.12	0.45
November	−0.09	0.35	0.16	0.88	−0.02	1.15	0.46
December	−0.11	0.37	0.17	0.89	−0.04	1.15	0.48

Table 2. Mean monthly variability of error measures between the two reanalyses (wind speed).

Month	Bias	RMSE	SI	CorrCoeff	NBias	NSTD	CRMSE
January	−0.04	1.62	0.23	0.79	−0.00	1.06	0.62
February	0.00	1.58	0.22	0.80	0.00	1.05	0.61
March	−0.01	1.58	0.22	0.80	0.00	1.06	0.61
April	0.05	1.57	0.22	0.79	0.01	1.03	0.60
May	0.05	1.56	0.22	0.77	0.01	1.03	0.60
June	0.01	1.54	0.22	0.78	0.01	1.02	0.60
July	0.09	1.61	0.23	0.78	0.01	1.03	0.62
August	−0.00	1.57	0.22	0.78	0.00	1.03	0.61
September	0.01	1.59	0.23	0.79	0.00	1.04	0.61
October	0.04	1.58	0.22	0.79	0.01	1.03	0.60
November	0.04	1.57	0.22	0.78	0.01	1.02	0.60
December	−0.01	1.58	0.22	0.79	0.00	1.03	0.61

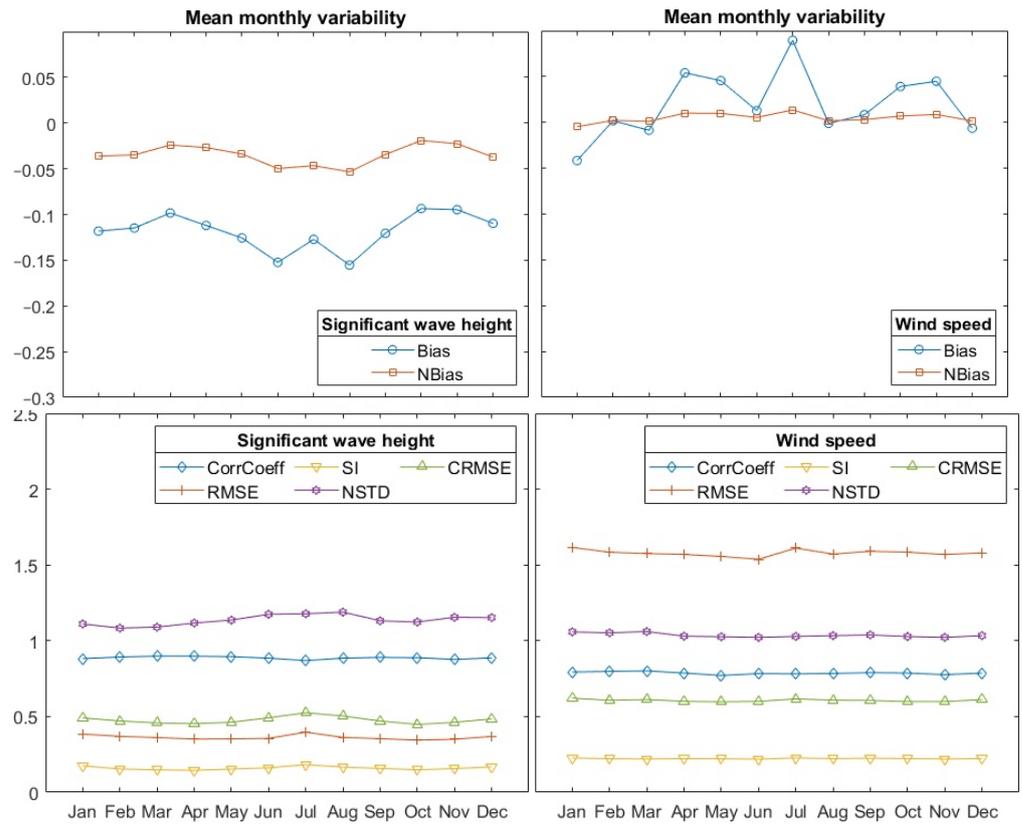


Figure 8. Mean monthly variability of error measures between the two reanalyses (left: significant wave height, right: wind speed).

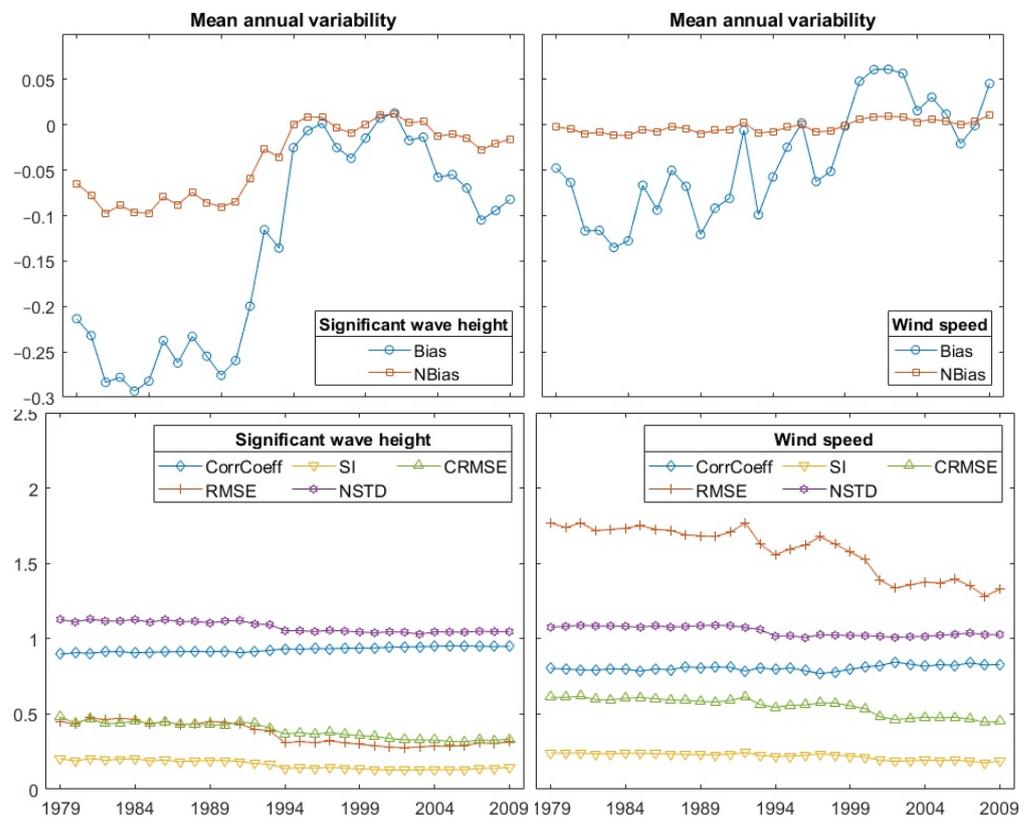


Figure 9. Mean annual variability of error measures between the two reanalyses (left: significant wave height, right: wind speed).

Further, the overall (averaged over all the years) mean annual variability of the same measures is depicted in Figures 10–12. In these figures, one can observe lower values of ERA5 (negative bias and NBias) in the Northern and the Southern Hemisphere, and higher values (positive bias and NBias) in the swell-dominated tropical zone. The greater differences, in absolute value, of the two datasets were exhibited in the Southern Hemisphere, as shown by RMSE. In addition, the study of NSTD revealed that the variability of WW3 was generally higher than the one exhibited by ERA5, while the reverse situation was exhibited in the areas near the coasts of South America for both wind and waves, and additionally near the coast of Canada for waves.

Additionally, following the spatial distribution of SI, it seemed that wind speed exhibited larger variability than wave height, especially in areas around the Equator in Southeast Asia, west coast of Central America and West Africa.

Further, the two datasets were more correlated in the extratropical zones in both the Northern and the Southern Hemisphere according to the CorrCoeff, with waves showing higher values than winds. This was also justified by CRMSE, which exhibited its lowest values (low relative variability of the differences) in the same areas.

One may argue that there was an inconsistency between the spatial distribution of the values of RMSE and those of CorrCoeff. However, the two error metrics should be seen as complementary, giving only partially overlapping (and not exactly the same) statistical information. According to [31], a more convenient measure to be related with CorrCoeff is the CRMSE, which is normalized with the standard deviation of the data; see Equation (25). Indeed, one can observe that in the areas where CorrCoeff suggested low correlation between the two datasets, CRMSE error exhibited its highest values.

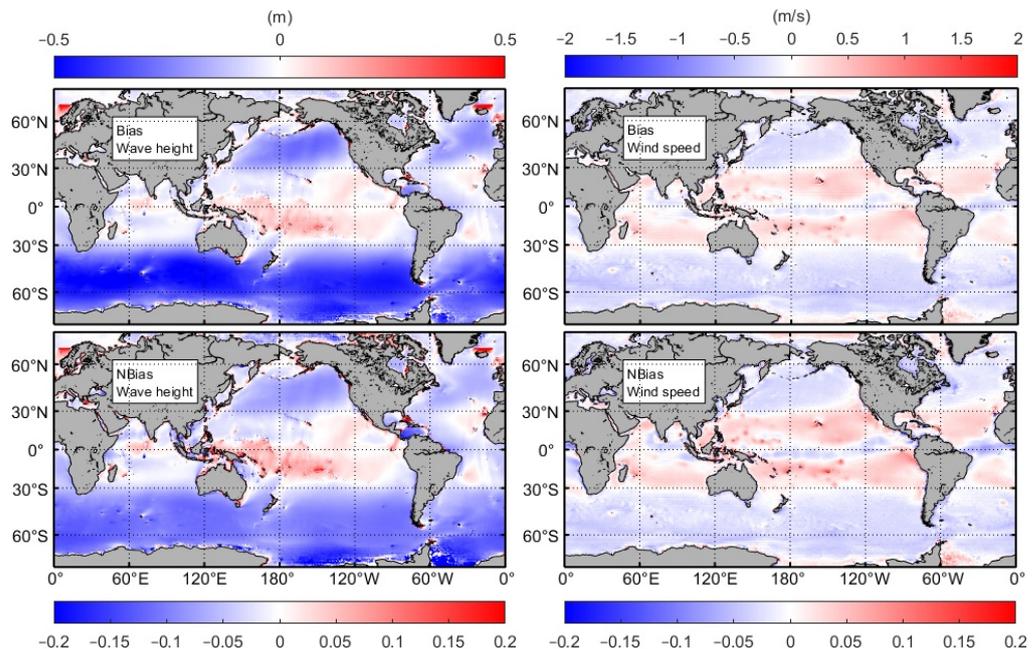


Figure 10. Mean annual variability of Bias, NBIAS (from top to bottom) between the two reanalyses (left: significant wave height, right: wind speed).

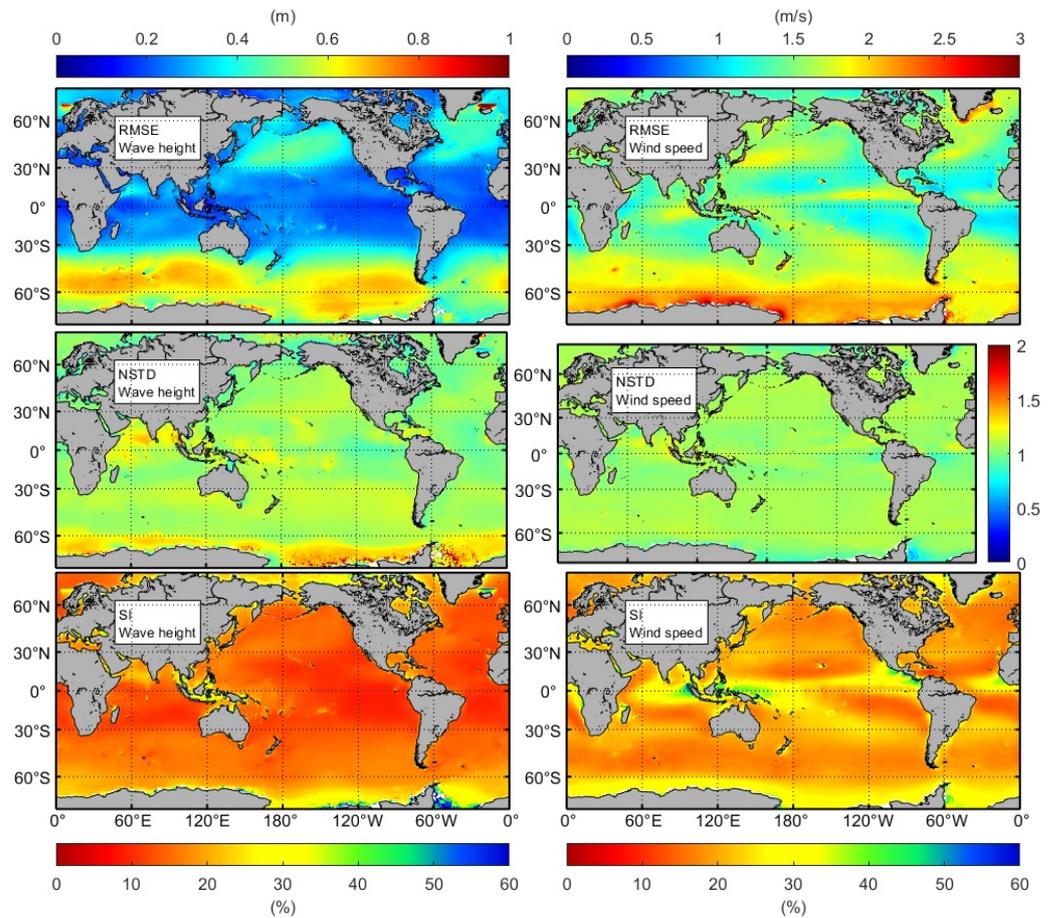


Figure 11. Mean annual variability of RMSE, NSTD, SI (from top to bottom) between the two reanalyses (left: significant wave height, right: wind speed).

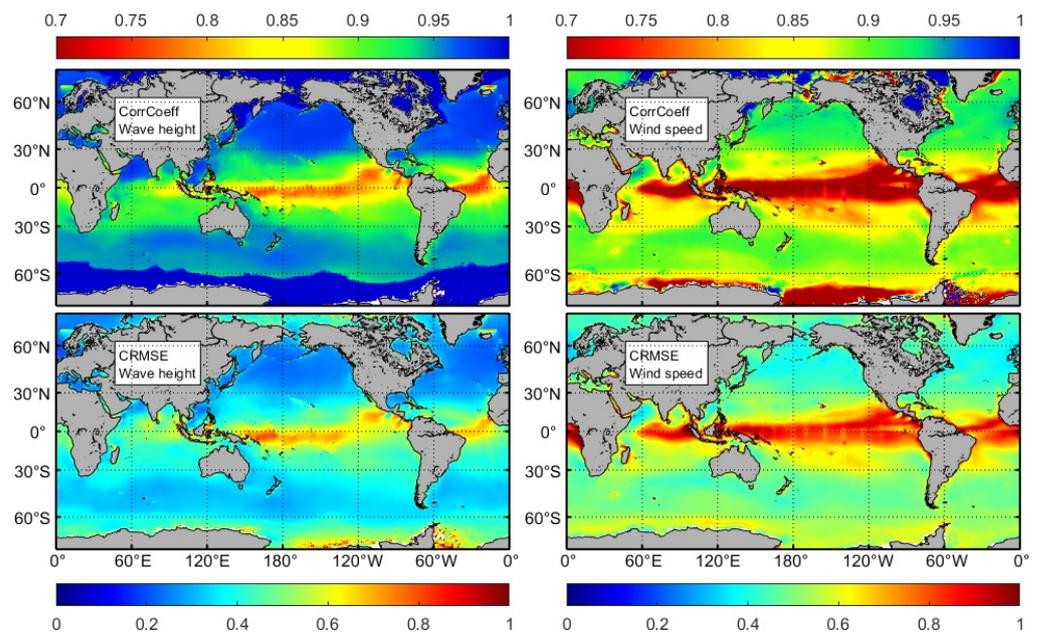


Figure 12. Mean annual variability of CorrCoeff, CRMSE, (from top to bottom) between the two reanalyses (left: significant wave height, right: wind speed).

Finally, the same error measures were calculated for the 13 subregions, using the monthly values of the datasets. The results both for (a) ERA5 vs Satellite, (b) WW3 vs

Satellite are given in Tables 3 and 4. In most of the regions, bias was negative (satellite data had greater values than ERA5 and WW3), except for regions: TESP (ERA5 greater than satellite), and TSAO (both ERA5 and WW3 greater than satellite). The values of RMSE were pretty close for each area. As already mentioned, CorrCoeff was near 1, which was an indication of the correlation of the datasets. NSTD, which showed the ratio of the standard deviations of the two datasets, varied between 1.17–1.71 for “WW3 vs satellite” (satellite had greater variability 17–71%), and 1.19–1.99 for “ERA5 vs. satellite” (satellite showed even greater variability compared to ERA5).

Table 3. Error measures (Bias, RMSE, SI, CorrCoeff) between (a) (E)RA5 and (S)atellite, (b) (W)W3 and (S)atellite for the 13 subregions (significant wave height).

	Bias		RMSE		SI		CorrCoeff	
	E-S	W-S	E-S	W-S	E-S	W-S	E-S	W-S
ETNP	−0.19	−0.21	0.46	0.46	0.19	0.19	0.91	0.90
ETNA	−0.44	−0.50	0.68	0.74	0.29	0.32	0.86	0.84
TNIO	−0.08	−0.11	0.38	0.38	0.26	0.26	0.81	0.80
TWNP	−0.13	−0.17	0.34	0.36	0.20	0.21	0.79	0.79
TENP	−0.01	−0.06	0.21	0.22	0.10	0.11	0.83	0.79
TNAO	−0.09	−0.09	0.23	0.24	0.13	0.14	0.81	0.80
TSIO	−0.07	−0.11	0.28	0.29	0.13	0.14	0.74	0.74
TWSP	−0.13	−0.24	0.26	0.34	0.15	0.20	0.34	0.34
TESP	0.03	−0.03	0.32	0.31	0.14	0.14	0.54	0.56
TSAO	0.00	0.01	0.25	0.26	0.13	0.13	0.58	0.49
ETSI	−0.18	−0.09	0.41	0.42	0.12	0.12	0.77	0.62
ETSP	−0.17	−0.14	0.38	0.41	0.12	0.12	0.79	0.65
ETSA	−0.16	−0.11	0.40	0.40	0.14	0.14	0.68	0.55

Table 4. Error measures (NBias, NSTD, CRMSE) between (a) (E)RA5 and (S)atellite , (b) (W)W3 and (S)atellite for the 13 subregions (significant wave height) (continued).

	NBias		NSTD		CRMSE	
	E-S	W-S	E-S	W-S	E-S	W-S
ETNP	−0.07	−0.08	1.19	1.17	0.53	0.52
ETNA	−0.18	−0.21	1.32	1.42	0.69	0.77
TNIO	−0.05	−0.07	1.42	1.27	0.81	0.70
TWNP	−0.07	−0.10	1.35	1.33	0.91	0.89
TENP	−0.01	−0.03	1.25	1.20	0.79	0.77
TNAO	−0.05	−0.05	1.26	1.31	0.74	0.79
TSIO	−0.03	−0.05	1.35	1.25	0.83	0.77
TWSP	−0.07	−0.14	1.99	1.62	1.86	1.60
TESP	0.01	−0.01	1.85	1.71	1.60	1.44
TSAO	0.00	0.01	1.77	1.59	1.37	1.26
ETSI	−0.05	−0.03	1.34	1.21	0.79	0.80
ETSP	−0.05	−0.04	1.35	1.17	0.89	0.88
ETSA	−0.05	−0.04	1.39	1.28	1.06	1.03

4. Concluding Remarks

In the present paper, the newly released ERA5 wind and wave climatology is compared against WW3 climatology, as well as a merged satellite database. The intercomparison covers a period of 31 years (1979–2009), and has been performed using three-hourly wave fields for the entire globe.

The mean monthly values of the datasets are used as cornerstones of the analysis. First, seasonal analysis is performed based on seasonal characteristics of them, showing a very

good agreement. Moreover, the mean annual and the inter-annual variability computed are in accordance with findings of other researchers [16].

Then, quantiles of the empirical cumulative distribution are calculated in order to get a picture of the variability of the distribution and of several percentiles of interest; especially in the tail of the distribution.

Finally, several error measures are derived in order to assess the agreement between the datasets.

In addition to the field analysis, data are considered for thirteen non-overlapping subregions following [23], and the above mentioned analysis (seasonal, probability, error) is performed for each one of them.

All in all, the two datasets are in a very good agreement, with WW3 having little greater variability than ERA5.

Some particular comments are as follows. The two datasets are well correlated in the extratropical zones in both the Northern and the Southern Hemisphere. ERA5 has lower values than WW3 in the Northern and the Southern Hemisphere, and higher values in the swell-dominated tropical zone. The variability of WW3 is generally higher than the one exhibited by ERA5, while the reverse situation is exhibited in the areas near the coasts of South America for both wind and waves, and additionally near the coast of Canada for waves. It also seems that wind speed exhibits larger variability than wave height, especially in areas around the Equator in Southeast Asia, west coast of Central America and West Africa.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data used for this paper have been retrieved from the following sites: <https://polar.ncep.noaa.gov/waves/hindcasts/nopp-phase2.php>, <https://cds.climate.copernicus.eu/#!/home>, and <ftp://ftp.ifremer.fr/ifremer/cersat/products/swath/altimeters/waves/> (last accessed: 10 September 2021).

Acknowledgments: The work in this paper has been funded by internal basic funds of SINTEF Ocean, Dept of Fisheries and New Biomarine Industry, which are gratefully acknowledged.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Gulev, S.K.; Grigorieva, V. Variability of the Winter Wind Waves and Swell in the North Atlantic and North Pacific as Revealed by the Voluntary Observing Ship Data. *J. Clim.* **2006**, *19*, 5667–5685. [[CrossRef](#)]
2. Lefèvre, J.M.; Cotton, P.D. Ocean surface waves. In *Satellite Altimetry and Earth Sciences*; Fu, L.-L., Cazenave, A., Eds.; Academic Press: San Diego, CA, USA, 2001; Chapter 7; pp. 305–328.
3. Young, I.R.; Zieger, S.; Babanin, A.V. Global Trends in Wind Speed and Wave Height. *Science* **2011**, *332*, 451–455. [[CrossRef](#)]
4. The WAMDI Group. The WAM model-A third generation ocean wave prediction model. *J. Phys. Oceanogr.* **1988**, *18*, 1775–1810. [[CrossRef](#)]
5. Tolman, H.L.; Balasubramanian, B.; Burroughs, L.D.; Chalikov, D.V.; Chao, Y.Y.; Chen, H.S.; Gerald, V.M. Development and Implementation of Wind-Generated Ocean Surface Wave Models at NCEP. *Weather Forecast.* **2002**, *17*, 311–333. [[CrossRef](#)]
6. Stopa, J.E. Wind forcing calibration and wave hindcast comparison using multiple reanalysis and merged satellite wind datasets. *Ocean Model.* **2018**, *127*, 55–69. [[CrossRef](#)]
7. Sterl, A.; Caires, S. Climatology, Variability and Extrema of Ocean Waves: The Web-based KNMI/ERA-40 Wave Atlas. *Int. J. Climatol.* **2005**, *25*, 963–977. [[CrossRef](#)]
8. Rasle, N.; Ardhuin, F. A global wave parameter database for geophysical applications. Part 2: Model validation with improved source term parameterization. *Ocean Model.* **2013**, *70*, 174–188. [[CrossRef](#)]
9. Cox, A.T.; Swail, V.R. A global wave hindcast over the period 1958–1997: Validation and climate assessment. *J. Geophys. Res. Ocean.* **2001**, *106*, 2313–2329. [[CrossRef](#)]
10. Chawla, A.; Spindler, D.M.; Tolman, H.L. Validation of a thirty year wave hindcast using the Climate Forecast System Reanalysis winds. *Ocean Model.* **2013**, *70*, 189–206. [[CrossRef](#)]

11. Guedes Soares, C.; Weisse, R.; Carretero, J.C.; Alvarez, E. A 40 Year Hindcast of Wind, Sea Level and Waves in European Waters. In Proceedings of the 21st International Conference on Offshore Mechanics and Arctic Engineering, OMAE'2002, Oslo, Norway, 23–28 June 2002; pp. 669–675. [[CrossRef](#)]
12. Reguero, B.G.; Menéndez, M.; Méndez, F.J.; Mínguez, R.; Losada, I.J. A Global Ocean Wave (GOW) calibrated reanalysis from 1948 onwards. *Coast. Eng.* **2012**, *65*, 38–55. [[CrossRef](#)]
13. Perez, J.; Menendez, M.; Losada, I.J. GOW2: A global wave hindcast for coastal applications. *Coast. Eng.* **2017**, *124*, 1–11. [[CrossRef](#)]
14. Caires, S.; Sterl, A.; Bidlot, J.R.; Graham, N.; Swail, V. Intercomparison of different wind wave reanalyses. *J. Clim.* **2004**, *17*, 1893–1913. [[CrossRef](#)]
15. Semedo, A.; Sušelj, K.; Rutgersson, A.; Sterl, A. A Global View on the Wind Sea and Swell Climate and Variability from ERA-40. *J. Clim.* **2011**, *24*, 1461–1479. [[CrossRef](#)]
16. Stopa, J.E.; Cheung, K.F.; Tolman, H.L.; Chawla, A. Patterns and cycles in the Climate Forecast System Reanalysis wind and wave data. *Ocean Model.* **2013**, *70*, 207–220. [[CrossRef](#)]
17. Stopa, J.E.; Cheung, K.F. Intercomparison of wind and wave data from the ECMWF Reanalysis Interim and the NCEP Climate Forecast System Reanalysis. *Ocean Model.* **2014**, *75*, 65–83. [[CrossRef](#)]
18. Campos, R.M.; Guedes Soares, C. Comparison and assessment of three wave hindcasts in the North Atlantic Ocean. *J. Oper. Oceanogr.* **2016**, *9*, 26–44. [[CrossRef](#)]
19. Hersbach, H.; Bell, B.; Berrisford, P.; Hirahara, S.; Horányi, A.; Muñoz-Sabater, J.; Nicolas, J.; Peubey, C.; Radu, R.; Schepers, D.; et al. The ERA5 global reanalysis. *Q. J. R. Meteorol. Soc.* **2020**, *146*, 1999–2049. [[CrossRef](#)]
20. Stefanakos, C. Intercomparison of Wave Reanalysis based on ERA5 and WW3 Databases. In Proceedings of the 29th International Offshore and Polar Engineering Conference, ISOPE'2019, Honolulu, HI, USA, 16–21 June 2019; pp. 2506–2512.
21. Tolman, H. A Third-Generation Model for Wind Waves on Slowly Varying, Unsteady, and Inhomogeneous Depths and Currents. *J. Phys. Oceanogr.* **1991**, *21*, 782–797. [[CrossRef](#)]
22. Queffeuilou, P.; Croizé-Fillon, D. Global altimeter SWH data set. In *Technical Report, Laboratoire d' Océanographie Physique et Spatiale*; IFREMER: Plouzané, France, 2017.
23. Alves, J.H.G. Numerical modeling of ocean swell contributions to the global wind-wave climate. *Ocean Model.* **2006**, *11*, 98–122. [[CrossRef](#)]
24. Hemer, M.A.; Trenham, C.E. Evaluation of a CMIP5 derived dynamical global wind wave climate model ensemble. *Ocean Model.* **2016**, *103*, 190–203. [[CrossRef](#)]
25. Stefanakos, C.; Athanassoulis, G.; Barstow, S. Time series modeling of significant wave height in multiple scales, combining various sources of data. *J. Geophys. Res. Sect. Ocean.* **2006**, *111*, 10001–10012. [[CrossRef](#)]
26. Athanassoulis, G.; Stefanakos, C. A nonstationary stochastic model for long-term time series of significant wave height. *J. Geophys. Res. Sect. Ocean.* **1995**, *100*, 16149–16162. [[CrossRef](#)]
27. Stefanakos, C.; Schinas, O. Forecasting bunker prices; A nonstationary, multivariate methodology. *Transp. Res. Part C Emerg. Technol.* **2014**, *38*, 177–194. [[CrossRef](#)]
28. Jiang, H. Evaluation of altimeter undersampling in estimating global wind and wave climate using virtual observation. *Remote Sens. Environ.* **2020**, *245*, 111840. [[CrossRef](#)]
29. Young, I. An intercomparison of GEOSAT, TOPEX and ERS1 measurements of wind speed and wave height. *Ocean Eng.* **1999**, *26*, 67–81. [[CrossRef](#)]
30. Izaguirre, C.; Méndez, F.J.; Menéndez, M.; Losada, I.J. Global extreme wave height variability based on satellite data. *Geophys. Res. Lett.* **2011**, *38*. [[CrossRef](#)]
31. Taylor, K.E. Summarizing multiple aspects of model performance in a single diagram. *J. Geophys. Res. Atmos.* **2001**, *106*, 7183–7192. [[CrossRef](#)]