



Research paper

Global evaluation of wave data reanalysis: Comparison of the ERA5 dataset to buoy observations

Victoria Bessonova ^{a,b}, Evdokia Tapoglou ^{b,c}, Robert Dorrell ^b, Nina Dethlefs ^b, Katharine York ^d

^a *Aura Centre for Doctoral Training, Kingston-upon-Hull, HU6 7RX, UK*

^b *University of Hull, Kingston-upon-Hull, HU6 7RX, UK*

^c *European Commission, Joint Research Centre, Petten, 1755 LE, The Netherlands*

^d *Offshore Renewable Energy Catapult, Grimsby, DN31 3LL, UK*

ARTICLE INFO

Keywords:

Wave data
Reanalysis
Buoy measurements
ERA5

ABSTRACT

Wave reanalysis data are critical to multiple industries including commerce, fisheries, petroleum, offshore wind, the design and the construction of maritime structures. One of the most used datasets is ERA5 developed by the European Centre for Medium Range Forecast. The accuracy and reliability of this dataset have not been rigorously assessed on a global scale. We present the first global evaluation of ERA5 significant wave height against measurements from 444 buoys worldwide. We demonstrate the influence of critical control factors including distance to shore, water depth, regional characteristics and seasonality. Results showed that ERA5 underestimates significant wave height in extreme conditions. This underestimation could lead to underestimating the severity of predicted wave conditions when designing vessels or foundations for maritime structures. ERA5 accuracy increases farther from shore, but does not change significantly with increasing water depth. In the Northern Hemisphere, the RMSE and absolute bias between ERA5 and measured significant wave height are up to 0.25 m lower in summer than in winter. Seasonal patterns are less pronounced in the Southern Hemisphere. The results provide global confidence levels for the robust use of ERA5 data, and establish a methodology to test future datasets.

1. Introduction

Accurate wave data are required for many applications, including coastal management, marine energy planning, the design and the construction of maritime structures and validation of wave projections (Storlazzi et al., 2018; Izaguirre et al., 2020; Li et al., 2022; Dowdy et al., 2014, e.g.). Historical wave height, wave direction and wave period can be obtained from direct observations such as drifting or moored buoys, voluntary observing ships or satellite altimetry (Wells, 2011).

Satellite measurement campaigns began in the 1980s, with increasing coverage and duration towards the 2000s–2010s (Ribal and Young, 2019). Satellites measure data along their tracks on Earth and repeat these tracks with various frequencies. These measurements have not been continuous globally, with the biggest gap between late-1989 and mid-1991. Moreover, the data from different satellites is not uniform in the calibration methods and data formats. These issues were mitigated

in an integrated dataset created by Ribal and Young (2019), but the temporal resolution remains insufficient for sub-daily analysis due to long orbiting times (Table 1).

Buoy measurements started in some regions as early as the 1970s and increased in frequency and prevalence particularly in the 1990s. Among a wide variety of buoys (e.g. Axyz Next Wave II, strapped-down accelerometers, etc.), Datawell directional Waverider is one of the most common types (Jensen et al., 2015). Over the history of buoy measurements, there were changes in buoy hulls and wave sensors that inevitably impact the quality of the data. These changes can lead to false conclusions about the wave climate if not taken into account (Gemmrich et al., 2011). Observations are also limited in spatial and temporal resolution and coverage (Table 1).

Studies of long-term global climate often use reanalysis datasets to overcome limitations of observational data (Liléo and Petrik, 2011; Sharmer and Markina, 2020, e.g.). Reanalysis is a modelled hindcast

* Corresponding author at: University of Hull, Kingston-upon-Hull, HU6 7RX, UK.

E-mail address: v.bessonova-2020@hull.ac.uk (V. Bessonova).

URL: <https://auracdt.hull.ac.uk/victoria-bessonova/> (V. Bessonova).

Table 1
Comparison of reference wave data.

| Dataset | Resolution | | Temporal coverage, years | Source |
|------------|------------|------------------------------|--------------------------|---|
| | Temporal | Spatial | | |
| Buoys | 0.5–1 h | Varied | 2–42, see Fig. 4 | Cefas (2021), NOAA (2021), Copernicus Marine Service (2021) and AODN (2021) |
| Satellites | 3–10 days | $\leq 3.5^\circ$ | 1985–1990, 1992–2024 | Ribal and Young (2019) |
| ERA5 | 1 h | $0.5^\circ \times 0.5^\circ$ | 1940–2024 | Hersbach et al. (2020) |
| WAVERYS | 3 h | 0.2° | 1980–2024 | CMEMS (2023) |

which is combined with various measurements through data assimilation (Compo et al., 2011). Reanalysis datasets vary in their accuracy due to differences in the underlying numerical models and data assimilation methods.

The temporal resolution of buoy and reanalysis data makes them suitable to analyse changes in diurnal patterns, which is important for understanding changing climate and could be used in studies of fatigue stress in marine structures or weather delays in marine operations (O'Connor et al., 2013). Buoy data availability varies significantly on a temporal and spatial scale, which leads to a more complex analysis of geographical patterns. Buoys, however, can produce low-quality measurements due to sensor malfunction or accidental unmooring of the device. These issues can affect the measurements over the long-term, since the maintenance requires marine operations.

There are two wave reanalysis products mainly used in marine industries and research, ECMWF ReAnalysis — ERA5 (Hersbach et al., 2020) and WAVERYS (CMEMS, 2023). ERA5 was developed by ECMWF (European Centre for Medium-Range Weather Forecast) and superseded ERA-40 (Källberg et al., 2004) and ERA-Interim (Berrisford et al., 2011). These three datasets resulted from the assimilation of high-resolution climate modelling with ground (except wave height) and satellite observations. Hersbach et al. (2020) do not report if the satellite measurements were calibrated prior to data assimilation as advised by Ribal and Young (2019), thus it is not clear if the altimetry errors and inconsistencies contribute to ERA5 accuracy.

The WAVERYS dataset was created by Copernicus Marine Service through numerical modelling and data assimilation with Sentinel-1 altimetry data. Unlike ERA5, WAVERYS does not have any atmospheric variables and benefits from higher spatial resolution of the wave climate data. On the other hand, this dataset has lower temporal resolution and shorter temporal coverage (Table 1). The latter could be a drawback in cases where reanalysis is used to validate or fine-tune wave models for longer-term climate studies (Wang et al., 2015; Semedo et al., 2018, e.g.).

Both ERA5 and WAVERYS provide significant wave height of combined wind waves and swell, peak wave period, mean wave direction, significant height of wind waves and height, period and direction of the first and second swell partitions. Additional variables in ERA5 which are not present in WAVERYS are height, period and direction of the third swell partition, mean zero crossing wave period and maximum individual wave height.

Given the widespread use of the ERA5 dataset in research and industry, understanding its accuracy and potentially weak areas is paramount. Hersbach et al. (2020) evaluated the ERA5 dataset, although did not report the number of buoys or their locations. ERA5 wave data show an improved agreement with observation compared to ERA-Interim (Hersbach et al., 2020; Bruno et al., 2020) due to better modelling of swell attenuation and bathymetry. Globally, the ERA5 dataset is 4%–5% closer to buoy observations than ERA-Interim. ERA5 is also up to 8% more accurate in the last 20 years compared to the previous 20 years (Hersbach et al., 2020).

Limited analysis of ERA5 performance that has been done to date was only at a regional scale. Similar to Hersbach et al. (2020) and Wang and Wang (2021) showed an increased ERA5 accuracy between 1980 and 2020 in the USA with a decreasing SI (-0.03), RMSE (-0.1 m) and

an increasing correlation (0.03). The increased accuracy could be explained by a wider availability of satellite data in recent decades (Ribal and Young, 2019). Since ERA5 is assimilated with satellite altimetry, higher availability of such measurements could improve the accuracy of the resulting dataset. Hisaki (2020), Wang and Wang (2021) and Shi et al. (2021) reported an underestimation of extreme wave heights by the ERA5 dataset, whereas Bruno et al. (2020) found an overestimation of significant wave height in a swell-dominated region. Bruno et al. (2020), however, based their analysis on a single buoy and their results may not be representative of the whole dataset.

Several studies explored potential factors that could affect the ERA5 accuracy, such as water depth, distance to shore and seasonal variations. Wang and Wang (2021) and Shi et al. (2021) showed that the correlation increases, while scatter index and bias decrease further from shore and in deeper waters. Hisaki (2020) reported a lower accuracy of ERA5 in fetch-limited conditions; however, Wang and Wang (2021) argued that ERA5 resolution is not sufficient for analysis of local conditions such as fetch or island shadowing. The only study with a dataset that covered a sufficiently long time span to examine the seasonal variations in accuracy is Wang and Wang (2021). Their assessment showed lower RMSE and absolute bias in summer, although the correlation was also lower, suggesting that absolute values of ERA5 are closer to the measurements in summer, but ERA5 might be less accurate in capturing the trends in the measurements.

The performance of ERA5 wave period has only been considered by Zhai et al. (2023) in the South China Sea. They concluded that ERA5 generally underestimates significant wave height (H_s) by up to 0.28 m and overestimates significant wave period (T_s) based on the 2-year data from six buoys. Interestingly, Li et al. (2022) analysed data from a buoy at one of the locations in Zhai et al. (2023) for a different period of 16 months and found 0.14 m larger bias for H_s and 0.5 s smaller bias for T_s , showing that comparison of ERA5 to buoy measurements is sensitive to the choice of the time period. Zhai et al. (2023) suggested that differences between ERA5 and buoys are caused by the inability of the ERA5 model to capture the influence of small islands on wave propagation, while Li et al. (2022) proposed a possible impact of the wet monsoon on the accuracy of the significant wave height reanalysis.

Our study provides the first-ever global benchmarking of the ERA5 significant wave height data via comparison with a globally unique wave buoy network. This study's findings support industry use, and academic study, of wave forecasting, global wave conditions and climate change impacts.

2. Data and methods

2.1. Data processing

The buoy measurements of wave parameters were not used in the assimilation of the ERA5 dataset, therefore, the buoy and ERA5 datasets are independent and suitable for comparison. We collected and quality-checked buoy measurements from several sources and then compared them to ERA5 time-series in corresponding locations.

We applied the methodology described in this section to reduce the dataset from 920 global locations to 444 locations, based on the quality of measurement, distance to shore and other factors. Generic significant

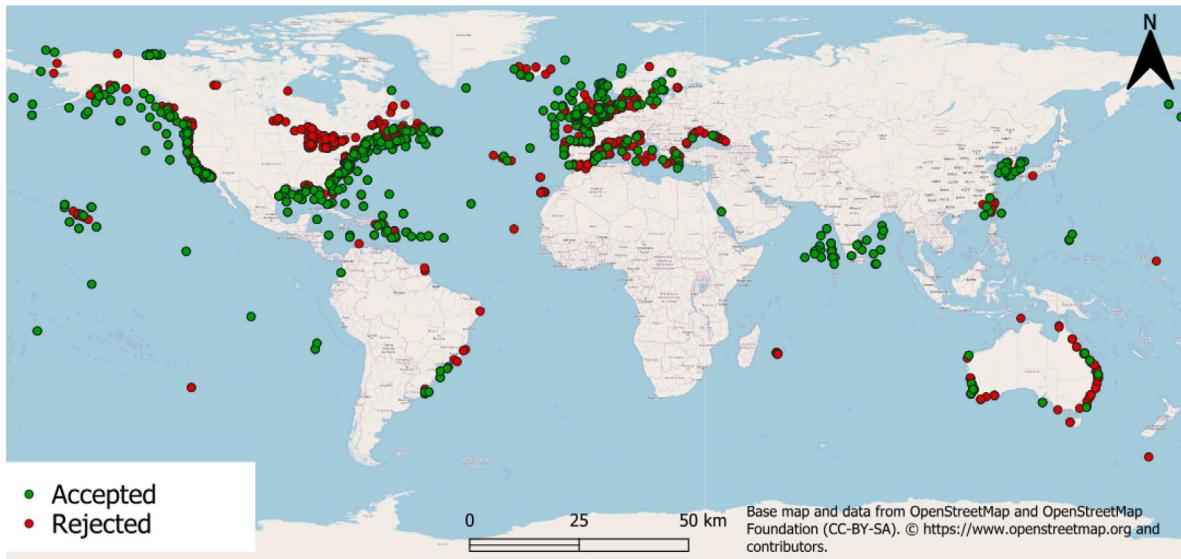


Fig. 1. Locations of the buoys used for analysis are shown in green. Buoys marked in red were rejected due to absence of significant wave height measurements, corresponding ERA5 data or having less than 3 months of measurements.

wave height and spectral significant wave height were used for the analysis and were treated as equivalent.

The data in Australia and the UK was recorded by *Datawell Waverider* buoys, which have 0.01 m wave height measurement accuracy (Cefas, 2021; AODN, 2021). Copernicus Marine Service (2021) aggregates the data from several measurement and research centres, which employ Waverider, *Oceanor Wavesense* and *Axys Triaxys* buoys. Alfonso et al. (2022) report 0.5%–1% uncertainty for measured heave and up to 5% uncertainty for estimated wave height, but do not report the accuracy for measured wave heights. Here, we assume that the measurement error does not exceed 0.1 m for all measurements based on the maximum error of typical wave measuring equipment (Lawrence et al., 2012).

Fifty-eight buoys had more than one H_s variable (e.g. both generic significant wave height and spectral significant wave height). The mean correlation between both variables in all such buoys is 0.99, so only the variable with the fewest data gaps was kept for further analysis.

Anomalously high values of H_s were detected in some buoy locations. Visual inspection of the corresponding time series showed that these values were erroneous (e.g. timesteps with values of 9999), and any values higher than 18 m were excluded from the measurements. These values accounted for less than 0.1% for any of the single locations. The threshold of 18 m may seem excessively high, however, the visual inspection of all measurements with $12 \leq H_s \leq 18$ m confirmed that these waves occurred in remote offshore locations during winter months (e.g. see Fig. 2). The time series also had comparably high H_s both preceding and following the maximum. This inspection, together with previous reports of significant wave height above 15 m during North Atlantic storms (Hanafin et al., 2012), allowed to conclude that all of the occurrences of $H_s \leq 18$ m are not erroneous.

The temporal frequency of the data varies for different locations of the dataset, from 10 min to 1 h and some time series have gaps in the measurements up to several years. The data were resampled to a 1-hr frequency to allow straightforward comparison with ERA5 data. The data with higher frequency was averaged hourly, and the data gaps, introduced when sampled at a lower frequency, were set to NaN (Not a Number). The equivalent number of years was found as the total number of records after resampling divided by 8760. All buoys with less than 3 months' worth of data were excluded from further analysis.

The ERA5 dataset was linearly interpolated to the buoys' locations. 200 buoys located close to coastlines are missing corresponding ERA5 values due to the relatively coarse spatial resolution of the dataset. The

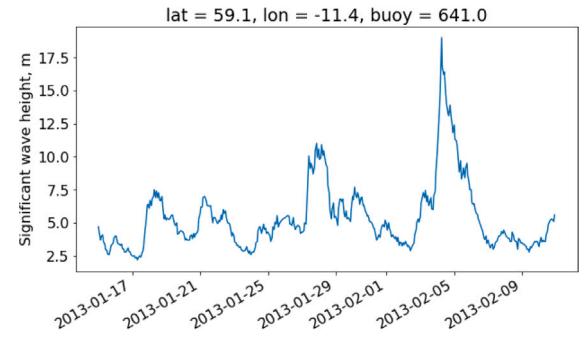


Fig. 2. An example of time series for a buoy with H_s measurements above 12 m.

final dataset includes 444 locations, as displayed in Fig. 1. The criteria used for buoy selection are:

- the location does not change by more than 0.5° in any direction;
- the data for at least 3 months' are available;
- there is a corresponding ERA5 time series after interpolation of the ERA5 dataset.

2.2. Performance evaluation

The main statistical metrics that can be used for the comparison of two datasets include Root Mean Squared Error (RMSE), HH metric, bias, scatter index (SI) and correlation (R^2). RMSE shows the spread of differences between the datasets, positive or negative bias represents an overestimation or underestimation of the values by a model. HH metric is similar to RMSE but it is less biased to negatively biased datasets. SI is a model error shown relative to the dataset mean and correlation can be understood as a measure of similarity between fluctuations of two datasets.

SI can be ambiguous since some researchers define it as RMSE divided by observations' mean (Wang and Wang, 2021; Shi et al., 2021, e.g.), while others (Mentaschi et al., 2013; Hisaki, 2020, e.g.) define it as

$$SI = \frac{\sqrt{\frac{1}{n} \sum_{i=1}^n \{(E(i) - \bar{E}) - [B(i) - \bar{B}]\}^2}}{\bar{B}},$$

where n is the total number of observations, i is the ordinal number of each observational pair, E and B are ERA5 (or other forecast/modelled data) and buoy observations respectively. Bryant et al. (2016) previously highlighted this issue, also noting the limitations of using a single metric from the aforementioned list. This study calculates both SI versions to allow comparison to previous research and understand the sensitivity of both metrics to ERA5 accuracy.

Bryant et al. (2016) also propose to complement such analyses with performance scores, e.g. the Willmott index (Willmott et al., 1985) or Interactive Model Evaluation and Diagnostic System (Hanson et al., 2009), however, such scores are as subjective as more traditional metrics. Mentaschi et al. (2013) also criticise the use of RMSE-based metrics, showing that they can falsely suggest a better performance for models that tend to underestimate environmental variables. They show that the HH metric (Hanna and Heinold, 1985) estimates model accuracy more reliably. Previous assessments of ERA5 performance have not used these metrics, hence it would be hard to pinpoint the metric values that would indicate good or bad performance. However, we include the HH index in case the ECMWF wave model underestimates ERA5 which could cause low RMSE and bias as highlighted by Mentaschi et al. (2013).

The formulae and accepted values for metrics used in this study are described in Table 2. We do not provide an acceptable value for RMSE since it depends on the range and the mean of the data and can be quite subjective. Here, it is used to compare the error in different regions or in different H_s bins. These metrics were calculated for each location and each year and later accumulated by regions defined by Alves (2006) based on the main swell-generating regions (Fig. 3). These regions are characterised by the frequency and intensity of storms generated within the regions and propagated between them. Alves (2006) did not provide the coordinates for region boundaries, so this study approximated these boundaries based on 3. The following latitudes were used for separating tropical and extratropical regions:

- 30°N for the North Pacific (ETNP-TENP/TWNP),
- 38°N for the North Atlantic (ETNA-TNAO),
- 30°S for all southern regions;

and the following longitudes for the east–west separation:

- 150°E for ETSI-ETSP,
- 150°W for TWSP-TESP,
- 180°E for TENP-TWNP.

Alves (2006) did not include the Mediterranean region in their assessment, possibly since the propagation of swell from the North Atlantic is limited, and the region has mostly wind waves (Toomey et al., 2022). Nevertheless, the Mediterranean has strong cyclones and storms causing high waves (Amores et al., 2020). These cyclones resemble the tropical cyclones due to the heat and moisture trapped between the mountainous land mass surrounding the Mediterranean Sea (Flaounas et al., 2015). The Mediterranean cyclones are not as long lived as the North Atlantic ones (mean lifetime is less than a day (Campins et al., 2011), compared to multiple days for the North Atlantic). In this paper, the separation between ETNA and TNAO regions was continued in the Mediterranean. However, the statistics were additionally calculated for the Mediterranean region as a whole to compare them to both TNAO and ETNA (Table 3).

These regions have drastically different distributions of buoys and duration of the available measurements. Fig. 4 demonstrates that the North Atlantic and the North Pacific have measurements spanning the largest number of years with the largest number of buoys. Fig. 1 shows that most of these measurements are concentrated on the North American and European coasts.

Distributions of two datasets can be compared using goodness-of-fit tests. Commonly used non-parametric tests include Kolmogorov-Smirnov (Kolmogoroff, 1941; Smirnov, 1948), Cramer-von-Mises

Table 2

Statistical metrics for comparison. RMSE — root mean squared error, SI — scatter index, CC — Pearson's correlation, HH — HH index (Hanna and Heinold, 1985).

| Name | Equation | Target value |
|--------------------|---|--------------|
| RMSE, m | $\sqrt{\frac{1}{n} \sum_{i=1}^n [E(i) - B(i)]^2}$ | — |
| Bias, m | $\frac{1}{n} \sum_{i=1}^n [E(i) - B(i)]$ | -0.1–0.1 |
| Normalised bias, m | $\frac{\sum_{i=1}^n [E(i) - B(i)]}{\sum_{i=1}^n B(i)}$ | — |
| SI 1 | $\sqrt{\frac{\frac{1}{n} \sum_{i=1}^n \{[E(i) - \bar{E}] - [B(i) - \bar{B}]\}^2}{\bar{B}}}$ | 0.25 |
| SI 2 | $\sqrt{\frac{\frac{1}{n} \sum_{i=1}^n [E(i) - \bar{E}]^2}{\bar{B}}}$ | 0.25 |
| CC | $\frac{\sum_{i=1}^n (E(i) - \bar{E})(B(i) - \bar{B})}{\sqrt{\sum_{i=1}^n (E(i) - \bar{E})^2} \sqrt{\sum_{i=1}^n (B(i) - \bar{B})^2}}$ | 0.7 |
| HH | $\sqrt{\frac{\sum_{i=1}^n [E(i) - B(i)]^2}{BB}}$ | — |

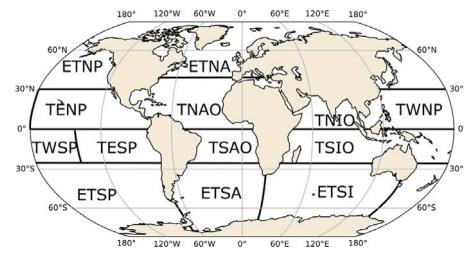


Fig. 3. Swell-generating regions, adapted from Alves (2006). In abbreviations, ET stands for extratropical, T - tropical, N/S/W/E - North/South/West/East, P - Pacific, A(O) - Atlantic (Ocean), I(O) - Indian (Ocean).

(Cramer, 1928; von Mises, 1931), Anderson–Darling (Anderson and Darling, 1952) and χ^2 tests (Chernoff and Lehmann, 1954).

These tests assume that the values in each dataset are independent and identically distributed. Environmental variables are rarely independent, as both temporal and spatial events affect the development of weather systems. However, the aforementioned tests are frequently used by the climate research community (Dookie et al., 2018; Rad et al., 2017; Chen, 2022), including in studies of ocean waves (Ferreira and Guedes Soares, 2000; Caires and Sterl, 2005). Visual inspection of the wave data during processing confirmed that the data do not display significant trends, hence it is identically distributed. In this study, the temporal correlations are assumed insignificant and the samples are considered independent and identically distributed.

Statistical tests are commonly compared based on their power — the ability to correctly reject the null hypothesis. The Anderson–Darling test tends to be the most powerful in a wide range of scenarios (sample size, underlying distribution, differences in tails versus in the middle), closely followed by Cramer–von-Mises and then Kolmogorov–Smirnov (Stephens, 1984; Laio, 2004; Engmann and Cousineau, 2011). The differences in the tests' power can reduce and eventually disappear for sample sizes over 100–1000, depending on the underlying distribution (Mohd Razali and Bee Wah, 2011). It is worth noting, that the Kolmogorov–Smirnov test requires a significant increase in sample size to reach the same power as Cramer–von-Mises or Anderson–Darling. Both Cramer–von-Mises and Anderson–Darling tests are used in this study to compare the distributions of ERA5 and buoy data.

Additionally, the influence of the distance to shore and water depth at the location of the measurements is assessed. Both metrics were estimated using GIS and publicly available bathymetry data (GEBCO, 2022). These metrics were included to test the ERA5 model's ability to capture the impact of local coastal and bathymetry features on wave propagation and dissipation. It should be noted that due to the



Fig. 4. Buoy measurement timeline by region with the colourscale representing annual data availability.

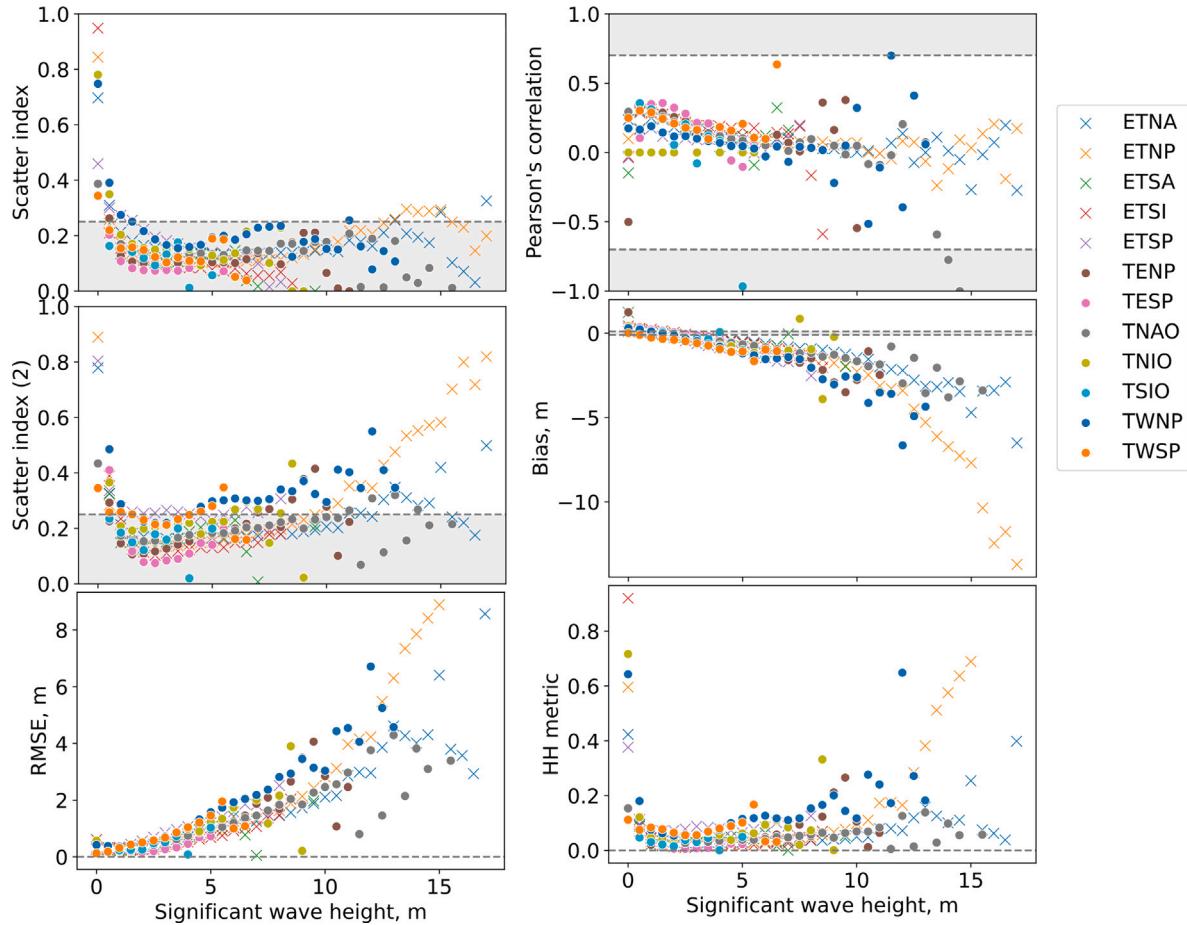


Fig. 5. Scatter index (1 and 2), RMSE, correlation, bias and HH metric calculated in 0.5 m/s H_s bins for each region (represented by different markers — see the legend). The shaded area shows ranges of acceptable values. The HH plot excludes four data points where $HH > 1$ caused by division by low H_s .

resolution of GEBCO data, the bathymetry around small islands may be not captured, which can impact the statistics related to the distance to shore and water depth.

3. Results

Application of the buoy selection criteria reduced the total dataset to 444 locations in 12 regions. Mean regional statistics are presented in

Table 3. Bias is close to or within acceptable values for all regions except TNIO and TWSP (Indian peninsula and eastern Australia). Results in [Table 3](#) show no particular relationship between the accuracy of the ERA5 dataset and the region. Higher values of H_s are expected to be more prevalent in limited areas, which may have statistics that are not representative of the regional or global results. The authors showed that in such regions, the statistics for the data where $H_s < 10$ m are similar to the results for the respective regions ([Appendix B](#)).

Table 3

Mean regional statistics of buoy measurements. Values in bold are outside acceptable limits defined in [Table 2](#). Regions are defined in [Fig. 3](#). The numbers for the Mediterranean region are included for comparison, and not considered in the mean across regions.

| Region | RMSE, m | Mean, m | Bias, m | NBias | CC | SI_1 | SI_2 | HH | Equiv. years |
|--------------------------------|---------|---------|--------------|--------|------|-------------|-------------|------|--------------|
| Extratropical North Atlantic | 0.46 | 1.86 | -0.06 | 0.034 | 0.91 | 0.29 | 0.25 | 0.08 | 7.6 |
| Extratropical North Pacific | 0.36 | 1.94 | -0.07 | -0.019 | 0.94 | 0.20 | 0.19 | 0.41 | 13.2 |
| Extratropical South Atlantic | 0.33 | 2.04 | -0.08 | -0.036 | 0.92 | 0.16 | 0.16 | 0.39 | 1.9 |
| Extratropical South Indian | 0.34 | 2.38 | 0.01 | 0.01 | 0.95 | 0.13 | 0.14 | 0.08 | 7.5 |
| Extratropical South Pacific | 0.47 | 1.65 | -0.09 | -0.055 | 0.83 | 0.25 | 0.29 | 0.1 | 29.1 |
| Tropical Eastern North Pacific | 0.28 | 2.11 | -0.07 | -0.032 | 0.91 | 0.12 | 0.13 | 0.68 | 14.0 |
| Tropical Eastern South Pacific | 0.23 | 2.18 | 0.06 | 0.034 | 0.90 | 0.10 | 0.11 | 0.03 | 5.9 |
| Tropical North Atlantic Ocean | 0.26 | 1.28 | -0.04 | -0.017 | 0.92 | 0.20 | 0.20 | 0.05 | 10.3 |
| Tropical North Indian Ocean | 0.49 | 1.43 | 0.02 | 0.067 | 0.82 | 0.34 | 0.34 | 0.11 | 0.89 |
| Tropical South Indian Ocean | 0.24 | 1.35 | 0.11 | 0.081 | 0.85 | 0.16 | 0.18 | 0.03 | 2.7 |
| Tropical Western North Pacific | 0.48 | 1.53 | 0.04 | 0.088 | 0.81 | 0.31 | 0.31 | 0.17 | 4.2 |
| Tropical Western South Pacific | 0.29 | 1.30 | -0.17 | 0.089 | 0.92 | 0.21 | 0.22 | 0.08 | 17.9 |
| Mean | 0.35 | 1.75 | -0.03 | 0.0005 | 0.89 | 0.205 | 0.21 | 0.18 | 9.6 |
| The Mediterranean | 0.33 | 1.15 | -0.10 | 0.087 | 0.94 | 0.275 | 0.26 | 0.18 | 5.5 |

The same statistics were calculated for the data binned by 0.5 m bins ([Fig. 5](#)). For location-bin pairs, where SI is higher than one, the values were omitted. There were two such points in the first bin caused by division by a mean close to zero.

Bias shows a growing underestimation of H_s by ERA5 at more extreme significant wave height in all regions. The absolute bias does not exceed 2 m when $H_s \leq 5$ m, however, the bias in each bin is still high. SI_2 is more sensitive to errors and exceeds the acceptable value in more bins and regions than SI_1 . The absolute bias and SI grow with growing H_s suggesting lower accuracy of ERA5 at extreme conditions. This trend in accuracy is also supported by the HH metric, which stays low in all regions up to 10 m significant wave height. The lack of significant variation between regional statistics confirms the aforementioned statement that the accuracy of ERA5 is unlikely to be governed by global scale variations.

[Fig. 6](#) shows that correlation increases slightly and SI decreases with increasing distance to shore, suggesting higher accuracy of ERA5 further offshore. Changes in water depth, on the other hand, do not create any particular trends for the ERA5 accuracy. The ERA5 accuracy does, however, depend on the time period, with the increased accuracy in recent years ([Fig. 7](#)). This pattern is further discussed in Section 4.1.

There is no particular difference in accuracy between extratropical and tropical regions ([Fig. 8](#)). The global average in this study compares well to the results of [Sifnioti et al. \(2020\)](#) and [Hisaki \(2020\)](#). Other similar studies ([Naseef and Kumar, 2020; Bruno et al., 2020; Wang and Wang, 2021; Shi et al., 2021](#)) did not provide the standard deviation metric to allow comparison on a Taylor diagram.

An Anderson–Darling test showed that the data from buoys and ERA5 data come from different distributions for 440 out of 444 locations at 5% significance level. A Cramer–von-Mises test did not detect differences in distributions in two additional buoys compared to the Anderson–Darling test. Five of the buoys that have similar distributions are located in the ETNP region and one in TNAO. The measurements' duration, depth and distance from shore of the buoys range between 0.5–24 years, 38–1650 m and 12–53 km respectively. All buoys with similar distributions between ERA5 and observations are relatively close to shore. However, the small number of such buoys (compared to the whole sample) does not allow us to conclude that the similarity in distributions is affected by the local climate.

Northern regions have lower ERA5 accuracy in winter months as suggested by larger absolute bias and RMSE ([Fig. 9](#)). The ERA5 accuracy in tropical and southern regions is more consistent throughout the seasons and overall is slightly higher than in the North. Analysis of individual swell-generating regions shows similar seasonal patterns ([Appendix A](#)) with some regions having high uncertainty in monthly results due to a smaller amount of available measurements.

Linear regression was fit to the full dataset ([Fig. 10](#)) and to subset of buoys in each swell-generating region ([Appendix C](#)). A single calibration factor was calculated as a ratio between the measured and ERA5

H_s . This factor was calculated for all buoys – 1.035, for buoys closer than 55.7 km to shore (the median distance) – 1.042, and for buoys further than 55.7 km from shore – 1.027.

4. Discussion

4.1. ERA5 accuracy

The results are generally in agreement with the previous findings and confirm the high accuracy of the ERA5 dataset on a much wider temporal and spatial scale, than the previous studies ([Table 4](#)).

Similar to [Hisaki \(2020\)](#), [Wang and Wang \(2021\)](#) and [Shi et al. \(2021\)](#), this study showed that ERA5 underestimates H_s in higher ranges. The inability to accurately model extreme conditions is a known issue in climate and weather prediction ([Timmermans et al., 2020](#), e.g.). The low frequency of extreme events causes a lack of validation data for model prediction, which stalls their improvement. Due to the limited accuracy of ERA5 data in extreme conditions, all applications that rely on such data should use it with care. Examples of such applications include the design of vessels, coastal and offshore structures, as well as climate model validation.

Similarly to [Wang and Wang \(2021\)](#), year-by-year analysis of the data showed that the accuracy of ERA5 wave height data increases in recent years. This can be explained by the increased frequency and distribution of satellite measurements in the last decade. Satellite measurements are assimilated with the numerical weather model to produce the ERA5 dataset.

[Law-Chune et al. \(2021\)](#) performed a comparison of ERA5 and WAVERYS with a buoy network similar to this study, concluding that WAVERYS outperforms ERA5 in most cases. Their study mostly focused on the accuracy of WAVERYS and did not provide the same level of analysis of ERA5 as performed in our study.

Other factors which could impact the accuracy of the ECMWF model for ERA5 data include regional weather and climate phenomena, water depth and distance to shore. The increasing accuracy of ERA5 with the increasing distance from shore could be caused by the inability of the ERA5 wave model to capture the impacts of local land features on wave dissipation. This could be linked to the relatively coarse resolution of the wave model. There are some seasonal patterns of ERA5 accuracy in several regions, but annual values vary minimally with the region. Seasonality of the ERA5 accuracy in the extratropical northern regions could indicate a lower ability of the ECMWF model to represent extratropical storms, which in the Northern Hemisphere occur predominantly in winter ([Alves, 2006; Befort et al., 2019](#)). It is unclear whether similar results are absent in the Southern Hemisphere due to the differences in physical relationships or the significantly shorter and less widespread measurement campaigns. An additional study would be required to attribute these regional variations in ERA5 accuracy to storm activity.

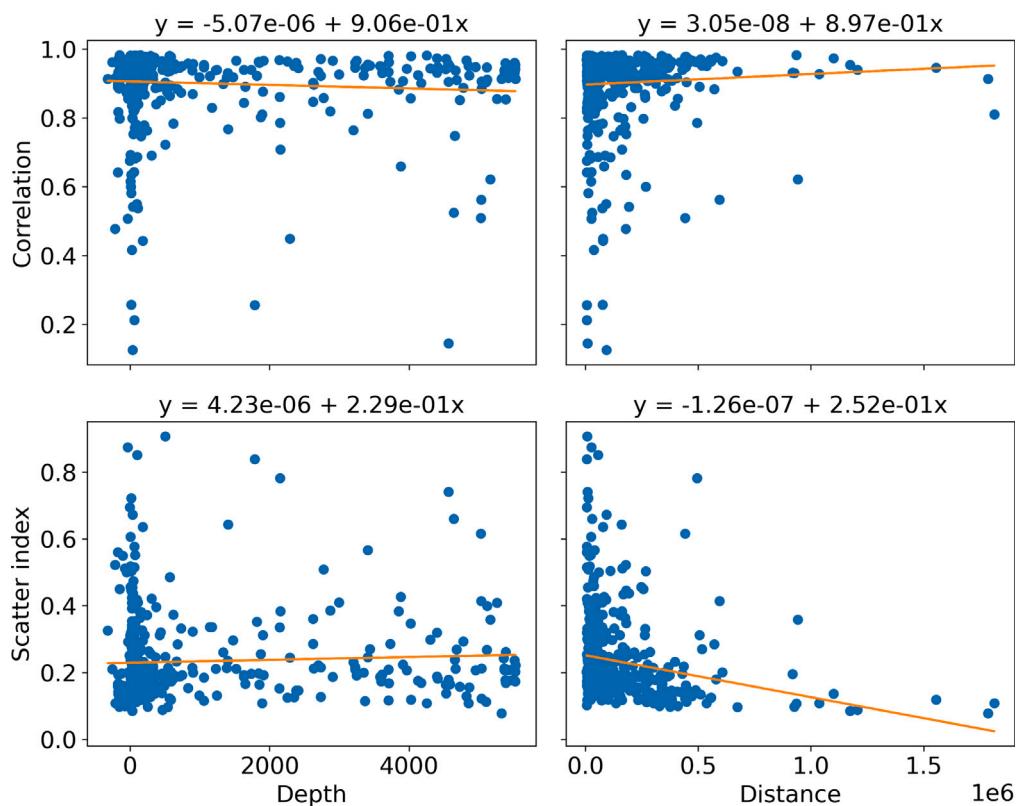


Fig. 6. Correlation (top row) and scatter index (bottom row) as functions of water depth (left) and distance from shore (right). The equation on top of each plot corresponds to the best fit linear regression.

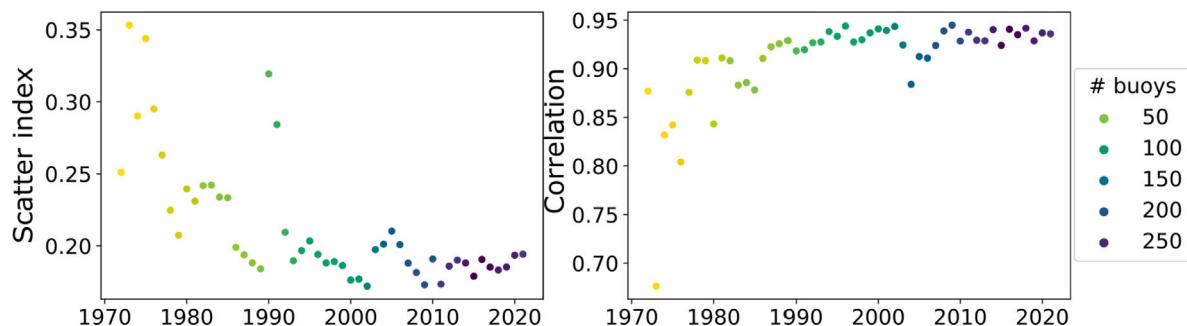


Fig. 7. Scatter index (left) and Pearson's correlation (right) between H_s measured by buoys and H_s in ERA5 dataset.

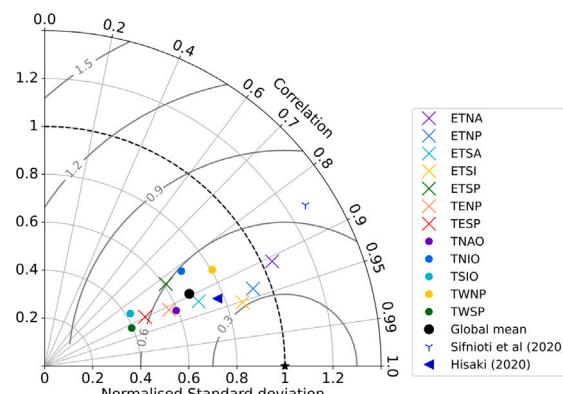


Fig. 8. Taylor diagram of regional ERA5 accuracy. The start shows the location of an idealised dataset with the correlation and the normalised standard deviation of one and RMSE of zero. The closer the data point to the star sign, the higher the accuracy of ERA5 for that data point.

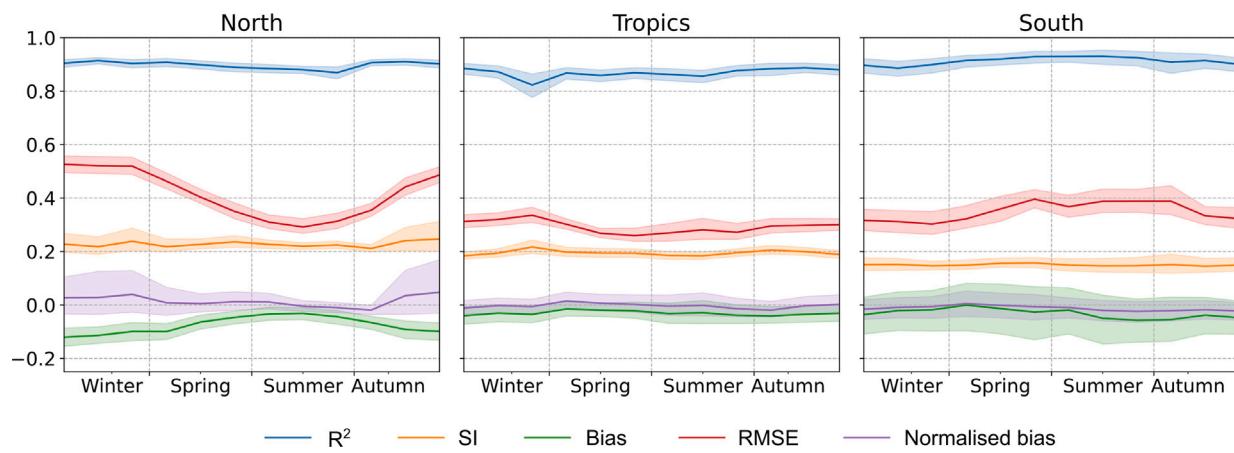


Fig. 9. Monthly trends in correlation, bias, normalised bias, RMSE and SI between ERA5 and buoy measurements. The shaded area represents 95% confidence interval. North includes ETNP and ETNA regions from Fig. 3, South — ETSP, ETSA and ETSI, the tropics — the rest of the regions.

Table 4

Comparison of ERA5 performance analysis in this study to previous research.

| Source | Region | No. buoys | Period | R ² | RMSE, m | Bias, m | SI 1 | SI 2 |
|-------------------------|-----------------------------|-----------|------------------------------|----------------|---------|-------------------|-------|------|
| Naseef and Kumar (2020) | Bay of Bengal, Arabian Sea | 2 | 2012, 2015 | 0.96 | 0.15 | -0.005 | | |
| Sifnioti et al. (2020) | North Sea | 2 | 2008–2017, 2010–2017 | 0.85 | 0.3 | -0.22 | | |
| Hisaki (2020) | Japan | 10 | 2014–2018 | 0.899 | | | 0.236 | |
| Bruno et al. (2020) | Arabian Sea | 1 | 08/2013–12/2013 | 0.9 | 0.29 | 0.03 ^b | | |
| Wang and Wang (2021) | US | 103 | 1979–2019 | 0.961 | 0.325 | -0.058 | 0.185 | |
| Shi et al. (2021) | Yellow Sea, South China Sea | 6 | 2011–2013, 2018 ^a | 0.89 | 0.23 | -0.04 | | 0.35 |
| Zhai et al. (2023) | South China Sea | 15 | 2020–2021 | 0.91 | 0.21 | 0.01 | | |
| Law-Chune et al. (2021) | Global | 200 | 1993–2019 | 0.97 | | -0.055 | 0.213 | |
| This study | Global | 454 | 1970–2021 | 0.953 | 0.39 | -0.05 | 0.205 | 0.21 |

^a Different years for different buoys, mainly 1 year per buoy.

^b Relative bias in %.

^c Exact number is not reported, we provide our estimate here.

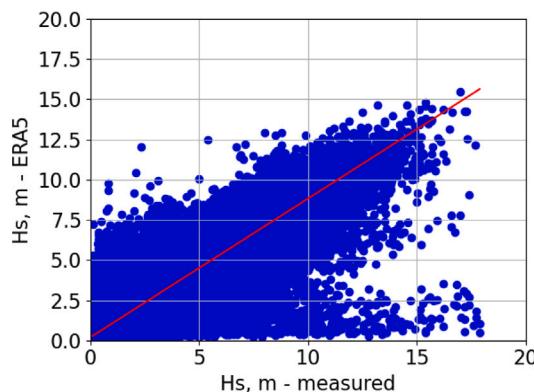


Fig. 10. Scatter plot showing the level of agreement between measured H_s and ERA5. The red line represents linear regression $y = 0.864 * x + 0.184$.

Uncertainty in both the ERA5 model and buoy measurements can affect the results of this comparison, however, these impacts could be hard to quantify. Sources of uncertainty in the ERA5 model include parametrisations of physical processes, inherited uncertainty from the input data (e.g. satellite altimetry, bathymetry), as well as limitations caused by relatively sparse spatial resolution. The effect of the latter would be especially important in coastal areas, where land features and changes in water depth drive localised wave behaviour. Instrument errors, differences in measurement techniques and data processing create uncertainty in buoy measurements.

Instrument errors, which are assumed to be between 0.01 m and 0.1 m, would be expected to have an impact at lower significant wave heights, where the relative error is higher. Even though the majority

of erroneous measurements were filtered during the processing in this study, the data can still have undetected spikes in measurements and other types of errors. The main concern is buoys with intermittent measurements since the reasons for prolonged gaps in measurement data are not documented. If any buoys sustained maintenance throughout the measurement period, there could be a change in calibration, change of measurement devices or other alterations that would introduce additional uncertainty to the whole time series.

4.2. Recommendations for future work

This study has not investigated the potential impacts of fetch and tides on the ERA5 performance. Fetch could be one of the indicators of ERA5's ability to capture the impacts of local coastal features on wave development and propagation. Tides can impact wave propagation in nearshore regions (Héquette et al., 2021; Lewis et al., 2019), which could affect how well ERA5 represents H_s in these areas. It is important to point out the recent release of ERA5 hindcast until 1940 (Hersbach et al., 2023). It would be difficult to assess the accuracy of the wave hindcast in this dataset due to the lack of instrumental measurements between 1940 and 1970.

The buoy measurements were not calibrated or analysed for changes in measurement equipment due to limited resources and information about each measuring device and the large amount of buoy data. The same sensors have been shown to both overestimate and underestimate the significant wave height by different authors (Gemmrich et al., 2011). Analysis of the potential impacts of equipment change at each buoy location was out of this study's scope. Future analysis of these impacts could improve the understanding of the differences shown between ERA5 and buoy measurements.

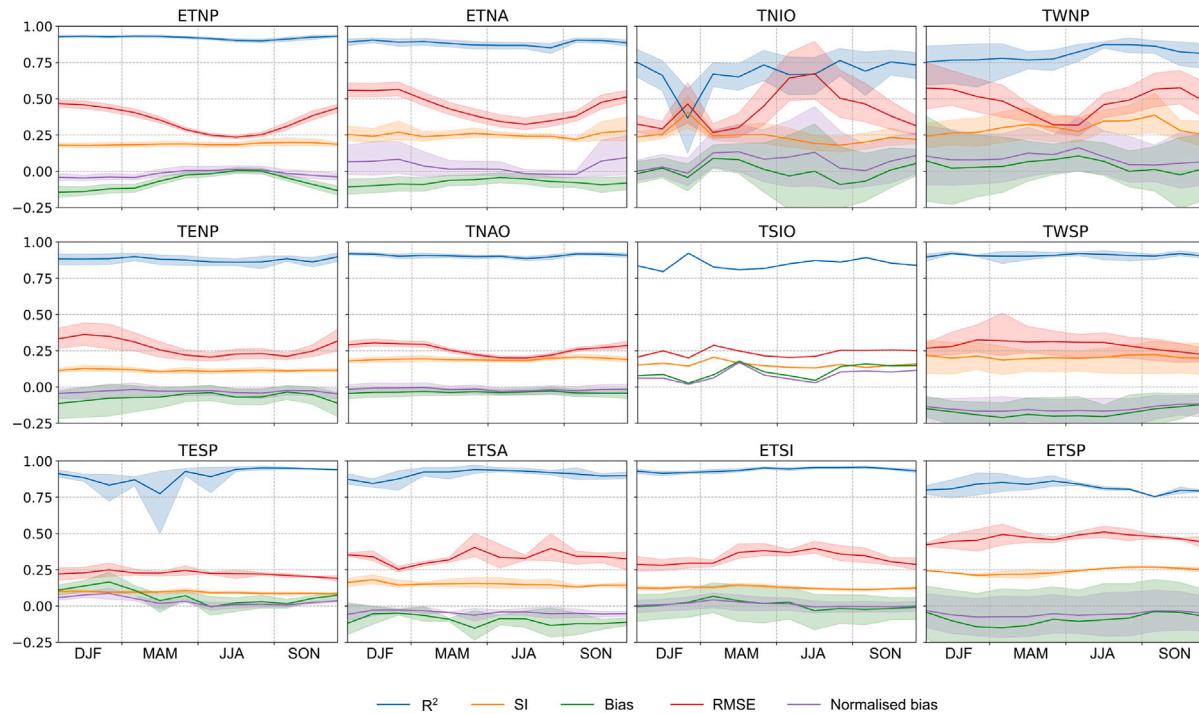


Fig. A.1. Monthly trends in correlation, bias, normalised bias, RMSE and SI between ERA5 and buoy measurements. The shaded area represents 95% confidence interval.

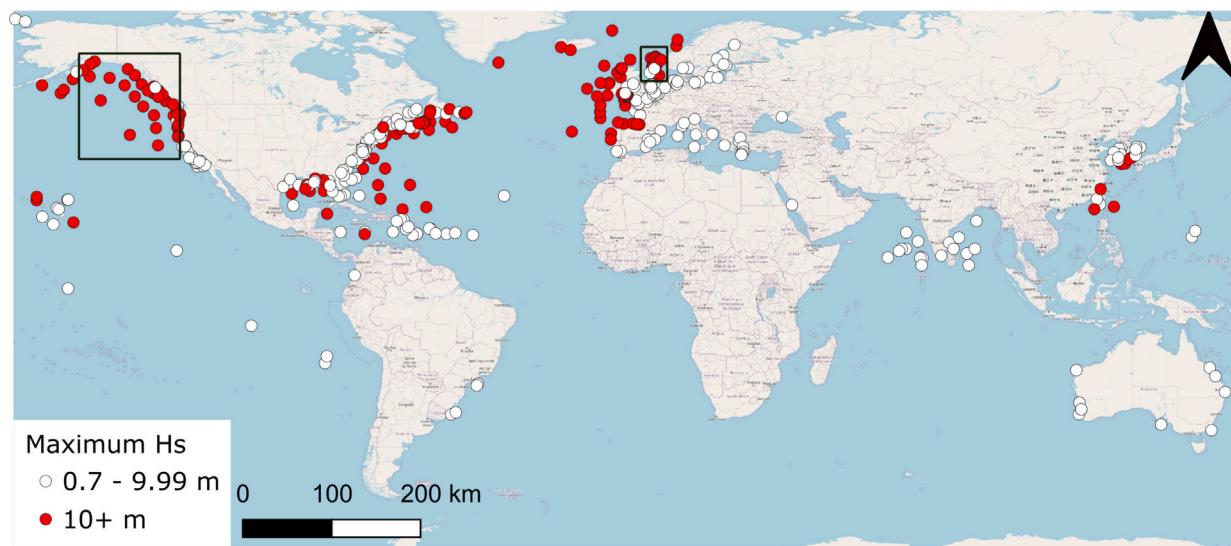


Fig. B.1. A map of buoys with the maximum recorded H_s below and above 10 m. The black polygons show regions chosen for validation of the comparison statistics, such as RMSE, correlation, HH index and bias.

The accuracy of ERA5 could be impacted by storms and cyclones, as the authors hypothesised based on the seasonal variations in extratropical Northern regions. A suggestion for future studies is to identify storm occurrences or split the data into weather types and assess the dataset within these categories. Some relationships between ERA5 accuracy and larger weather and climate patterns could be noticed from the analysis of monthly statistics on a local scale. For example, Li et al.

(2022) showed that RMSE between ERA5 and buoy observations is larger during the summer monsoon season in South East Asia.

In regions with a relatively high density of man-made structures, such as oil and gas rigs and offshore wind farms, it would be interesting to study the impact of such structures on the wave climate. In Chinese and North European waters, the density of wind farms is growing, and the atmospheric wake effects from them persist hundreds of kilometres

Table B.1
Comparison statistics for subsets with recorded $H_s > 10$ m.

| Region | Bias | RMSE | Buoy mean | NRMSE | SI | CC | HH |
|------------------------------------|--------|-------|-----------|-------|-------|--------|-------|
| Global | -0.030 | 0.350 | 1.750 | 0.200 | 0.205 | 0.890 | 0.180 |
| Extratropical North Atlantic | -0.060 | 0.460 | 1.860 | 0.247 | 0.290 | 0.910 | 0.080 |
| Northern North Sea, $H_s < 10$ m | -0.051 | 0.410 | 2.471 | 0.166 | 0.163 | 0.962 | 0.021 |
| Northern North Sea, $H_s < 8$ m | -0.050 | 0.403 | 2.447 | 0.165 | 0.164 | 0.960 | 0.021 |
| Northern North Sea, $H_s > 10$ m | -1.012 | 1.425 | 10.823 | 0.132 | 0.093 | 0.340 | 0.019 |
| Northern North Sea, $H_s > 8$ m | -0.387 | 1.042 | 8.944 | 0.117 | 0.108 | 0.563 | 0.014 |
| Extratropical North Pacific | -0.070 | 0.360 | 1.940 | 0.186 | 0.200 | 0.940 | 0.410 |
| Northeastern Pacific, $H_s < 10$ m | -0.090 | 0.413 | 2.397 | 0.172 | 0.168 | 0.958 | 0.024 |
| Northeastern Pacific, $H_s < 8$ m | -0.087 | 0.402 | 2.382 | 0.169 | 0.165 | 0.960 | 0.024 |
| Northeastern Pacific, $H_s > 10$ m | -4.034 | 5.546 | 11.470 | 0.484 | 0.332 | -0.373 | 0.369 |
| Northeastern Pacific, $H_s > 8$ m | -2.022 | 2.872 | 9.134 | 0.314 | 0.223 | -0.010 | 0.127 |

downstream of wind farms (Dong et al., 2022; Schneemann et al., 2020). The ECMWF model is unlikely to include such impacts as this effect is recent and less studied. A valid hypothesis would be to suggest that lower wind speeds in the wake of turbines and the subsea parts of turbines and oil platforms slow down wave generation and propagation. This would mean that in these regions, ERA5 would be likely to overestimate significant wave height in recent and later years when the density of wind farms is high.

5. Conclusions

This study assessed the accuracy of the ERA5 significant wave height with respect to buoy measurements. For the first time, the impact of such factors as wave generation regions, water depth and distance to shore were assessed in a global scale comparison of ERA5 with buoy measurements. Metrics used for this comparison include RMSE, correlation, bias and scatter index. This comparison is useful to the research community and to the metocean and maritime industries.

We showed that ERA5 is increasingly less accurate for significant wave heights above 10 m. ERA5 accuracy increases with increasing distance to shore, as well as in recent years (post-1990). On the other hand, water depth and wave generation region do not have significant effects on the reanalysis accuracy, except a trend of increased accuracy in summer in the Northern Hemisphere. The distribution analysis suggested the distributions of measured and modelled wave height are different for almost all studied locations. We also compared two definitions of the scatter index used by the research community, and concluded that both definitions produce similar results for our dataset.

These weaknesses of the ERA5 wave data are important to consider when using these data in vessel and offshore structure design, planning of marine operations, assessing wave energy potential and validating other wave models.

CRediT authorship contribution statement

Victoria Bessonova: Writing – review & editing, Writing – original draft, Visualization, Software, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Evdokia Tapogloou:** Writing – review & editing, Supervision, Methodology, Conceptualization. **Robert Dorrell:** Writing – review & editing, Supervision, Funding acquisition, Conceptualization. **Nina Dethlefs:** Writing – review & editing, Supervision. **Katharine York:** Writing – review & editing, Supervision, Resources.

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 work was supported by Engineering and Physical Sciences Research Council, United Kingdom and the Offshore Renewable Energy Catapult [grant number EP/S023763/1].

Appendix A. Monthly regional statistics

Fig. A.1 shows monthly statistics by region. In most regions, the accuracy of ERA5 shows very little seasonal variation. The North Pacific and the extratropical North Atlantic present lower RMSE and absolute bias in summer compared to winter, while correlation and SI stay stable. Lower confidence in the monthly statistics of the TNIO region is likely due to the insufficient average duration of the measurements (less than 1 year).

Appendix B. Comparison of statistics in areas with higher H_s

The validation of the comparison statistics was performed for two areas with prevalent high values of significant wave height. An area in the Northern North Sea and an area in the Northeastern Pacific were chosen for this validation (**Fig. B.1**), with 21 and 42 buoys with a maximum recorded $H_s > 10$ m respectively.

The data in the chosen areas were split into higher and lower H_s based on two thresholds, 8 m and 10 m. Both thresholds show that the statistics for lower H_s in these areas are similar to the statistics of the respective swell-generating regions (**Table B.1**).

Appendix C. Scatter plots of measured versus modelled H_s

Linear regression was fit to find relationship between the modelled ERA5 and the measured H_s in each swell-generation region (**Fig. C.1**). The result supports the previous finding that ERA5 underestimates H_s . The underestimation is present in all regions, including regions with relatively lower H_s , such as Tropical Eastern and Western South Pacific, and Tropical South Indian Ocean.

Data availability

The raw buoy data were collected from [Cefas \(2021\)](#), [AODN \(2021\)](#), [Copernicus Marine Service \(2021\)](#). ERA5 data were downloaded from the Copernicus Climate Change Service's Climate Data Store ([Hersbach et al., 2020](#)). The final dataset with buoy measurements, corresponding ERA5 data and the buoy registry with calculated parameters are available in the public repository ([Bessonova, 2024](#)).

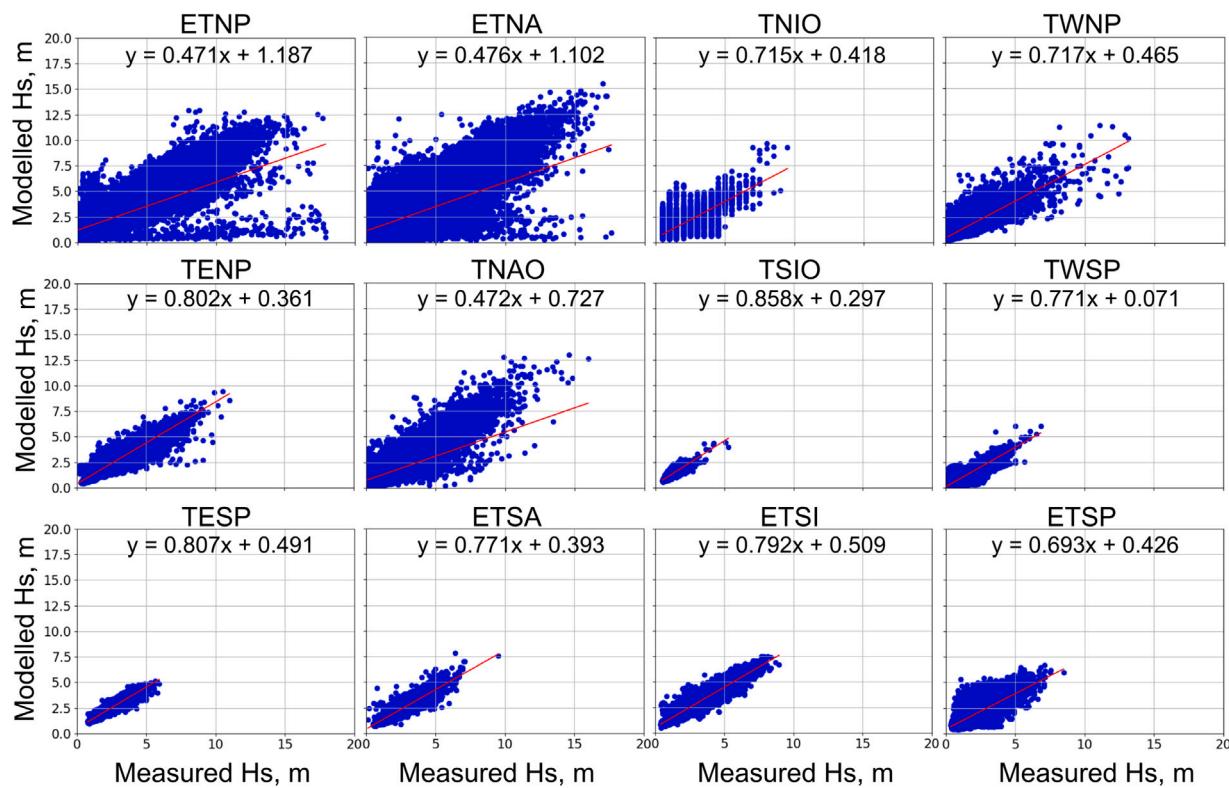


Fig. C.1. Scatter plots of measured versus modelled H_s for each swell-generating region. The red line and equation on each subplot corresponds to the best linear regression fit.

References

- Alfonso, M.d., Manzano, F., Gallardo, A., 2022. Quality Information Document for Global Ocean - Delayed Mode Wave Product. Technical Repor, Copernicus Marine Service, URL: <http://dx.doi.org/10.1214/aoms/117728726>.
- Alves, J.H.G., 2006. Numerical modeling of ocean swell contributions to the global wind-wave climate. *Ocean. Model.* 11 (1–2), 98–122. <http://dx.doi.org/10.1016/j.ocemod.2004.11.007>.
- Amores, A., Marcos, M., Carrión, D.S., Gómez-Pujol, L., 2020. Coastal impacts of storm gloria (January 2020) over the north-western mediterranean. *Nat. Hazards Earth Syst. Sci.* 20 (7), 1955–1968. <http://dx.doi.org/10.5194/nhess-20-1955-2020>.
- Anderson, T.W., Darling, D.A., 1952. Asymptotic Theory of Certain "Goodness of Fit" Criteria Based on Stochastic Processes. *Ann. Math. Stat.* 23 (2), 193–212. <http://dx.doi.org/10.1214/aoms/117729437>.
- ODN, 2021. Waverider buoys observations - Australia,delayed. In: Australian Ocean Data Network.
- Befort, D.J., Wild, S., Knight, J.R., Lockwood, J.F., Thornton, H.E., Hermanson, L., Bett, P.E., Weisheimer, A., Leckebusch, G.C., 2019. Seasonal forecast skill for extratropical cyclones and windstorms. *Q. J. R. Meteorol. Soc.* 145 (718), 92–104. <http://dx.doi.org/10.1002/qj.3406>.
- Berrisford, P., Dee, D., Poli, P., Brugge, R., Fielding, K., Fuentes, M., Källberg, P., Kobayashi, S., Uppala, S., Simmons, A., 2011. The ERA-40 archive. Version 2.0.
- Bessonova, V., 2024. Dataset to support the publication "global evaluation of wave data reanalysis: comparison of the ERA5 dataset to buoy observations". <http://dx.doi.org/10.5281/zendodo.14544121>.
- Bruno, M.F., Molfetta, M.G., Totaro, V., Mossa, M., 2020. Performance Assessment of ERA5 Wave Data in a Swell Dominated Region. *J. Mar. Sci. Eng.* 2020 8 (3), 214. <http://dx.doi.org/10.3390/JMSE8030214>, URL: <https://www.mdpi.com/2077-1312/8/3/214>.
- Bryant, M.A., Hesser, T.J., Jensen, R.E., 2016. Evaluation Statistics Computed for the Wave Information Studies (WIS). US Army Corps of Engineers, Vicksburg, URL: <https://apps.dtic.mil/sti/citations/AD1013235>.
- Caires, S., Sterl, A., 2005. 100-Year Return Value Estimates for Ocean Wind Speed and Significant Wave Height from the ERA-40 Data. *J. Clim.* 18 (7), 1032–1048. <http://dx.doi.org/10.1175/JCLI-3312.1>, URL: <https://journals.ametsoc.org/view/journals/clim/18/7/jcli-3312.1.xml>.
- Campsins, J., Genoves, A., Picornell, M., Jansa, A., 2011. Climatology of mediterranean cyclones using the ERA-40 dataset. *Int. J. Climatol.* 31, <http://dx.doi.org/10.1002/joc.2183>.
- Cefas, 2021. WaveNet interactive map. URL: <https://wavenet.cefas.co.uk/Map>.
- Chen, H., 2022. A comprehensive statistical analysis for residuals of wind speed and direction from numerical weather prediction for wind energy. *Energy Rep.* 8, 618–626. <http://dx.doi.org/10.1016/J.EGMR.2022.07.080>.
- Chernoff, H., Lehmann, E.L., 1954. The Use of Maximum Likelihood Estimates in χ^2 Tests for Goodness of Fit. *Ann. Math. Stat.* 25 (3), 579–586. <http://dx.doi.org/10.1214/aoms/117728726>.
- CMEMS, 2023. Global ocean waves reanalysis. <http://dx.doi.org/10.48670/moi-00022>.
- Compo, G.P., Whitaker, J.S., Sardeshmukh, P.D., Matsui, N., Allan, R.J., Yin, X., Gleason, B.E., Vose, R.S., Rutledge, G., Bessemoulin, P., Broennimann, S., Brunet, M., Crouthamel, R.I., Grant, A.N., Groisman, P.Y., Jones, P.D., Kruk, M.C., Kruger, A.C., Marshall, G.J., Maugeri, M., Mok, H.Y., Nordli, O., Ross, T.F., Trigo, R.M., Wang, X.L., Woodruff, S.D., Worley, S.J., 2011. The Twentieth Century Reanalysis Project. *Q. J. R. Meteorol. Soc.* 137 (654), 1–28. <http://dx.doi.org/10.1002/QJ.776>, URL: <https://onlinelibrary.wiley.com/doi/full/10.1002/qj.776>.
- Copernicus Marine Service, 2021. Global Ocean - Delayed Mode Wave product. <http://dx.doi.org/10.48670/moi-00042>.
- Cramer, H., 1928. On the Composition of Elementary Errors. Vol. 11, Skand Aktuarietidskr.
- Dong, G., Li, Z., Qin, J., Yang, X., 2022. How far the wake of a wind farm can persist for? *Theor. Appl. Mech. Lett.* 12, <http://dx.doi.org/10.1016/J.TAML.2021.100314>.
- Dookie, I., Rocke, S., Singh, A., Ramlal, C.J., 2018. Evaluating wind speed probability distribution models with a novel goodness of fit metric: a Trinidad and Tobago case study. *Int. J. Energy Environ. Eng.* 9 (3), 323–339. <http://dx.doi.org/10.1007/S40095-018-0271-Y/FIGURES/20>, URL: <https://link.springer.com/article/10.1007/S40095-018-0271-y>.
- Dowdy, A.J., Mills, G.A., Timbal, B., Wang, Y., 2014. Fewer large waves projected for eastern Australia due to decreasing storminess. *NATURE CLIMATE CHANGE* | 4, <http://dx.doi.org/10.1038/NCLIMATE2142>, URL: www.nature.com/natureclimatechange.
- Engmann, S., Cousineau, D., 2011. Comparing distributions: the two-sample Anderson-Darling tes as an alternative to the Kolmogorov-Smirnoff test. *J. Appl. Quant. Methods* 6 (3).
- Ferreira, J.A., Guedes Soares, C., 2000. Modelling distributions of significant wave height. *Coast. Eng.* 40 (4), 361–374. [http://dx.doi.org/10.1016/S0378-3839\(00\)00018-1](http://dx.doi.org/10.1016/S0378-3839(00)00018-1).
- Flaounas, E., Raveh-Rubin, S., Wernli, H., Drobinski, P., Bastin, S., 2015. The dynamical structure of intense mediterranean cyclones. *Clim. Dyn.* 44, <http://dx.doi.org/10.1007/s00382-014-2330-2>.
- GEBCO, 2022. Gridded Bathymetry Data. URL: https://www.gebco.net/data_and_products/gridded_bathymetry_data/.
- Gemmrich, J., Thomas, B., Bouchard, R., 2011. Observational changes and trends in northeast Pacific wave records. *Geophys. Res. Lett.* 38 (22), n/a-n/a. <http://dx.doi.org/10.1029/2011GL049518>.

- Hanafin, J.A., Quilfen, Y., Ardhuin, F., Sienkiewicz, J., Queffeulou, P., Obrebski, M., Chapron, B., Reul, N., Collard, F., Corman, D., de Azevedo, E.B., Vandemark, D., Stutzmann, E., 2012. Phenomenal sea states and swell from a north atlantic storm in february 2011: A comprehensive analysis. *Bull. Am. Meteorol. Soc.* 93 (12), 1825–1832. <http://dx.doi.org/10.1175/BAMS-D-11-00128.1>.
- Hanna, S.R., Heinold, D.W., 1985. *Development and Application of a Simple Method for Evaluating Air Quality Models*. Vol. 4409, American Petroleum Institute.
- Hanson, J.L., Tracy, B.A., Tolman, H.L., Scott, R.D., 2009. Pacific Hindcast Performance of Three Numerical Wave Models. *J. Atmos. Ocean. Technol.* 26 (8), 1614–1633. <http://dx.doi.org/10.1175/2009JTECHO650.1>.
- Héquette, A., Cartier, A., Schmitt, F.G., 2021. The effects of tidal translation on wave and current dynamics on barred macrotidal beach, northern France. *J. Mar. Sci. Eng.* 9 (8), <http://dx.doi.org/10.3390/jmse9080909>.
- Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., Rozum, I., Schepers, D., Simmons, A., Soci, C., Dee, D., Thépaut, J.-N., 2023. ERA5 hourly data on single levels from 1940 to present. <http://dx.doi.org/10.24381/cds.adbb2d47>.
- 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., Abellán, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., 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., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., Thépaut, J.N., 2020. The ERA5 global reanalysis. *Q. J. R. Meteorol. Soc.* 146 (730), 1999–2049. <http://dx.doi.org/10.1002/qj.3803>.
- Hisaki, Y., 2020. Intercomparison of assimilated coastal wave data in the Northwestern Pacific Area. *J. Mar. Sci. Eng.* 8 (8), <http://dx.doi.org/10.3390/JMSE8080579>.
- Izaguirre, C., Losada, I.J., Camus, P., Vigh, J.L., Stenek, V., 2020. Climate change risk to global port operations. *Nat. Clim. Chang.* 2020 11: 1 11 (1), 14–20. <http://dx.doi.org/10.1038/s41558-020-00937-z>, URL: <https://www.nature.com/articles/s41558-020-00937-z>.
- Jensen, R.E., Swail, V.R., Bouchard, R.H., Riley, R., Hesser, T.J., Blasckie, M., MacIsaac, C., 2015. Field laboratory for ocean sea state investigation and experimentation: FLOSSIE intra-measurement evaluation of 6N wave buoy systems. URL: <https://api.semanticscholar.org/CorpusID:221855370>.
- Kolmogoroff, A., 1941. Confidence limits for an unknown distribution function. *Ann. Math. Stat.* 12 (4), 461–463. <http://dx.doi.org/10.1214/AOMS/1177731684>.
- Källberg, P., Simmons, A., Uppala, S., Fuentes, M., 2004. The ERA-40 archive. [revised october 2007]. (17), p. 31,
- Laio, F., 2004. Cramer-von Mises and Anderson-Darling goodness of fit tests for extreme value distributions with unknown parameters. *Water Resour. Res.* 40 (9), <http://dx.doi.org/10.1029/2004WR003204>.
- Law-Chune, S., Aouf, L., Dalphinet, A., Levier, B., Drillet, Y., Drevillon, M., 2021. WAVERYS: a CMEMS global wave reanalysis during the altimetry period. *Ocean. Dyn.* 71 (3), 357–378. <http://dx.doi.org/10.1007/S10236-020-01433-W/FIGURES/17>.
- Lawrence, J., Holmes, B., Bryden, I., Magagna, D., Yago, P.U., Eve, T.-E., Rousset, J.-M., Smith, H., Paul, M., Aau, L.M., Cândido, J., 2012. Work Package 2: Standards and Best Practice D2.1 Wave Instrumentation Database. Technical Report, MARINET, URL: https://tethys.pnnl.gov/sites/default/files/publications/D2_01_Wave_Instrumentation_Database.pdf.
- Lewis, M.J., Palmer, T., Hashemi, R., Robins, P., Saulter, A., Brown, J., Lewis, H., Neill, S., 2019. Wave-tide interaction modulates nearshore wave height. *Ocean. Dyn.* 69, <http://dx.doi.org/10.1007/s10236-018-01245-z>.
- Li, B., Chen, W., Li, J., Liu, J., Shi, P., Xing, H., 2022. Wave energy assessment based on reanalysis data calibrated by buoy observations in the southern South China Sea. *Energy Rep.* 8, 5067–5079. <http://dx.doi.org/10.1016/J.EGYR.2022.03.177>.
- Liléo, S., Petrik, O., 2011. Investigation on the use of NCEP/NCAR, MERRA and NCEP/CFSR reanalysis data in wind resource analysis. In: European Wind Energy Conference and Exhibition 2011, EWEC 2011. (March 2011), pp. 181–185.
- Mentaschi, L., Besio, G., Cassola, F., Mazzino, A., 2013. Problems in RMSE-based wave model validations. *Ocean. Model.* 72, 53–58. <http://dx.doi.org/10.1016/J.OCEMOD.2013.08.003>.
- von Mises, R., 1931. *Wahrscheinlichkeitsrechnung und Ihre Anwendung*. Stat. Und Theor. Phys..
- Mohd Razali, N., Bee Wah, Y., 2011. Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests. *J. Stat. Model. Anal.* 2 (1), 13–14.
- Naseef, T., Kumar, V.S., 2020. Climatology and trends of the Indian Ocean surface waves based on 39-year long ERA5 reanalysis data. *Int. J. Climatol.* 40 (2), 979–1006. <http://dx.doi.org/10.1002/JOC.6251>.
- NOAA, 2021. National Data Buoy Center. URL: <https://www.ndbc.noaa.gov/>.
- O'Connor, M., Lewis, T., Dalton, G., 2013. Weather window analysis of Irish west coast wave data with relevance to operations & maintenance of marine renewables. *Renew. Energy* 52, 57–66. <http://dx.doi.org/10.1016/j.renene.2012.10.021>.
- Rad, A.M., Ghahraman, B., Khalili, D., Ghahremani, Z., Ardakani, S.A., 2017. Integrated meteorological and hydrological drought model: A management tool for proactive water resources planning of semi-arid regions. *Adv. Water Resour.* 107, 336–353. <http://dx.doi.org/10.1016/J.JADWATRES.2017.07.007>.
- Ribal, A., Young, I.R., 2019. 33 years of globally calibrated wave height and wind speed data based on altimeter observations. *Sci. Data* 2019 6: 1 6 (1), 1–15. <http://dx.doi.org/10.1038/s41597-019-0083-9>, URL: <https://www.nature.com/articles/s41597-019-0083-9>.
- Schneemann, J., Rott, A., Dörenkämper, M., Steinfeld, G., Kühn, M., 2020. Cluster wakes impact on a far-distant offshore wind farm's power. *Wind. Energy Sci.* 5, 29–49. <http://dx.doi.org/10.5194/WES-5-29-2020>.
- Semedo, A., Dobrynin, M., Lemos, G., Behrens, A., Staneva, J., de Vries, H., Sterl, A., Bidlot, J.R., Miranda, P.M., Murawski, J., 2018. CMIP5-derived single-forcing, single-model, and single-scenario wind-wave climate ensemble: Configuration and performance evaluation. *J. Mar. Sci. Eng.* 6 (3), <http://dx.doi.org/10.3390/jmse6030090>.
- Sharmar, V., Markina, M., 2020. Validation of global wind wave hindcasts using ERA5, MERRA2, ERA-Interim and CFSRv2 reanalyses. *IOP Conf. Series: Earth Environ. Sci.* 606 (1), <http://dx.doi.org/10.1088/1755-1315/606/1/012056>.
- Shi, H., Cao, X., Li, Q., Li, D., Sun, J., You, Z., Sun, Q., 2021. Evaluating the Accuracy of ERA5 Wave Reanalysis in the Water Around China. *J. Ocean. Univ. China* 20 (1), <http://dx.doi.org/10.1007/S11802-021-4496-7>.
- Sifnioti, D.E., Dolphin, T.J., Vincent, C.E., 2020. Performance of Hindcast Wave Model Data used in UK Coastal Waters. *J. Coast. Res.* 95 (sp1), 1284–1290. <http://dx.doi.org/10.2112/SI95-248.1>.
- Smirnov, N., 1948. Table for Estimating the Goodness of Fit of Empirical Distributions. *Ann. Math. Stat.* 19 (2), 279–281. <http://dx.doi.org/10.1214/aoms/1177730256>.
- Stephens, M.A., 1984. Tests based on EDF statistics. In: D'Agostino, R.B., Stephens, M.A. (Eds.), *Goodness of Fit Techniques*. Marcel Dekker, New York, pp. 166–171.
- Storlazzi, C.D., Gingerich, S.B., Van Dongeren, A., Cheriton, O.M., Swarzenski, P.W., Quataert, E., Voss, C.I., Field, D.W., Annamalai, H., Piniak, G.A., McCall, R., 2018. Most atolls will be uninhabitable by the mid-21st century because of sea-level rise exacerbating wave-driven flooding. *Sci. Adv.* 4 (4), [http://dx.doi.org/10.1126/SCIADV.AAP9741/SUPPL_FILE/AAP9741_\(SM\).PDF](http://dx.doi.org/10.1126/SCIADV.AAP9741/SUPPL_FILE/AAP9741_(SM).PDF), URL: <https://www.science.org/doi/10.1126/sciadv.aap9741>.
- Timmermans, B., Shaw, A.G., Gommenginger, C., 2020. Reliability of extreme significant wave height estimation from satellite altimetry and in situ measurements in the coastal zone. *J. Mar. Sci. Eng.* 8 (12), 1–19. <http://dx.doi.org/10.3390/jmse8121039>.
- Toomey, T., Amores, A., Marcos, M., Orfila, A., 2022. Coastal sea levels and wind-waves in the mediterranean sea since 1950 from a high-resolution ocean reanalysis. *Front. Mar. Sci.* 9, <http://dx.doi.org/10.3389/fmars.2022.991504>.
- Wang, X.L., Feng, Y., Swail, V.R., 2015. Climate change signal and uncertainty in CMIP5-based projections of global ocean surface wave. *J. Geophys. Res.: Ocean.* 120 (5), 3859–3871. <http://dx.doi.org/10.1002/2015JC010699>.
- Wang, J., Wang, Y., 2021. Evaluation of the ERA5 Significant Wave Height against NDBC Buoy Data from 1979 to 2019. *Mar. Geod.* <http://dx.doi.org/10.1080/01490419.2021.2011502>.
- Wells, N., 2011. *The Atmosphere and Ocean: A Physical Introduction*, third ed.
- Willmott, C.J., Ackleson, S.G., Davis, R.E., Feddema, J.J., Klink, K.M., Legates, D.R., O'Donnell, J., Rowe, C.M., 1985. Statistics for the evaluation and comparison of models. *J. Geophys. Res.* 90 (C5), 8995. <http://dx.doi.org/10.1029/JC090iC05p08995>.
- Zhai, R., Huang, C., Yang, W., Tang, L., Zhang, W., 2023. Applicability evaluation of ERA5 wind and wave reanalysis data in the South China Sea. *J. Ocean. Limnol.* 41 (2), 495–517. <http://dx.doi.org/10.1007/S00343-022-2047-8-METRICS>.