IOP Conference Series: Earth and Environmental Science

**PAPER** • **OPEN ACCESS**

# Validation of global wind wave hindcasts using ERA5, MERRA2, ERA-Interim and CFSRv2 reanalyzes

To cite this article: V Sharmar and M Markina 2020 *IOP Conf. Ser.: Earth Environ. Sci.* **606** 012056

# Validation of global wind wave hindcasts using ERA5, MERRA2, ERA-Interim and CFSRv2 reanalyzes

**V Sharmar, M Markina**

Shirshov Institute of Oceanology, Moscow, Russia.

E-mail: `sharmvit@gmail.com`

December 2019

**Abstract.** Four global wind wave hindcasts based on ERA5, MERRA2, ERA-I and CFSR reanalyses and spectral wave model WAVEWATCH III for the period from 1980 to 2017 have been validated against satellite altimetry and NDBC buoys for 1 year. Hindcast based on newly released ECMWF reanalysis ERA5 demonstrated the best agreement with both satellite altimetery (with normalized bias being up to 5%, and RMSE up to 0.5 m) and buoy measurements, including values of upper percentiles. In general, all hindcasts show good correspondence with the observational data and thus can be used in further wind wave climate studies.

*Keywords*: global wave climate, validation, WAVEWATCHIII, ERA5, MERRA2

## 1. Introduction

In recent years wind wave hindcasts based on numerical wave modelling became one of the sources of information about ocean surface waves. Quality of hindcasts heavily depends on the atmospheric forcing usually provided by reanalyses. Various atmospheric datasets differ in time and space resolution and also have different assimilation algorithms and data assimilation input [8, 10]. In this way, they may not necessarily demonstrate the same signals in 10-m winds which reflects in the simulated wave fields. Here we used WaveWatch III spectral wave model forced by

ERA5, ERA-Interim, MERRA-2 and NCEP CFSR winds and developed wind wave hindcasts for the period from 1980 to 2017. The aim of this paper is to provide detailed validation of wave heights from hindcasts and wind speed from reanalyses against satellite altimetry and buoy data. We compare both winds and waves with observational data in order to identify which uncertainties in the simulated wave heights are attributed solely to the local atmospheric conditions. This paper is organized as follows. The configuration of numerical experiments with wave model and boundary conditions are described in Section 2. The results of validation against satellite altimetry and buoy data are presented in Section 3 and 4 respectively. The conclusions are summarized in Section 5.

## 2. Data and Methods

This study provides the validation of four long-term wind wave hindcasts forced by modern atmospheric reanalyses: ERA-5 [4, 5] and ERA-Interim [2] from the European Centre for Medium-range Weather Forecasts (ECMWF), MERRA2 [3] from the National Aeronautics and Space Administration (NASA) and NCEP CFSv2 [9] from the National Center of Environmental Prediction (NCEP). Information about spectral and vertical resolution as well as the resolution of the available output of these reanalyses is given in Table 1. The wave hindcasts were performed using the latest version 5.16 of WaveWatch III (WW3) spectral wave model with source term function ST4 [1] and with default settings of wind-wave growth parameter for the period from 1980-2017.

Table 1. Set-up of the hindcasts

| Reanalysis | Model Resolution | Data Assimilation | Spatial Resolution (lat x lon) | Time steps (sec) in WaveWatch III | | | |
|---|---|---|---|---|---|---|---|
| | | | | $\Delta t_g$ | $\Delta t_{xy}$ | $\Delta t_k$ | $\Delta t_s$ |
| ERA5 | T636 L137 | 4D-Var + EDA | 0.25° x 0.25° | 1800 | 225 | 900 | 60 |
| ERA-Interim | T255 L60 | 4D-Var | 0.7° x 0.7° | 1200 | 600 | 600 | 60 |
| NCEP CFSR (v2) | T382 L64 | 3D-Var SSI | 0.312° x 0.312° | 1800 | 300 | 900 | 30 |
| MERRA2 | 1/2 x 2/3 L72 | 3D-Var + IAU | 0.5° x 0.625° | 1200 | 600 | 600 | 60 |

Here validation of the obtained datasets has been performed using two data sources:

quality controlled Jason-1 altimeter data archive maintained at IFREMER [11, 12] and the historical buoy data archive from the National Data Buoy Center (NDBC; https://www.ndbc.noaa.gov/). The comparison is performed for both wind speed and wave heights for 1 year (2011). In order to smooth the altimetry measurements along the satellite track we use 6-point moving average for ERA5, 11-point moving average for MERRA-2, 7-point moving average for CFSR and 15-point moving average for ERA-Interim which approximately corresponds to spatial resolution of these datasets. Statistical estimates are based on the following error metrics: normalized bias (NBIAS), scatter index (SI) and root mean square error (RMSE) following [8]. We also use data that passed quality control and consider the areas where the distance from the coast is 5.8 km and more. The NDBC buoys provide quality controlled measurements of significant wave height (SWH) from a global network with the particular focus on US coastal areas. Since here we aim to assess overall errors on the global scale, we chose 35 buoy records containing both information about 10-m winds and significant wave heights located in shallow and deep water for winter season (January – March) of 2011.

## 3. Validation against altimeter data.

Comparisons between 10-m winds from reanalyses and Jason-1 altimeter data for 2011 are presented in Figure 1. Analysis of normalized biases (Fig. 1 a-d) between reanalyses and satellite shows that CFSR, ERA-Interim and ERA5 mostly overestimate wind speeds with the values of NBIAS ranging from 10-15 % except for the equatorial regions where NBIAS is negative and ranges from -5 to -15 % (equivalent to 1-1.5 $\mathrm{m\,s^{-1}}$). Moreover, moderately negative values of NBIAS are observed for ERA-Interim, ERA5 and MERRA2 in the subpolar North Atlantic and Southern Ocean where winds from reanalyses are underestimated compared to altimeters by 0.5-0.7 $\mathrm{m\,s^{-1}}$. MERRA2 is characterized by the lowest magnitudes of both positive and negative NBIAS ranging within ±5 %. The lowest scatter index is observed for MERRA2 and ERA5 reanalyses (up to 10 – 15 %) and the highest values (up to 30 %) are found for ERA-Interim and CFSR. The highest RSME in all reanalyses is found in the regions with the highest climatological wind speed in the mid- and high latitudes where error amounts up to 1.8 $\mathrm{m\,s^{-1}}$. In the trade wind zones RMSE is the lowest and ranges between 0.6 and 1.2 $\mathrm{m\,s^{-1}}$. Overall the highest RMSE is observed for ERA-Interim and ERA5 while the lowest is observed for MERRA2 and CFSR.

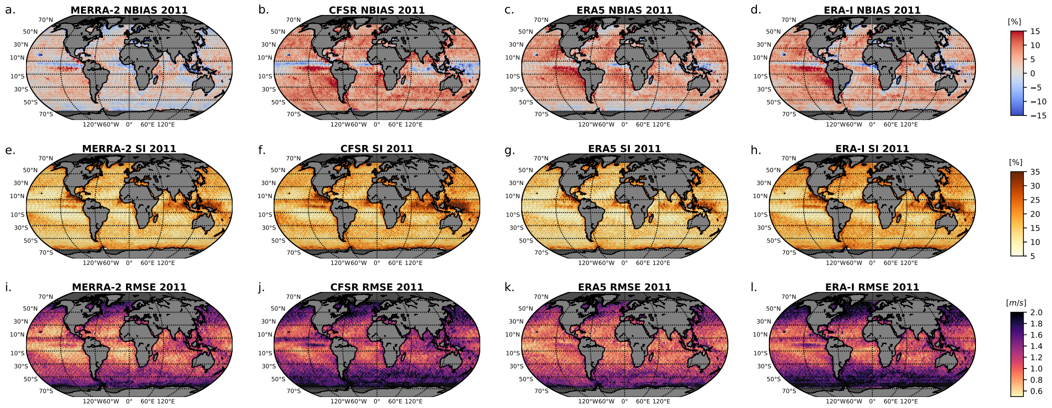

Figure 1. The normalized bias (NBIAS; a-d), scatter index (SI; e-h) and root-mean square error (RMSE; i-l) for 10-m wind speeds from MERRA-2 (a, e, i respectively), CFSR (b, f, j respectively), ERA5 (c, g, k respectively) and ERA-Interim reanalyses (d, h, i respectively) collocated with Jason-1 altimeter in 2011.

Evaluation metrics for the modelled SWH are presented in Figure 2. Interestingly, the patterns of NBIAS are different from those obtained for the winds (Figure 1). Three out of four hindcasts (MERRA2-WW3, ERA5-WW3 and ERAI-WW3) show overall underestimated SWH compared to Jason-1 altimeter with the highest NBIAS of -13 % for MERRA2-WW3. In addition, these hindcasts show moderately positive NBIAS in the Southern Ocean where bias for winds is close to zero or moderately negative. At the same time, CFSR-WW3 is characterized by the positive NBIAS values in the mid- and subpolar latitudes of the southern hemisphere with the values in the Southern Ocean being 15-25 % (equivalent to 0.5 - 1 meter). This implies qualitative consistency between the biases in wind speed and SWH for CFSR reanalysis and CFSR-WW3 hindcast respectively. Similar findings have been reported by [13] where they validated WW3-based wind wave hindcasts against ENVISAT satellite altimetry in 2011. All hindcasts show negative NBIAS compared to satellite data in the maritime continent in the Southeast Asia which is probably due to the complexity of the coastlines. The study of [14] reported that this region also exhibits the largest uncertainties among different wind wave projections due to intra-model discrepancies. Values of SI are relatively small for CFSR-WW3, ERA5-WW3 and ERAI-WW3 (5 to 15 %) indicating better agreement of these three hindcasts with satellite data compared to MERRA2-WW3 where SI exceed 25 % in the mid-latitudes. The RMSE is the lowest for ERA5-WW3 and ERAI-WW3 hindcasts with values being 0.3 m in the tropics and increasing to 0.4-0.5 m (ERA5-WW3) and 0.6-0.7 m (ERAI-WW3) in the Southern Ocean. CFSR-WW3 shows the largest RMSE of up to 0.9-1.0 m in the Southern Ocean. The highest overall RMSE of 0.3-0.5 m is identified in MERRA2-WW3 in the tropics and is approaching 1.0 m in the subpolar latitudes of the Northern Hemisphere.

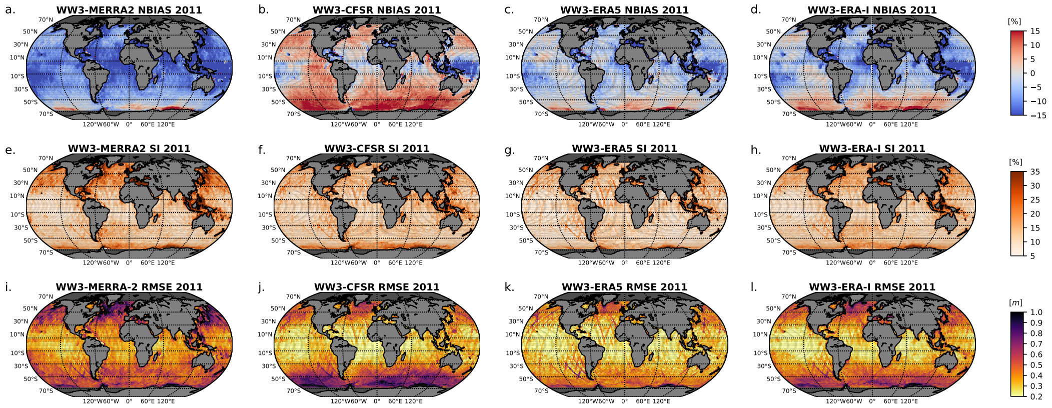

Figure 2. The normalized bias (NBIAS), scatter index (SI), root-mean square error (RMSE) for collocated significant wave heights from the MERRA2-WW3 (a, e, i respectively), CFSR-WW3 (b, f, j respectively), ERA5-WW3 (c, g, k respectively) and ERAI-WW3 hindcasts (d, h, i respectively) and Jason-1 altimeter for 2011.

## 4. Validation against buoy measurements.

The results of comparison between reanalyses/hindcasts and NDBC buoy measurements in the Atlantic and Pacific Oceans are presented on Taylor diagrams [6]. Figure 3 show multiple statistical metrics for time series of both 10-m wind speed (Fig.3 a - d) and SWH (Fig. 3 e - h). In general, MERRA2 (Fig. 3 a) and ERA5 (Fig. 3 c) demonstrate Pearson correlation coefficient higher than 0.8 (varying from 0.85 to 0.95 and from 0.9 to 0.97, respectively) and CRMSE $< 0.5$ $\mathrm{m\,s^{-1}}$. Normalized standard deviation for both MERRA2, ERA5 and ERAi (Fig. 3

c) reanalyses is close to 1. The exception is the buoy №41025 (North-West boundary of the Sargasso Sea), where correlation coefficient is less or equal 0.8 and CRMSE is high probably due to the shallow water limit. Figure 3d shows that CFSR seems to be overestimating wind speeds which is reflected in the higher normalized standard deviation values (varying from 1 to 1.5 $\mathrm{m\,s^{-1}}$) for most of the selected buoys. The validation of significant wave heights from wind wave hindcasts against buoy data (Figs. 3 e-f) shows similar patterns to the ones obtained for satellite measurements. In general, Fig. 3 shows that for most of the buoys the correlation is between 0.8 and 0.9 with the CRMSE being around 0.5 m. Values of normalized standard deviation in CFSR-WW3 lie within 0.75-1.25.

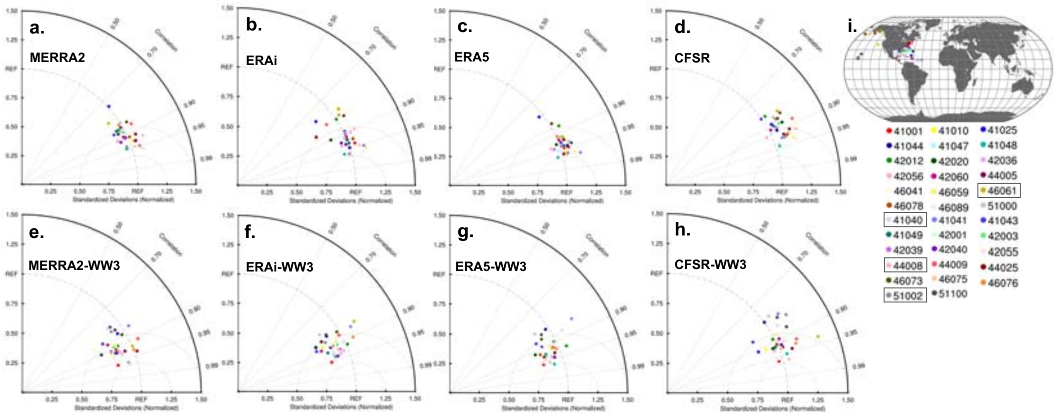

Figure 3. Taylor diagrams with error metrics for wind speed (a - d) and wave height (e - h) from reanalyses/hindcasts compared to NDBC buoys (the colors on diagrams correspond to buoy locations shown in (i)). The radial dotted lines represent correlation coefficients, the positions on X and Y axes represent the normalized standard deviation ($\mathrm{STD}_{model}/\mathrm{STD}_{buoy}$), the contour lines represent the normalized centered RMSE (CRMSE).

In order to compare the probability distributions, we use quantile-quantile (Q–Q) plots shown in Figure 4 for winds (a – d) and wave heights (e – h). For this purpose, we analyzed in-situ data in two distinct climate regions in both deep water and shallow water areas (shown in Fig. 1i in stars). In general, wind speed shows better agreement with buoy data than SWH, which is due to the fact that wind speed measurements are assimilated in reanalyses. For deep water buoys (Fig. 4 a, d) wind speed magnitudes are well captured in all reanalyses with a moderate overestimation in CFSR for the values of highest percentiles. The same is true for shallow water buoys (Fig. 4 c, b) where both CFSR and ERA-Interim overestimate higher wind speeds in particular for buoy № 46061 (in the Gulf of Alaska) while compared to buoy № 44008 (in shallow water in the Northwest Atlantic) all reanalyses indicate the higher winds than observed in buoy.

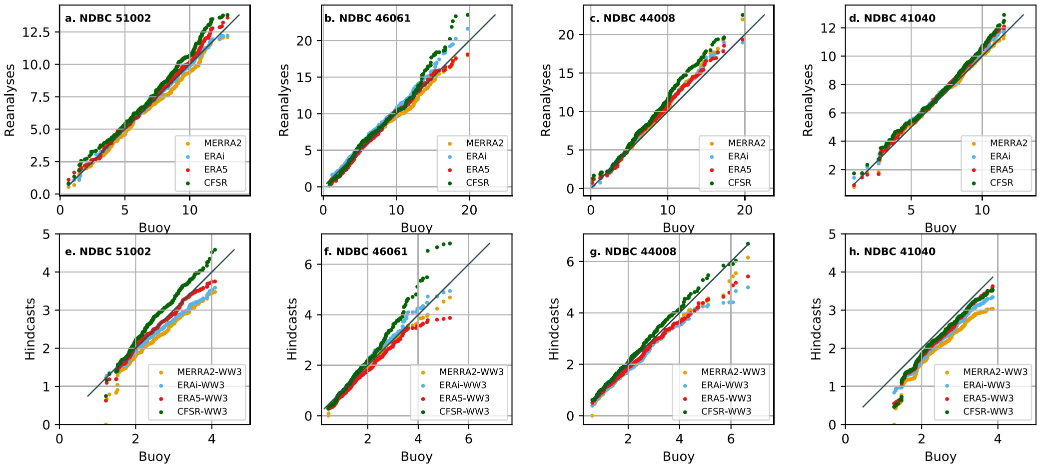

Figure 4. Q–Q plots of wind speed (a – d) and wave height (e – h) along the y-axis and data along the x-axis at selected buoys for 2011 (January - March); buoy locations are shown in stars in Fig.1i: №51002 (a, e) is in deep water near Hawaii, №46061 (b, f) is in shallow water in the Gulf of Alaska, №44008 (c, g) is in shallow water in the Northwest Atlantic, №41040 (d, h) is in the deep water in the tropical North Atlantic.

As for significant wave heights, for deep water buoys (Fig. 4 e,h) ERA5-WW3 is the closest to the buoy measurements, while ERAi-WW3 and MERRA2-WW3 moderately underestimate wave heights in all ranges of percentiles. CFSR-WW3 overestimates values higher than 2 m compared to buoys №51002 and №46061 in Pacific Ocean (Fig. 3 e,f). ERA5-WW3 shows a good correspondence in SWH with deep water buoy in the Pacific Ocean for all percentiles (Fig. 4 e), while it underestimates the high percentiles compared to shallow water buoy near Alaska (Fig. 4 f). MERRA2-WW3 and ERAI-WW3 hindcasts demonstrate overall underestimation of SWH for the deep water buoy (Fig. 4 e) and are very close to observations for the shallow water buoy (Fig. 4 f).

## 5. Conclusions

Four wind wave hindcasts based on WaveWatch III spectral wave model and modern atmospheric reanalyses ERA5, ERA-Interim, MERRA2 and CFSR have been developed for the 38-year period (1980-2017). The lowest overall bias for both winds and waves is observed for ERA5-based hindcast. MERRA2 and associated hindcast MERRA2-WW3 demonstrate overall relatively low biases in winds and also low biases in waves in midlatitudes, however the large discrepancies were found for this product in the equatorial and tropical regions. The comparison with satellite altimetry shows consistent patterns for winds and waves only for CFSR-WW3 hindcast with the largest errors found in the Southern Ocean. All obtained hindcasts corresponds well to buoy measurements both in the deep and shallow waters in the Atlantic and Pacific Oceans. ERA5-WW3 hindcast shows the best correspondence with deep water buoys in the whole range of percentiles, including high values. In general, all hindcasts show good correspondence with the observational data and thus can be used for further investigation of the common patterns of interannual variability in winds and wave heights in the wave climate studies.

## Acknowledgments

This research was supported by the Russian Ministry of Science and Higher Education (grant No 05.604.21.0210 project ID RFMEFI60419X0210).

## References

[1] Ardhuin F *et al* 2010 Semi-empirical Dissipation Source Functions for Ocean Waves. Part I: Definition, Calibration, and Validation *J. Phys. Oceanogr* **40(9)** pp 1917–41
[2] Dee D P, Uppala S M, Simmons A J, Berrisford P, Poli P, Kobayashi S, Vitart F 2011 The ERA-Interim reanalysis: configuration and performance of the data assimilation system *Quarterly Journal of the Royal Meteorological Society* **137(656)** pp 553–97
[3] Gelaro R *et al* 2017 The modern-era retrospective analysis for research and applications, Version 2 (MERRA-2) *J. Clim.* **30** pp 5419-54

[4] Hersbach H and Dee D 2016 ERA5 reanalysis is in production *ECMWF Newsletter* 147 (Reading: United Kingdom,ECMWF) [Available online at www.ecmwf.int/sites/default/files /elibrary/2016/16299¬newsletter¬no147¬spring¬2016 .pdf.]

[5] Hersbach H 2018 Operational global reanalysis: progress, future directions and synergies with NWP *ERA Report Series* 27 (Reading: United Kingdom,ECMWF)

[6] Taylor K E 2001 Summarizing multiple aspects of model performance in a single diagram *J. Geophys. Res.* **106 (D7)** pp 7183–92

[7] Saha S *et al* 2010 The NCEP Climate Forecast System Reanalysis *Bulletin of the American Meteorological Society* **91(8)** pp 1015–58

[8] Stopa J E and Cheung K F 2014 Intercomparison of wind and wave data from the ECMWF Reanalysis Interim and the NCEP Climate Forecast System Reanalysis *Ocean Modelling* **75** pp 65–83

[9] Saha S *et al* 2014 The NCEP Climate Forecast System Version 2 *J. Clim.* **27(6)** pp 2185–208

[10] Ardhuin F *et al* 2018 Measuring currents, ice drift, and waves from space: the Sea Surface KInematics Multiscale monitoring (SKIM) concept *Ocean Sci.* **14 (3)** pp 337-54

[11] Queffeulou P 2004 Long-Term Validation of Wave Height Measurements from Altimeters *Marine Geodesy* **27(3-4)** pp 495–510

[12] Queffeulou P and Croizé-Fillon D 2010 Global altimeter SWH data set *Tech. rep* version 7 (Ifremer)

[13] Stopa J E, Ardhuin F, Bababin A V, and Zieger S 2016 Comparison and validation of physical wave parameterizations in spectral wave models *Ocean Modell* **103** pp 2-17

[14] Morim J *et al* 2019 Robustness and uncertainties in global multivariate wind-wave climate projections *Nature Climate Change* **9** pp 711-8