Intercomparison and assessement of wave models at global scale

Technical Notes
May 2020
Issue TN0287

CMCC – Ocean and Data Assimilation (ODA) division

By ***Fabrizio Baordo, Emanuela Clementi, Dorotea Iovino and Simona Masina***
Fondazione CMCC Centro Euro-Mediterraneo sui Cambiamenti Climatici
simona.masina@cmcc.it

SUMMARY This report analyses the performance of two state of the art spectral wave models: the European model WAM (cycle 4.6.2) and the American model WAWEWATCH III (WW3, version 5.16). In this study, WAM and WW3, configured to run at global scale, are treated as 'stand-alone' models and are forced considering 10-meter winds and sea-ice cover from Numerical Weather Prediction (NWP) systems. To assess the sensitivity of the wave models to the spatial and temporal resolution of the wind input data, we use two different ECMWF datasets: the ERA5 reanalysis and the analysis of the operational high-resolution forecast system. Initially we evaluate the configurations of the wave models (named as CGWAM and CGWW3, where CG is the acronym to indicate 'CMCC Global'), by means of the Significant Wave Height (SWH) compared against the values provided by the CMEMS global wave analysis and forecast system (MFWAM) and the ECMWF ERA5 ocean wave reanalysis (ECWAM). This evaluation indicates that the SWH predicted by CGWAM and CGWW3 agrees with that provided by MFWAM and ECWAM in terms of large-scale patterns and seasonal variability, although with different magnitude. A more accurate statistical assessment for CGWAM and CGWW3 is successively conducted by means of CMEMS near real time in-situ and satellite observations. This validation shows that the wave models are both positively affected by the use of the ECMWF high resolution winds: if we consider the SWH, we observe a global reduction in the root mean square error of 4.7% (CGWAM) and 2% (CGWW3), when we validate the models against in-situ observations, and, 6.6% (CGWAM) and 2.4% (CGWW3), when Jason-3 altimetry measurements are used in the assessment. Overall, CGWW3 looks more skillful than CGWAM and this result is most likely due to the different formulation of the input and swell dissipation source terms implemented in the models. The assessment against in-situ and satellite observations also reveals that the modelled SWH is characterized by a global negative bias of few centimeters. To further investigate this systematic bias, we cluster the observed SWH according to different thresholds and it appears that the magnitude of the model bias increases with the increasing value of the significant wave height. In these circumstances (SWH larger than 2 m), the wave models, depending on the configuration and wind forcing, are affected by a bias which can vary from 20 to 70 cm.

Keywords WAM, WW3, MFWAM, ECWAM, Significant Wave Height, CMEMS products

1. INTRODUCTION

Nowadays, ocean wave modelling has reached a high level of accuracy and many weather forecasting centers run an operational wave forecast system. A wave model is mainly a wind-driven application and it has been shown that the accuracy of the wind data from Numerical Weather Prediction (NWP) systems plays an essential role in determining good wave forecasts (e.g. Cavaleri and Bertotti, 2004; Cavaleri et al., 2007). Recently, Janssen and Bidlot (2018) pointed out that about the 75% of the enhancement in wave height forecast skill, which has happened over the past 25 years, is related to the fact that, across the same time, atmospheric models have remarkably improved the quality of surface winds. In this analysis, the remaining 25% of the increased performance is associated with the actual improvements of the wave model which the authors assign to different causes such as: better representation of unresolved bathymetry, increase in spatial and spectral resolution, improved representation of wave dissipation and assimilation of altimeter wave height data.

The purpose of this study is therefore to intercompare the accuracy of the two state of the art spectral wave models, Wave Action Model – WAM (The WAMDI Group, 1988; Komen et al. 1994) and WAVEWATCH III – WW3 (The WAWEWATCH III Development group, WW3DG, 2016), when they are configured to run at global scale as “stand-alone” models. In this configuration the wave model is forced by 10-meter wind data from NWP systems and, as additional input, the sea-ice cover is used to distinguish ocean grid points at high latitudes. It is worth mentioning that sensitivity studies associated with the use of different source term parameterizations are beyond the scope of this report and they might be considered for future research. Hence, for both WAM and WW3, we consider a configuration which is our best choice for studies at global scale (the WAM and WW3 experiments are named respectively as CGWAM and CGWW3, where CG is the acronym to indicate ‘CMCC Global’). However, in this assessment, we consider the sensitivity of the wave models to the spatial and temporal resolution of the wind input data, and to do that, we used two different datasets: the ECMWF ERA5 reanalysis and the analysis of the operational ECMWF high resolution forecast system.

The report is organized as follows: Section 2 gives a general overview of the numerical wave models; Section 3 firstly provides a more detailed description of the wave models which have been used in this study and secondly, as a preliminary assessment, it shows a qualitatively intercomparison of the simulated significant wave height; Section 4 describes the observational dataset which has been chosen for the validation; Section 5 presents the results of the statistical assessment which validates the performance of the wave models by means of in-situ and satellite observations. Summary and conclusions are given in Section 6. In the appendix, a technical discussion on the computational resources to run CGWAM and CGWW3 is also provided.

2. NUMERICAL WAVE MODELS

In the third generation spectral wave models, the ocean waves are described with a two-dimensional wave action density spectrum N which is defined as the energy spectrum F divided by the intrinsic frequency σ. Neglecting diffraction and scattering effects which are generally not relevant at scales larger than 1 km (e.g. larger than the wavelength of ocean waves), the wave action density spectrum evolves in time and space according to:

$$\frac{\partial N}{\partial t} + \frac{\partial}{\partial \lambda}\left(\dot{\lambda} N\right) + \frac{\partial}{\partial \phi}(\dot{\phi} N) + \frac{\partial}{\partial k}(\dot{k} N) + \frac{\partial}{\partial \theta}(\dot{\theta} N) = \frac{S}{\sigma}, \quad (1)$$

where t is the time, λ and ϕ are the longitude and the latitude, κ is the wave number and θ is the wave propagation direction. The term S on the right-hand site of eq.1 (generally called the source function) describes the sources and the sinks of wave energy due to various physical processes. Basically, the physics of a wave model is exclusively related to the parameterizations adopted to implement S and different formulations have an impact on the resultant solution of the wave model. Usually, the source function is the sum of a number of terms which can be individually parametrized (e.g. $S = S_1 + S_2 + \ldots + S_n$). The formulation of S represents the effects due to processes that primary characterize the physics of wave modeling: the growth of the wave energy as a result of the variation of the stress on the wave surface by the wind (commonly identified as the input source term, S_{in}); the reduction of wave momentum and energy caused by different processes such as wave breaking, white-capping and bottom friction (the dissipation source function, S_{ds}); the non-linear quadruplet wave-wave interaction (generally called the non-linear source term, S_{nl}). For additional details on the formulation of the source terms, the reader can find exhaustive explanation and more references, for instance, in WW3DG (2016) and Janssen and Bidlot (2018). In the following section we provide a more detailed technical description of the wave models which have been used in this study and, as a preliminary assessment, we qualitatively intercompared the SWH generated by every spectral model.

3. GLOBAL OCEAN WAVE MODELS SET-UP AND QUALITATIVE INTERCOPARISON

3.1 CGWAM AND CGWW3

The ocean wave model WAM is an open source code developed by the WAMD Group and whose latest version is maintained by the German institute HZG (Helmholtz-Zentrum Geesthacht). The version used in this study is the 4.6.2 (Alari et al., 2016; Staneva et al., 2017) which is the stable and tested version of the code made available under the “coupled ocean-wave model development in forecast environment” Wave2NEMO project (https://www.mercator-ocean.fr/en/portfolio/wave2nemo-2). WW3 is also a community wave model which was originally developed in the United States by the National Oceanic and Atmospheric Administration (NOAA) and the National Centers for Environmental Prediction (NCEP). The version used here is the 5.16 (WW3DG, 2016) which was the last available at the time of this study. It is worth mentioning that a more recent version (6.07) was released in April 2019. WAM and WW3 are well-established wave models which have been widely used for both research and operational purposes. For instance, WAM is the wave model used as a component of the Earth system model and the operational forecast system at the European Center for the Medium-Range Weather Forecasts (ECMWF) and it is also the model implemented by the Copernicus Marine Environment Monitoring Service (CMENS) to deliver wave products. WW3, on the other hands, is operationally used by NOAA and NCEP as well as by the UK Met Office.

At CMCC, we configured WAM 4.6.2 and WW3 5.16 to run at global scale and to identify the outputs of the wave models we named our experiments as CGWAM and CGWW3, where CG is the acronym to indicate 'CMCC Global'. Before describing the configuration chosen for both CGWAM and CGWW3, it is worth mentioning technical differences between the two models regarding the source term formulations and the model grid. WAM contains only a determined number of source term formulations (e.g. 1 package for the S_{in}, 1 package for S_{nl} and so on) which can be activated and tuned by modifying the numerical value of the variables which characterize that specific parameterization. WW3, in this respect, provides more options for the parameterization of the source terms and consequently the user can choose between several different packages. In this way, the configuration of WW3 is identified by a set of switches which are necessary as an input to build the programs. Successively, a namelist file is used to set the numerical value of the variables for the selected source terms. The additional basic difference between WAM and WW3 is in the generation of the model grid. WAM has its own source conde (program "Preproc") which generates the model grid, and, in this case, 2 options are available: a regular grid (constant increments in latitude and longitude) and a "reduced grid". The latter is implemented to reduce the number of grid points for each latitude with increasing longitude. Basically, since the distance between longitudes is reduced towards the poles, this results in a non-homogeneous grid resolution and a strong reduction of the propagation time step is required to avoid numerical instability. The “reduced grid” approach overcomes this problem and, additionally, the computational efficiency is improved. Finally, the program “Preproc” uses ETOPO2 data to generate the bathymetry and it is also capable to generate the blocking mask to handle unresolved or poorly resolved sub grid features (e.g. islands). The technique to handle the sub-grid bathymetric features was originally developed by Tolman (2003) who showed the importance of considering the amount of wave energy that can be advected through these unresolved obstacles. Model grid (regular or curvilinear) for WW3 can be generated by an external Matlab package, named Gridgen, provided by NOAA (https://github.com/NOAA-EMC/gridgen). Gridgen is a software package that, in order to create the model bathymetry, uses two types of datasets: 1) a high-resolution global bathymetry (currently the choice is between ETOPO2 or ETOPO1 data); 2) a high-resolution shoreline database (GSHHS − Global Self - consistent Hierarchical High - resolution Shoreline; Wessel, P. and W. Smith, 1996) that is employed to determine the coastal boundaries. Gridgen is also capable to generate the blocking mask used by WW3 to deal with obstructions and poorly resolved features. Another intrinsic difference between WAM and WW3 regards the way in which the two models solve eq.1. WAM solves the balance equation in terms of spectral energy, while WW3 does it in the wave number space. Numerically, WAM solves the wave transport equation explicitly where the source terms and the propagation are computed with different methods and time steps. The source term integration is carried out using a semi implicit integration scheme while the propagation scheme is a first order upwind flux scheme. In WW3, on the contrary, the action balance equation is solved using a fractional step method. The first step considers temporal variations of the depth, and corresponding changes in the wave number grid. Other fractional steps consider spatial propagation, intra-spectral propagation and source terms. In WW3, the propagation scheme is configurable.

Finally, in **Table 1**, we summarize the configurations for CGWAM and CGWW3 which were selected to conduct the assessment study. Essentially, CGWAM and

CGWW3 have been configured as close as possible apart from those choices which are the default options for WAM. However, it is important to highlight that these configurations probably represent the optimal choice for both the models. From one side, the configuration for CGWAM should be similar to that implemented in the operational wave model at ECMWF (ECWAM cycle 38R1). From the other side, the set-up for CGWW3 has been already tested and validated by previous studies (e.g. Ardhuin et al., 2010; Rascle et al., 2008; Rascle and Ardhuin, 2013). Nevertheless, it is important to list the primary differences which might have an impact on the final gridded outputs of CGWAM and CGWW3:

a) bathymetry: at the resolution of the model grid, most likely the use of ETOPO1 or ETOPO2 data does not have a significant impact.

b) the diverse treatment of ice concentrations may generate differences only in a limited number of model grid points localized at high latitudes.

c) a slightly more significant impact might be generated by the different propagation schemes, which however, are the default (WAM) and recommended (WW3) options for the wave models.

d) at large scale, the different formulation of the input and dissipation source terms is the parameterization which more likely plays the dominat role in generating differences between the two wave models.

We expect that the las point is the key to interpret the results of our validation: as discussed by Ardhuin et al. (2010), Rascle et al. (2008), Rascle and Ardhuin (2013), the Ardhuin’s wave growth and dissipation parameterization, which was assessed by in-situ and remote sensing observations of significant wave height, peak and mean periods, produces more accurate results with respect to those generated by the equivalent source term package implemented in WAM 4.6.2.

For general nomenclature, the configuration of the source term package used in the GCWW3 experiment is called TEST471 (WW3DG, 2016) which was found to provide the best results at global scale when ECMWF winds are used as atmospheric forcing. It is worth mentioning that some work is in progress to implement the Ardhuin et al. (2010) physic package in WAM and most likely it will be available in the next code release (WAM 4.7).

Table 1: Summary of the configuration selected for CGWAM and CGWW3.

	CGWAM (WAM 4.6.2)	CGWW3 (WW3 5.16)
Model Grid	Regular 0.25 x 0.25, 80S-89N, generated by Preproc	Regular 0.25 x 0.25, 80S-89N, generated by Gridgen
Bathymetry	Default - ETOP2 + Unresolved sub grid features	ETOP1 + Unresolved sub grid features
Sea Ice Concentration	Default - The wave spectra at all grid points marked as sea ice is set to zero after a propagation	Sea ice concentrations greater than 75% are treated as land, ice concentrations less than 25% are treated as open ocean; partial sub-grid blocking otherwise, as described by Tolman (2003)
Wave Spectrum Discretization	24 directions; 30 frequencies starting from 0.035 Hz	
Global Time Step	6 Minutes	
Propagation Scheme	Default - First order upwind flux scheme	Third order with GSE alleviation (Tolman, 2002)
Source Terms a) Input + Dissipation b) Non-linear Interaction c) Bottom friction d) Depth-induced breaking	Default a) Bidlot (2007, 2012) b) Discrete interaction approximation (Hasselmann et al., 1985) c) JONSWAP d) Battjes and Janssen (1978)	 a) Ardhuin et al. (2010) b) Discrete interaction approximation (Hasselmann et al., 1985) c) JONSWAP d) Battjes and Janssen (1978)
Spatial/Temporal Resolution Atmospheric	ECMWF ERA5: • u10m + v10m 0.25° x 0.25°/3-hourly	

Forcing	• sea ice cover 0.25°x 0.25°/daily (at 00z) ECMWF High Resolution Model (ECHRES): • u10m + v10m 0.125° x 0.125°/6-hourly • sea ice cover 0.125° x 0.125°/daily (at 00z)
Experiments Duration	1 May 2017 00z – 01 April 2018 00z (starting from calm conditions)
Spatial/Temporal Resolution Model Outputs	0.25° x 0.25°/3-hourly

As reported in **Table 1**, the sensitivity of the wave models to the spatial and temporal resolution of the wind input data were tested using atmospheric forcing from two distinct dataset:

- the ECMWF ERA5 reanalysis which is characterized by a horizontal resolution of 0.25° x 0.25°. The 10-meter u and v10-meter components of wind and the sea-ice cover were selected respectively considering a temporal resolution of 3 (00z, 03z, 06z and so on) and 24 (at 00z) hour. The dataset used in this study was downloaded from the Copernicus Climate Data Store (CDS):

  https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-single-levels?tab=form (last access October 2019).

- the ECMWF high resolution forecast system (ECHRES): in this case, the operational analysis of winds and sea-ice were retrieved directly from the Meteorological Archival and Retrieval System (MARS) on a regular grid with horizontal resolution of 0.125° x 0.125°. The temporal resolution of the sea-ice cover is equivalent to that of the ERA5 dataset, while the frequency of the wind data is reduced to 6 hours (analysis at 00z, 06z, 12z and 18z).

In the assessment presented in the following sections, to distinguish the experiments which are characterized by different atmospheric forcing, together with the name of the wave model (CGWAM or CGWW3) we use the string "era5" and "echres" (e.g. CGWAM-era5, CGWAM-echres, CGWW3-era5 and CGWW3-echres).

3.2 BENCHMARK SYSTEMS: MFWAM AND ECWAM

The statistical assessment of CGWAM and CGWW3 is conducted by validating the model outputs against in-situ and satellite observations. However, to generally assess the consistency of the configuration selected for CGWAM and CGWW3, we qualitatively intercompared the SWH with the outputs of two additional operational wave models:

- MFWAM: The operational global wave analysis and forecast system of the Copernicus Marine Environment Monitoring Service (CMEMS) which is implemented at Météo-France. Essentially, the computing code is based on WAM (for this reason this system is named as MFWAM) and particularly on the ECWAM cycle 38R2. The basic difference with ECWAM is related to the implementation of the input and dissipation terms which was upgraded to that developed by Ardhuin et al. (2010). In terms of atmospheric forcing, MFWAM is driven by 6-hourly analysis and 3-hourly forecasted winds from the ECMWF high resolution forecast system, while the wave spectrum is characterized by 24 directions and 30 frequencies starting from 0.035 Hz. MFWAM is not configured as "stand-alone" model, but it is also forced by the surface currents provided by the CMEMS global ocean forecasting system (product identified as "GLOBAL ANALYSIS FORECAST PHY 001 024") with a daily update and it operationally assimilates, every 6 hours, altimeter wave data from Jason 2, Jason 3, Cryosat-2 and Sentinel 3A together with synthetic aperture radar (SAR) wave spectra from Sentinel 1A. The global CMEMS wave product (GLOBAL ANALYSIS FORECAST WAV 001 027) consists of 3-hourly wave data provided on a regular global grid at horizonal resolution of 1/12°. More details about the model and the wave products can be found on the documentation available online (CMEMS Product User Manual and CMEMS Quality Information Document, see bibliography for details). In the results presented in this report, the MFWAM wave outputs were download from the CMEMS ftp site:

ftp://nrt.cmems-du.eu/Core/GLOBAL_ANALYSIS_FORECAST_WAV_001_027/global-analysis-forecast-wav-001-027 (last access October 2019).

- The operational ECMWF ERA5 ocean waves reanalysis (hereafter simply identified as ECWAM). ERA5, the successor of ERA-Interim, is produced using the cycle 41R2 of the ECMWF 4D-Var Integrated Forecast System (IFS). In this configuration, the IFS is coupled to a soil model and an ocean wave model. The latter provides the outputs of the wave model which were used in this study. In this configuration, ECWAM is discretized with 24 directions and 30 frequencies (starting from 0.035 Hz) and the output parameters are provided on a regular grid at horizontal resolution of 0.5°. This version of ECWAM uses the input and dissipation source term as that adopted in our CGWAM experiments (Bidlot et al., 2007, 2012). Only recently, the wave physics package in ECWAM cycle 46R1 (June 2019) was updated with the parameterization developed by Ardhuin et al. (2010). As in the MFWAM wave model, also ECWAM implements the assimilation of altimeter and SAR wave observations. Additional details on ECWAM, can be found on the documentation available online (ECWAM-Cy41R2, see bibliography for details). The ERA5 ocean wave reanalysis dataset used in this study was downloaded from the Copernicus CDS website:

  https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-single-levels?tab=form (last access October 2019).

3.3 QUALITATIVE INTERCOMPARISON

In this section we provide a first assessment of CGWAM and CGWW3 which is not based on the use of observations, but we intercompare the Significant Wave Height (SWH) in order to verify that, on average, our experiments provide similar results to those of MFWAM and ECWAM. To do that, **Figure 1** and **Figure 2** explore the modelled SWH respectively in terms of latitude/longitude maps and time series. The SWH, which is at the native resolution of the models, in the case of the geographical maps, is averaged considering the model outputs at 00z for the month of July 2017 (note that similar results are obtained if we consider outputs at different times, e.g. 06z, 12z and so on). Although with difference in magnitude, **Figure 1** shows that all the models present comparable features at global scale. MFWAM is the model which provides the largest values of SWH especially in the storm track regions in the Southern Ocean. We might speculate that the capability of MFWAM to resolve these features in more detail is probably related to the higher spatial resolution together with a more complete configuration of the system (e.g. data assimilation and use of ocean currents as forcing fields). The other general characteristic which is visible in the maps is the effect of the different wind forcing in the CGWAM/CGWW3 experiments. The ECMWF high resolution wind data, with respect to the case when the ERA5 dataset is used as atmospheric forcing, produces an increase in the SWH magnitude particularly visible at high latitudes, but also in the Arabian Sea.

Time series shown in **Figure 2** are computed considering the spatial mean of SWH (at the model native resolution) for every model output (3-hourly) from 1 June 2017 to 1 April 2018 and for 3 distinct latitude bands: Northern Hemisphere (30N-89N), Tropics (30S-30N) and Southern Hemisphere (80S-30S). **Figure 2** provides a general overview of the SWH seasonal variability in the Northern and Southern Hemisphere. Overall, we can again conclude that the models show a similar behaviour. The sensitivity of wave models to the different atmospheric forcing is also evident: the use of ECMWF high resolution wind data systematically brings an increase in the magnitude of SWH, in both the CGWAM and CGWW3 experiments, across all the time series. As additional qualitative assessment, **Figure 3** extends the content of **Figure 2** showing the time series of the standard deviation of SWH. Statistics confirm the good agreement of the wave models at global scale.

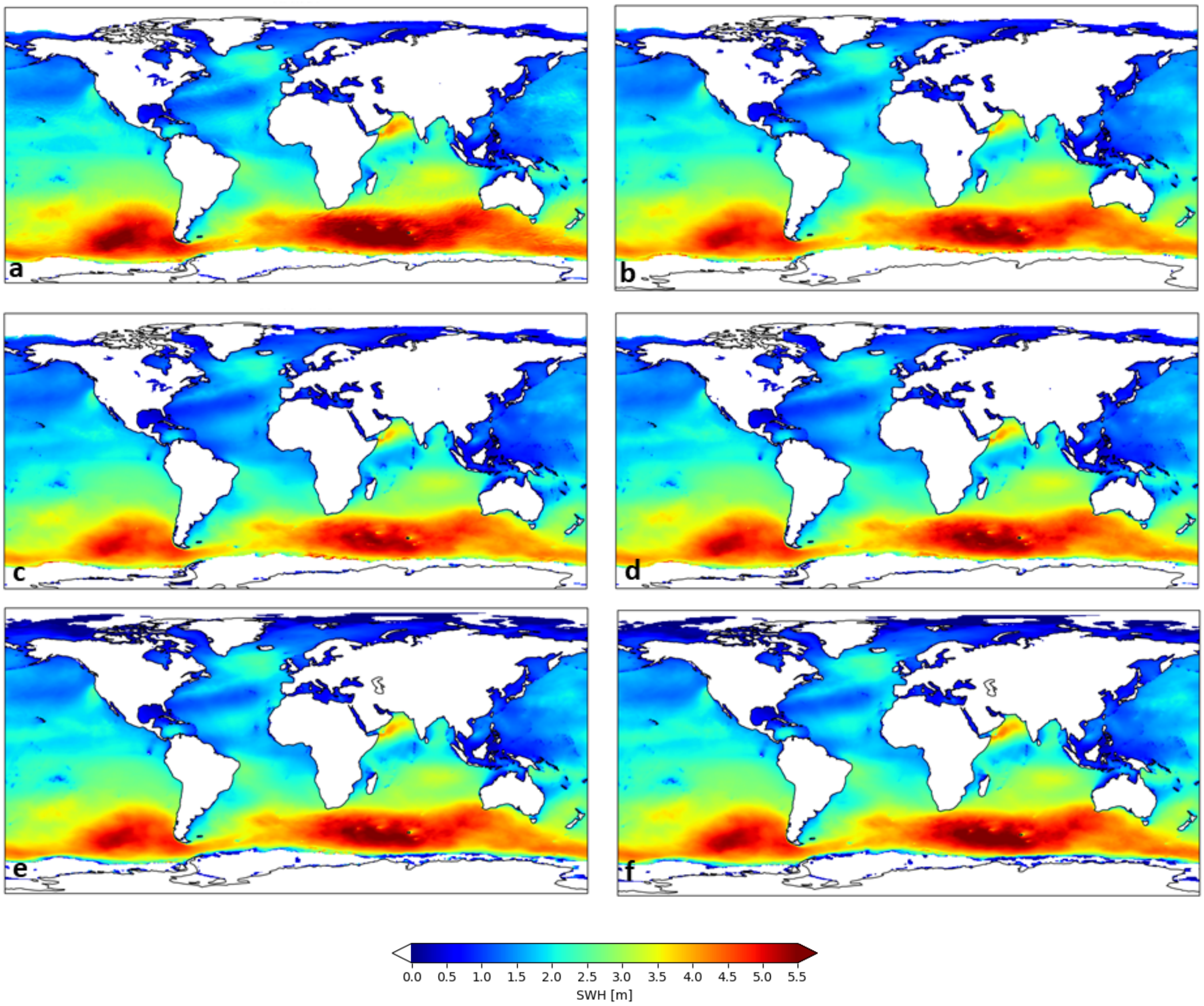

Figure 1: *Maps of mean Significant Wave Height (SWH) for July 2017 at 00z: a) MFWAM, b) ECWAM, c) CGWAM -era5, d) CGWAM-echres, e) CGWW3-era5 and f) CGWW3-echres. The mean value is calculated considering the gridded outputs of the wave model at the native resolution.*

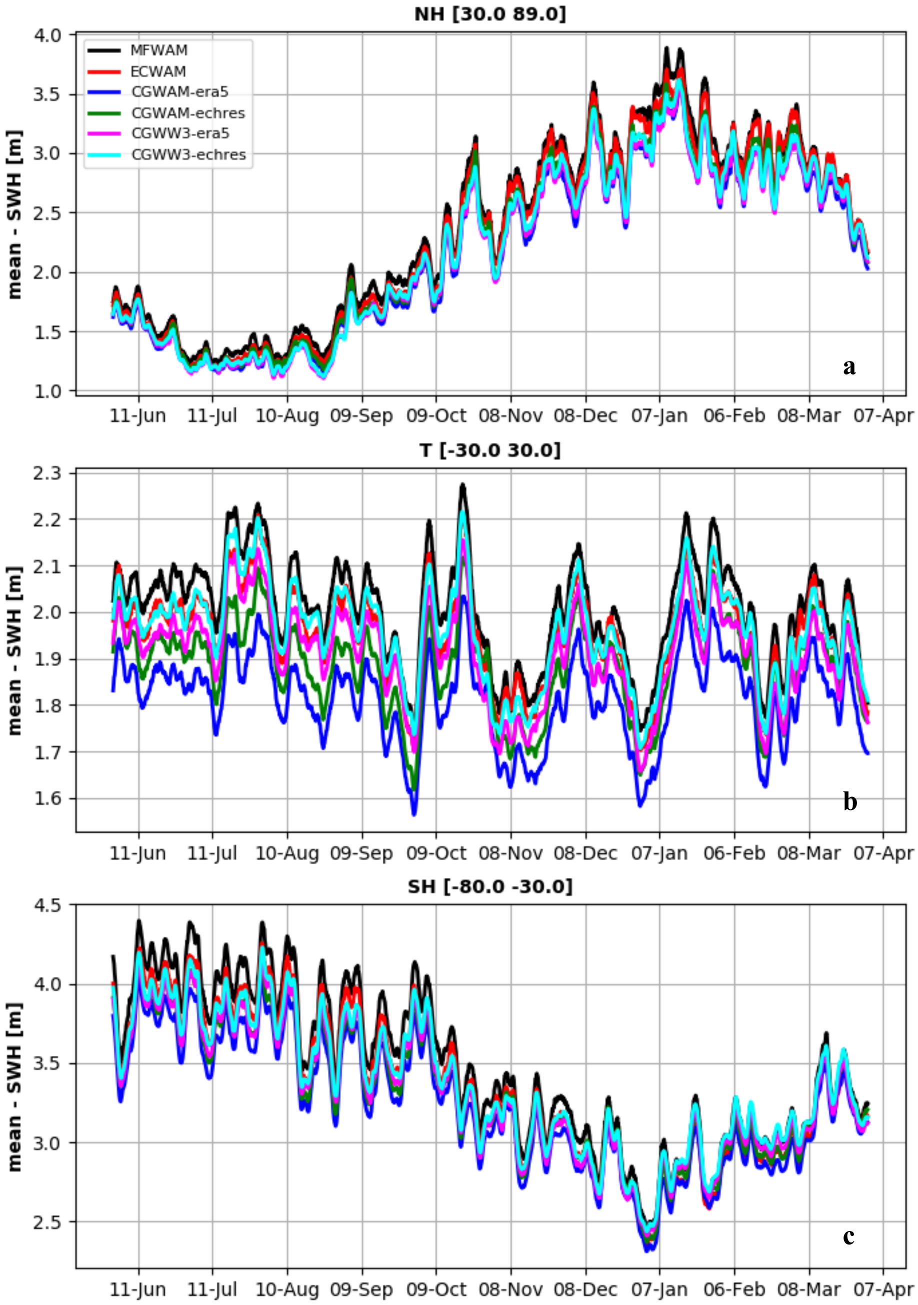

Figure 2: *Time series (from 1 June 2017 to 1 April 2018) of mean Significant Wave Height (SWH) in meters: MFWAM, black line, ECWAM, red line, CGWAM-era5, blue line, CGWAM-echres, green line, CGWW3-era5, purple line, and CGWW3-echres, cyan line. Every point of the time series represents the mean of the 3-hourly gridded outputs of the wave model at the native resolution. As a reference, on the x-axis, an interval of 30 days is given. Time series are shown for 3 distinct latitude bands: a) Northern Hemisphere (30N-89N), b) Tropics (30S-30N) and c) Southern Hemisphere (80S-30S).*

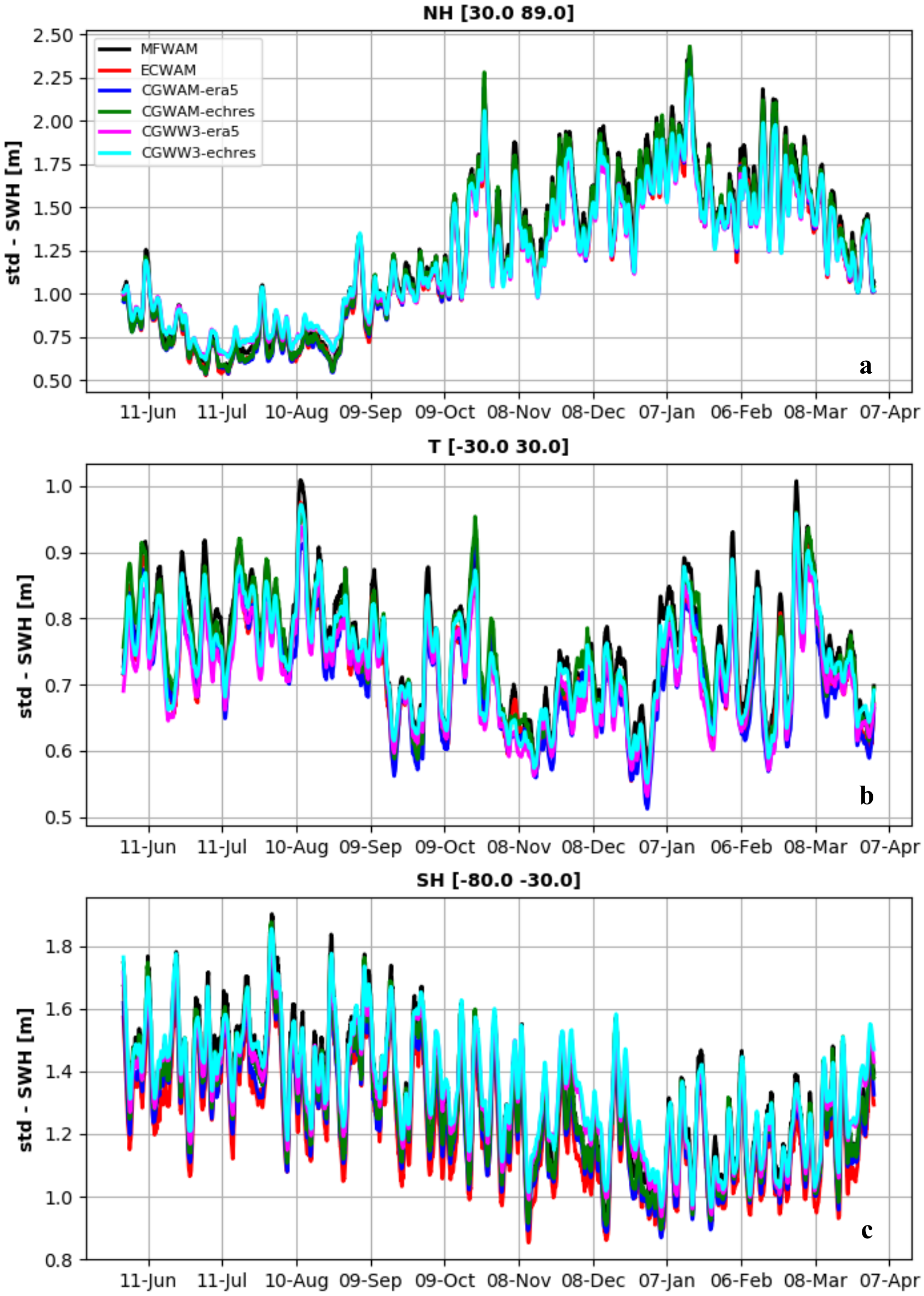

Figure 3: *As* ***Figure 2****, but showing the time series of the standard deviation (std) of Significant Wave Height (SWH).*

4. OBSERVATIONAL DATASET

In a numerical spectral model, the actual parameter which is calculated at each grid point is the two-dimensional wave spectrum (F) that describes how the wave energy is distributed according to the selected spectral discretization (number of frequencies, $\square$, and directions, $\square$). Usually, to simplify the analysis of wave models, the resultant output parameters are computed considering different weighted integrals of the wave spectrum. Among the possible outputs of a wave model, the following 3 parameters are those which can be compared with measurements collected by in-situ instruments (e.g. moored buoy stations):

- Significant Wave Height (SWH): $4\sqrt{\iint F(\square,\square)d\square d\square}$;
- Mean Wave Direction (MWD): $\arctan\left(\frac{\iint sen\square F(\square,\square)d\square d\square}{\iint cos\square F(\square,\square)d\square d\square}\right)$;
- Wave Peak Period (WPP): corresponds to the period of the most energetic wave component.

The modelled SWH can additionally be verified with altimeter data observed by satellite platforms. The statistical assessment, which is presented in the next section, was conducted using 2 distinct CMEMS observational dataset which are available on the ftp site:

- The Global Near Real Time In-situ Observations (Moorings), available online at:

  ftp://nrt.cmems-du.eu/Core/INSITU_GLO_NRT_OBSERVATIONS_013_030 (last access October 2019). The CMEMS in-situ observational dataset was used to validate CGWAM and CGWW3 outputs of SWH, MWD and WPP.

- The Global Near Real Time Satellite Observations, available online at:

  ftp://nrt.cmems-du.eu/Core/WAVE_GLO_WAV_L3_SWH_NRT_OBSERVATIONS_014_001 (last access October 2019). Altimeter data collected by Jason-3 and Sentinel-3a were used to validate CGWAM and CGWW3 outputs of SWH.

5. WAVE MODELS VALIDATION

In this section we present the results of the statistical assessment to validate the performance of CGWAM and CGWW3. It is important to highlight some of the choices made in the validation:

- To minimize the spin-up impact, the model outputs for the entire month of May 2017 was discarded in our assessment.
- The bias is calculated as the difference between model and observation.
- The error is expressed in terms of Normalized Root Mean Square Error (NRMSE): the RMSE is normalized relative to the root mean square value of the observations, so that the error signal is expressed in term of the observed signal. This in general allows a quantitative comparison between widely different sea state regimes.

5.1 VERIFICATION AGAINST IN-SITU OBSERVATIONS

The CMEMS Global Near Real Time In-situ Observations are stored in a NetCDF file which, for a specific buoy (e.g. GL_TS_MO_42001.nc), contains a time series of data with time resolution of 10 minutes. For every buoy, the criteria which was implemented to select observations and the equivalent model values are as follows:

a) Firstly, observations are quality checked by means of the QC flag which is available in the CMEMS NetCDF file. We keep only those observations which are flagged as “good data” or “probably good data”.

b) Secondly, we selected only those observations which are distant in time no longer than 30 minutes with respect to the model time (e.g. if model time is 00z the allowed observation time window is 23:30z-00:30z).

c) If more than one good observation is available after step a and b, as final measurement we consider the mean value and, from the NetCDF file, we also extract the latitude and longitude of the mooring location.

d) Finally, as model value, we consider the nearest neighbor to the mooring location.

Figure 4 shows the location of the 249 moorings which were selected to conduct the assessment study.

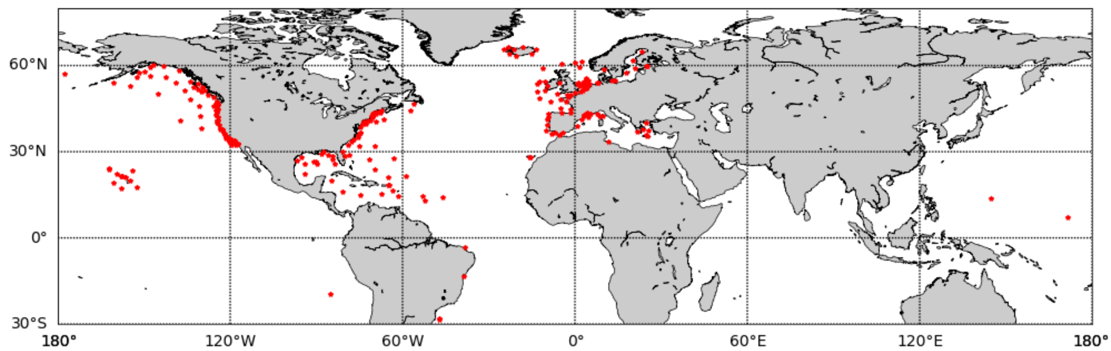

Figure 4: *Location of the 249 moorings used in the assessment study.*

Statistics are computed considering all the observations collected from 1st June 2017 to 1st April 2018 by the moorings in **Figure 4**. However, to diversify the assessment, the equivalent analysis was also performed considering a subset of moorings located in different geographical areas. The regions selected for this additional assessment (shown in **Figure 5**) are as follows: Gulf of Mexico and the Caribbean (33 moorings), North America (102 moorings) and Europe (95 moorings).

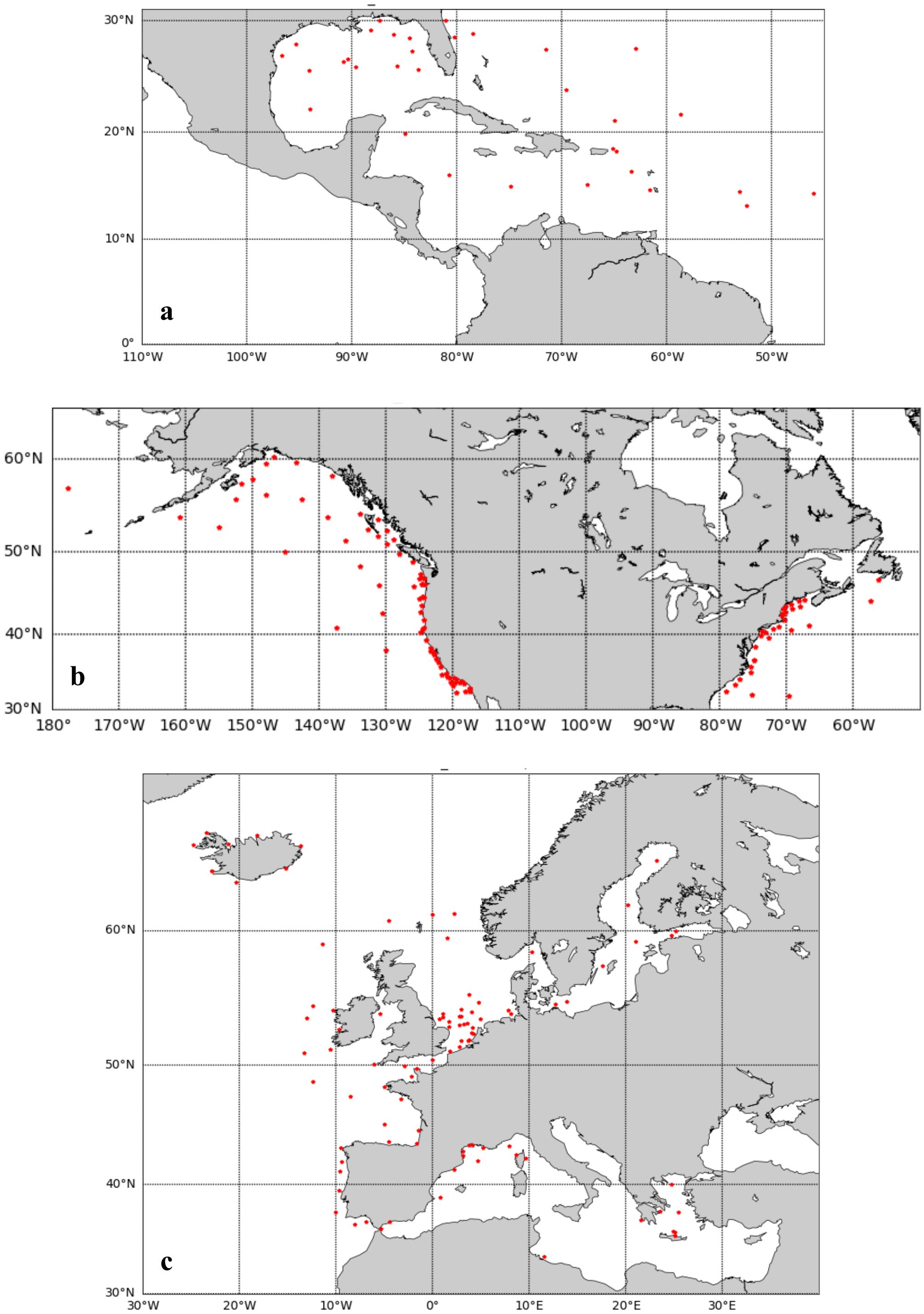

Figure 5: *Location of moorings selected for regional assessment: a) Gulf of Mexico and Caribbean, b) North America and c) Europe.*

An overall summary of the assessment for the four geographical areas is provided by **Figure 6**, **Figure 8** and **Figure 9** which respectively explore the statistical analysis for the Significant Wave Height, the Mean Wave Direction and the Wave Peak Period. In every figure, we compare bias, NRMSE and, by means of Taylor diagrams, the Pearson correlation coefficients, the centered root mean square differences and the standard deviations. Number of observations used to calculate the statistics is also shown for every geographical region.

Let us focus on the results for the SWH (**Figure 6**) which is a critical parameter for wave modelling in both coastal and large-scale applications. An accurate prediction is extremely important in decision-making particularly for those intense storm events in which wave heights can be larger than 10 meters. The overall outcome can be summarized as follows:

a) All the experiments show particularly good skills in simulating the SWH and this can be generally stated looking at the high value of the correlation coefficients in every geographical area (always greater than 0.94).

b) The impact of the different atmospheric forcing is noticeable: statistics suggest that higher horizontal resolution, which is most likely associated with an improved quality of the wind data, plays a fundamental role in the performance of the wave model. Both CGWAM and CGWW3 show a reduction in the bias and NRMSE as well as higher values of correlation coefficients when the input wind data comes from the ECMWF high resolution system. From this assessment, we might conclude that, in wave modelling for large scale applications, a higher temporal frequency (e.g. 3-hourly with respect to 6-hourly) is not as crucial as the spatial resolution of the NWP wind data. However, for coastal application higher frequency wind forcing might have an impact.

c) In the tested configurations, CGWW3 looks in general more skillful that CGWAM. As discussed in section 3.1, we may infer that the reason of the better performance of CGWW3 is essentially due to the different formulation of the input and dissipation source terms.

d) Overall, the experiments show a systematic negative bias.

To investigate the systematic negative bias in the SWH in more detail, we clustered the global observations of significant wave height (those of **Figure 6a**) according to 3 different thresholds

1) 0.2 m < SWH <= 2 m (THR1);

2) 2 m < SWH <= 4 m (THR2);

3) SWH > 4 m (THR3);

and, successively, we computed the corresponding model mean bias. Results of this analysis is summarized in **Figure 7**. The general outcome is very interesting: all the wave experiments show a slight bias of few centimeters for small values of SWH (less than 2 m) and, under this regime, CGWW3 is characterized by the lowest bias (-1.24 cm and 1.03 cm respectively for CGWW3-era5 and CGWW3-echres). As the observed SWH increases (THR2 and THR3), also the wave models are characterized by an increased negative bias which, in case of large values of SWH (higher than 4 m), can be greater than 56 cm. Overall, CGWW3-echres and CGWAM-echres always show a smaller bias than the equivalent 'era5' experiments and this confirms the positive impact associated to the use of the ECMWF high resolution winds. As a general conclusion, **Figure 7** provides an indication that the negative bias, which affects the wave models at global scale, is largely linked to those very high values of SWH that the modeling has difficulty in representing. We might speculate that this behavior is mainly connected to bias in the wind forcing: it is known that in some intense storm conditions, the NWP systems may not be able to properly predict the right direction and intensity of the winds. In these conditions, the NWP wind speed is often underestimated, and the obvious consequence is that the magnitude of the modelled SWH is too small in comparison with the observations.

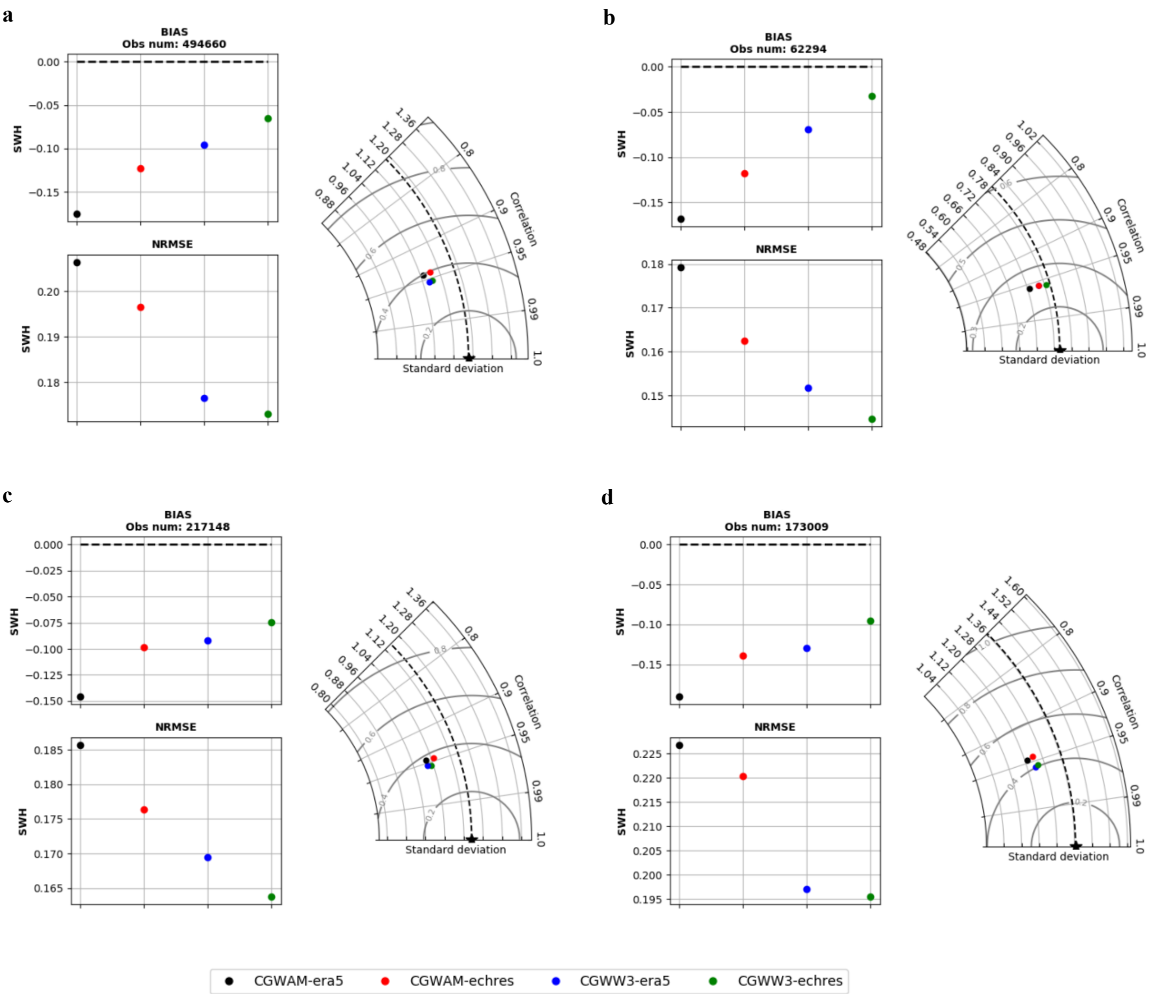

Figure 6: *Statistics for the Significant Wave Height (unit in meters) calculated considering the observations collected from 1 June 2017 to 1 April 2018 by moorings in different geographical areas: a) Global, b) Gulf of Mexico and Caribbean, c) North America and d) Europe. In every panel, the following statistics are displayed: Bias, Normalized Root Mean Square Error (NRMSE), and in the Taylor diagram, Correlation, Standard Deviation and Centered Root Mean Square difference (gray contours). As a reference, the number of observations used to calculate the statistics is also shown and the black star in the Taylor diagrams indicates the standard deviation of the observations. CGWAM and CGWW3 experiments are identified by the colored dots as specified in the legend.*

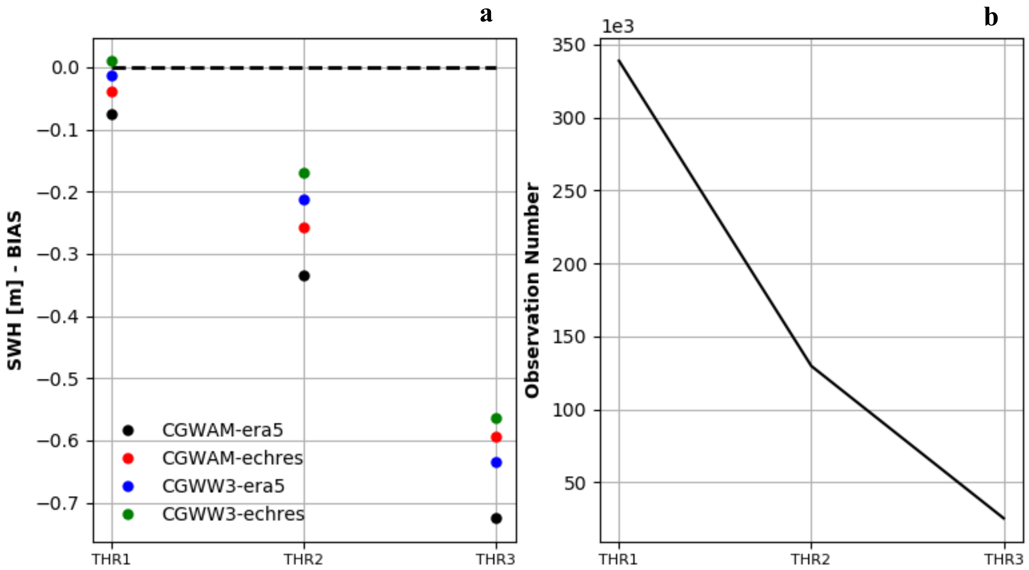

*Figure 7: Model mean bias (a) computed considering the observational dataset of **Figure 6a**, but clustering the value of SWH according to 3 different thresholds: THR1: 0.2 m < SWH <= 2 m; THR2: 2 m <= SWH <= 4 m; THR3: SWH > 4 m. Number of observations used to calculate statistics in every cluster is shown in panel b. CGWAM and CGWW3 experiments are identified by the colored dots as specified in the legend.*

To expand further the statistical assessment, for every geographical area, **Table 2** presents the percentage difference of SWH NRMSE between two experiments calculated as $100*(NRMSE_{exp2} - NRMSE_{exp1})/NRMSE_{exp1}$. As mentioned above, the higher resolution atmospheric forcing has positive impact on the wave model skill: from **Table 2** we can infer that, at global scale, the reduction in NRMSE is about 1.98% and 4.75% respectively for CGWW3-echres and CGWAM-echres when they are compared with the equivalent experiments CGWW3-era5 and CGWAM-era5. However, we can also observe that the reduction in NRMSE is much larger in the Gulf of Mexico (4.68% and 9.43%) than in Europe (0.81% and 2.82%). These results might suggest that in some areas of the globe, the quality and resolution of the wind data remain the critical key for improving the skills, but in other regions, implementation and tuning of different configurations and parameterizations in the wave model might be the main cause of the enhancement in the performance. In saying that, the significance of the source term formulation is highlighted in **Table 2**: CGWW3 outperforms CGWAM and, independently of the wind forcing, the reduction of NRMSE at global scale is larger than 11%.

Table 2: *Analysis of percentage difference in NRMSE (calculated using in-situ observations) for the Significant Wave Height*

Percentage difference in NRMSE (%) is computed as: 100 *($NRMSE_{exp2}$ - $NRMSE_{exp1}$)/$NRMSE_{exp1}$, so that negative/positive values indicate an overall decrease/increase in the error for exp2.				
	Percentage difference in NRMSE (%) for SWH			
exp1/exp2	**Global**	**Gulf of Mexico and the Caribbean**	**North America**	**Europe**
same wave model, different forcing				
CGWAM-era5/CGWAM-echres	-4.75	-9.43	-5.01	-2.82
CGWW3-era5/CGWW3-echres	-1.98	-4.68	-3.36	-0.81
same forcing, different wave model				
CGWAM-era5/CGWW3-era5	-14.44	-15.34	-8.72	-13.10
CGWAM-echres/CGWW3-echres	-11.95	-10.90	-7.14	-11.30

To conclude, it is worth comparing some of our statistics with those obtained by similar studies already done in literature. For instance, Saulter et al. (2016) assessed scientific and technical changes in the operational UKMO wave forecast system. In that study, the authors investigated a new WW3 configuration at global scale based on the physical package of Ardhuin et al. (2010) and the use of a Spherical Multiple-Cell grid. The validation was performed considering a one-year period covering September 2014 to August 2015 and in-situ measurements from the Joint Commission On Marine Meteorology's operational Wave Forecast Verification Scheme (JCOMM-WFVS, Bidlot et al., 2007a). There are obvious differences with the configuration of our wave models and the in-situ dataset (type and period) used to compute statistics at global scale, but it is very interesting to show that the results we obtained are similar to that presented by Saulter et al. (2016). **Table 3** summaries the comparison between our statistics and those presented in the study of the UKMO.

Table 3: *Comparison with the UKMO statistics (Saulter et al., 2016) for SWH at global scale.*

<table>
<tr><th colspan="2">Wave Model</th><th>SWH [m]
NRMSE</th><th>SWH [m]
Correlation</th></tr>
<tr><td colspan="2">UKMO before upgrade</td><td>0.18</td><td>0.95</td></tr>
<tr><td colspan="2">UKMO after upgrade</td><td>0.14</td><td>0.97</td></tr>
<tr><td colspan="2">CGWW3-echres</td><td>0.17</td><td>0.95</td></tr>
<tr><td colspan="2">CGWAM-echres</td><td>0.20</td><td>0.94</td></tr>
<tr><td colspan="4">Observed mean and standard deviation for SWH [m]</td></tr>
<tr><td></td><td>Mean</td><td colspan="2">Standard Deviation</td></tr>
<tr><td>UKMO
Sep 14 – Aug 15</td><td>1.82</td><td colspan="2">1.23</td></tr>
<tr><td>This study
Jun 17 – Apr 18</td><td>1.75</td><td colspan="2">1.18</td></tr>
</table>

Figure 8 and **Figure 9** represent the equivalent statistical analysis shown for the SWH, but respectively for the Mean Wave Direction and the Wave Peak Period. Results of **Figure 8** and **Figure 9** show that these parameters are poorly replicated by the wave models which are generally characterized by a small degree of correlation with the observed measurements (correlation coefficients always lower than 0.9). The

peak period is a parameter which is prone to large errors which can occur in the presence of multi-modal seas. In these cases, a small variation in the energy assigned to one spectral component can lead to a large variation in the identification of the peak frequency. In the same way, the wave direction is made up of contributions from wind-sea and swell (with different frequencies and directions) which interact in a complex manner, so errors in the representation of these fields directly affect the performance of the model in reproducing mean directions. The error in WPP and MWD can also increase depending on local conditions such as storms or strong currents. Additionally, it is important to consider that bias in the direction of the input wind data may be another source of error which impacts on the wave directions.

In summary, statistics for MWD and WPP demonstrate similar results to those previously discussed for the significant wave height: the high-resolution atmospheric forcing brings a positive impact in the CGWAM and CGWW3 experiments; CGWW3 looks more skillful than CGWAM. As also discussed by Rascle and Ardhuin (2013), the improvements in the mean wave direction is mostly likely due to the physical package of Ardhuin et al. (2010) which provides a better representation of swell fields. It is worth pointing out that these conclusions are generally true except for the statistics of WPP in the North America area (**Figure 9c**). For some reasons, in this area the use of the high-resolution wind data seems to penalize both the experiments (CGWAM-echres and CGWW3-echres) and consequently the statistics at global scale are also affected. Nevertheless, **Table 4** shows a very good agreement between our global WPP statistics and those presented in Saulter et al. (2016, where statistics for MWD are not provided).

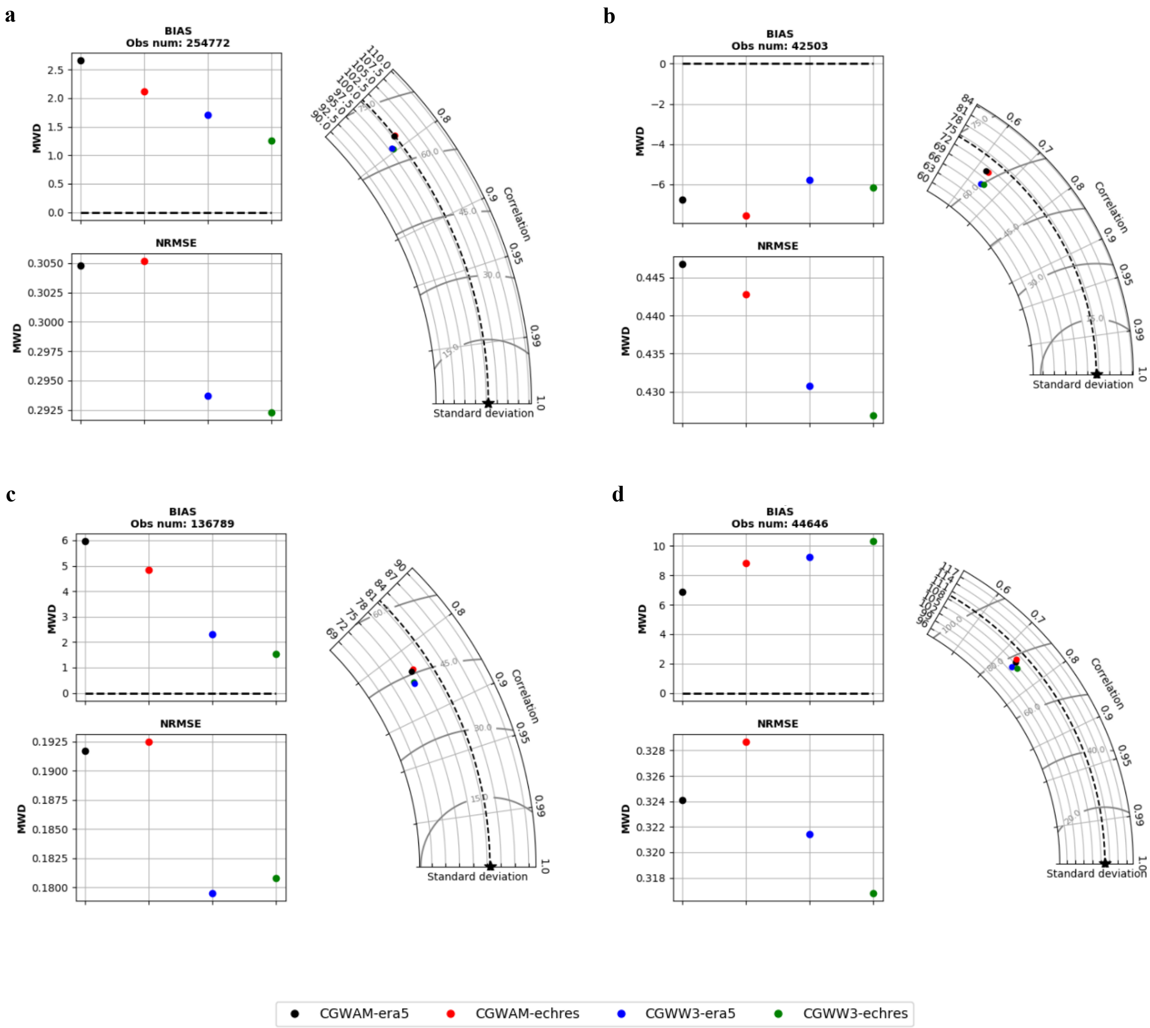

***Figure 8**: AS **Figure 6**, but for the Mean Wave Direction (unit in degree).*

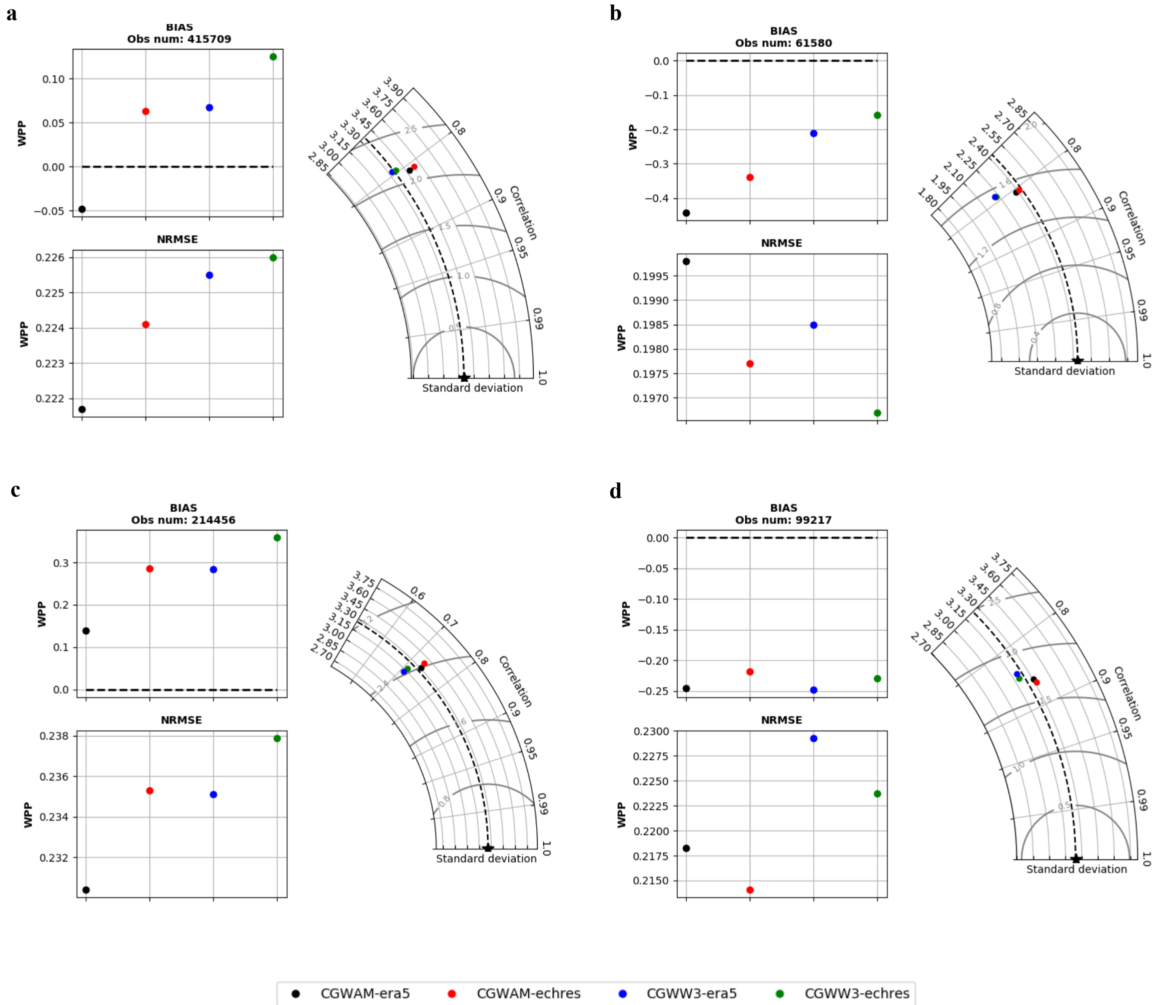

Figure 9: *As **Figure 6**, but for the Wave Peak Period (unit in seconds).*

***Table 4**: Comparison with the UKMO (Saulter et al., 2016) for WPP at global scale.*

Wave Model		**WPP [s] NRMSE**	**WPP [s] Correlation**
UKMO before upgrade		0.29	0.73
UKMO after upgrade		0.24	0.78
CGWW3-echres		0.23	0.80
CGWAM-echres		0.22	0.81
Observed mean and standard deviation for WPP [s]			
	Mean	**Standard Deviation**	
UKMO Sep 14 – Aug 15	8.88	3.17	
This study Jun 17 – Apr 18	8.94	3.35	

5.2 VERIFICATION AGAINST SATELLITE OBSERVATIONS

Statistical assessment based on in-situ observations give a measure of the model's overall performance which, however, is limited by the poor observational spatial coverage. Satellite altimeter data can be used to overcome this issue and, providing an almost complete global coverage, they also allow to observe a larger variety of ocean conditions. In our study, we used the CMEMS dataset (Global Near Real Time Satellite Observations) which offers altimeter observations collected by different satellite platforms such as AltiKa, CryoSat-2, Jason-3, Sentinel-3a and Sentinel-3b. To cover the period of our assessment study, from the CMEMS ftp site we downloaded Jason-3 and Sentinel-3a data which are the measurements available starting from July 2017. The data are organized in NetCDF files which contain the values of the significant wave height that are observed along the satellite track (observations collected in a satellite pass of about 50 minutes with a horizontal resolution of 7 km).

The criteria which was implemented to select satellite data and the equivalent model value is as follows:

a) observations are thinned in time: we filter the data so that 2 consecutive observations are no less than 10 seconds apart. With this choice the distance between 2 measurements is about 60 km.

b) after step a, all the measurements which are acquired within 30 minutes with respect to the model time are kept and associated to that model time (e.g. if model time is 00z, the acceptable time of observations is 23:30z-00:30z).

c) Finally, the latitude and longitude of the observations selected in b are used to calculate the corresponding model value which is considered as the nearest neighbor to the observation location.

Table 5 provides a statistical summary of the observational dataset which have been selected with the criteria described above. The observations were acquired by the satellite altimeter Jason-3 and Sentinel-3a during the period 1st July 2017 and 1st April 2018.

***Table 5**: Statistical summary relative to the satellite observational dataset (Jason-3 and Sentinel-3a) used in the assessment study (1 July 2017 to 1 April 2018).*

Satellite	Dataset size	SWH [m]			
		Min	Max	Mean	Std
Jason-3	497973	0.18	23	2.66	1.39
Sentinel-3a	446758	0.21	22.03	2.57	1.32

An overall summary of the assessment using Jason-3 and Sentinel-3a observations is provided respectively by **Figure 10** and **Figure 11** which show the statistical analysis in terms of bias, NRMSE and the Pearson correlation coefficients for the Significant Wave Height. To explore the skills of the models in different geographical areas, statistics are computed at global scale, but also considering 3 different latitude bands: Northern Hemisphere (30N-89N), Tropics (30S-30N) and Southern Hemisphere (80S-30S). Analyzing **Figure 10** and **Figure 11**, we can state that, although with very slight different magnitude, statistics calculated using Jason-3 and Sentinel-3a observations are very similar. The outcome of this assessment also confirms the results obtained comparing with in-situ observations:

- CGWAM and CGWW3 show a reduction in the bias and NRMSE as well as higher values of correlation coefficients when the atmospheric forcing comes from the ECMWF high resolution system.
- In the tested configurations, CGWW3 looks in general more skillful that CGWAM.
- There is systematic negative bias in SWH.

As previously done in Section 5.1 and presented in **Figure 7**, we investigate the systematic negative bias in the SWH clustering the satellite observational dataset of significant wave height and computing the corresponding model mean bias. **Figure 12** shows the results of this analysis when Janson-3 observations are used (similar results are obtained using Sentinel-3a observations, figure not shown). Findings encapsulated in **Figure 12** confirm what was highlighted by the analysis conducted using in-situ observations: a) CGWAM and CGWW3, independently of the wind forcing, are characterized by a bias which increases with the increasing magnitude of the observed SWH; b) the beneficial effect of the ECMWF high resolution winds which always help to reduce the bias; c) the higher accuracy of CGWW3 respect to CGWAM.

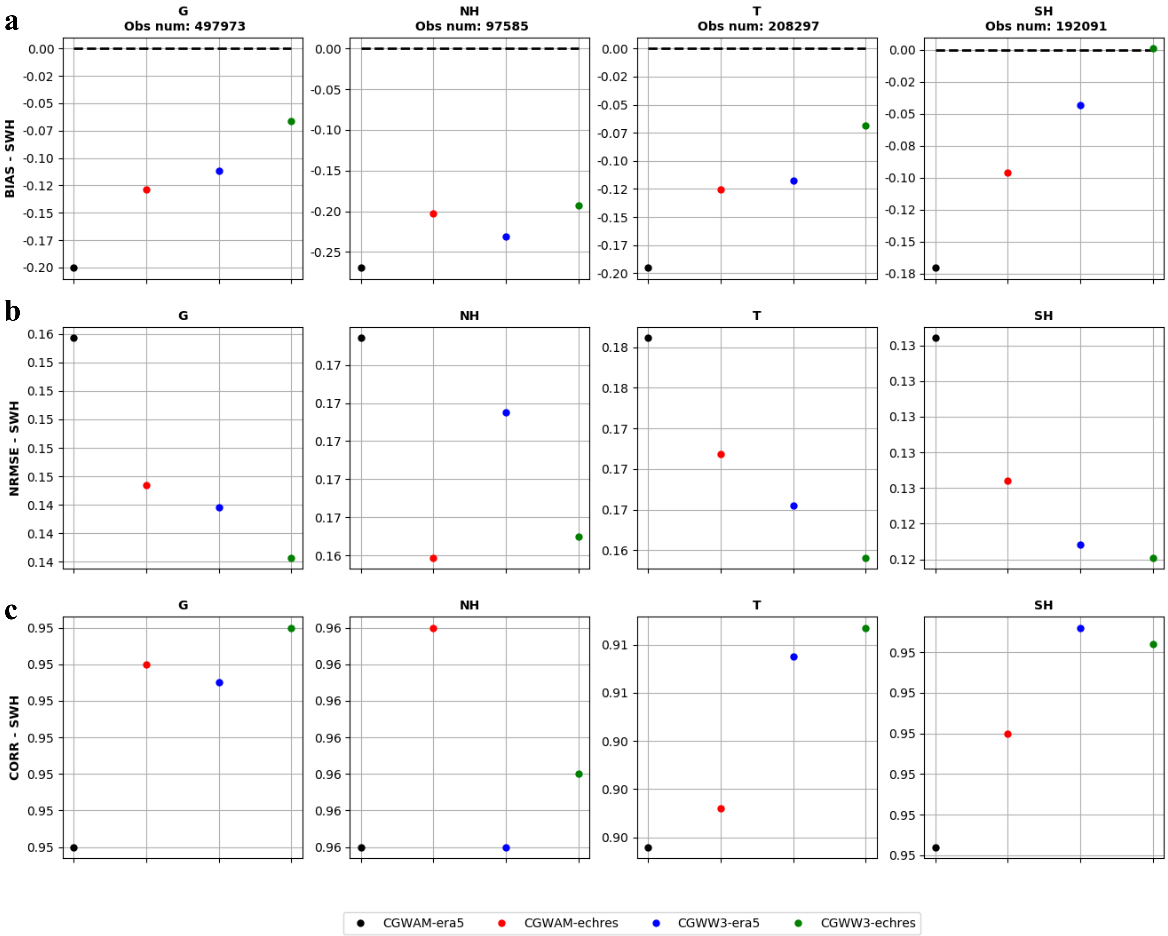

Figure 10: *Statistics for the Significant Wave Height (unit in meters) calculated considering the observations acquired by Jason-3 from 1 July 2017 to 1 April 2018: a) Bias, b) NRMSE and c) Pearson correlation coefficient. In panels a, b and c, statistics are shown for different geographical areas: global scale (G) and for 3 different latitude bands, Northern Hemisphere (30N-89N), Tropics (30S-30N) and Southern Hemisphere (80S-30S). Number of Jason-3 observations used to calculate the statistics is also shown in panel a. CGWAM and CGWW3 experiments are identified by the colored dots as specified in the legend.*

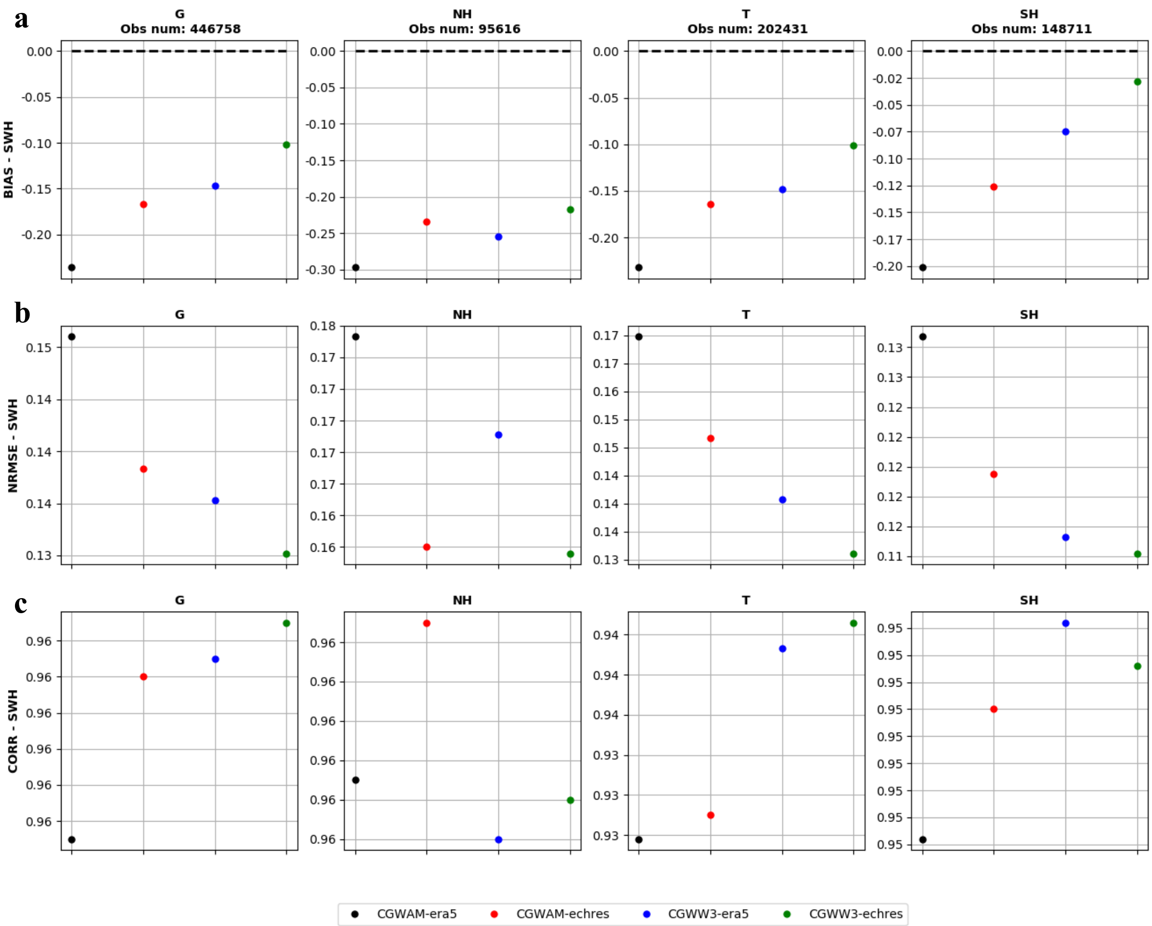

Figure 11: *As* ***Figure 10****, but considering SWH observations from Sentinel-3a.*

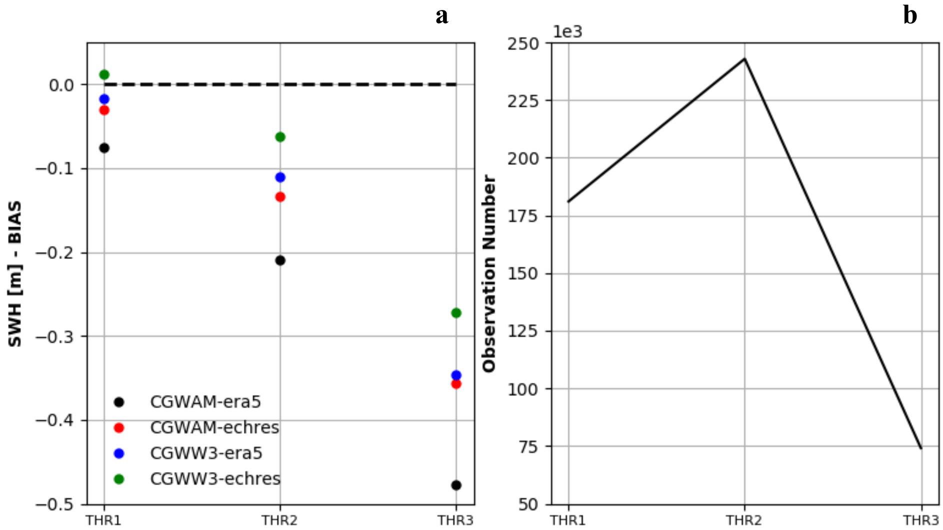

Figure 12: As ***Figure*** *7, but statistics are computed considering observations acquired by Jason-3 from 1 July 2017 to 1 April 2018.*

To evaluate in more details the differences between CGWAM and CGWW3, in **Table 6**, we analyze the percentage difference in NRMSE for the SWH when Janson-3 observations are used to compute the statistics. Results of **Table 6** clearly indicate, for both the wave models, the positive impact associated to the use of high-resolution wind data which, in different magnitude, decrease the error in SWH at global scale and in the different latitude bands. CGWAM appears to largely benefit from the higher horizontal resolution of the atmospheric forcing. This may be an indication of the diverse sensitivity of CGWAM and CGWW3 to the winds which is due to the different parameterizations of the input and dissipation source terms. Nevertheless, CGWW3 shows overall a smaller error and the significance of the formulation of the input and dissipation source terms is confirmed by the reduction in NRMSE which is observed in the Tropics when we compare CGWAM and CGWW3 (CGWAM-era5 with CGWW3-era5 and CGWAM-echres with CGWW3-echres). This agrees with the conclusions stated by Ardhuin et al. (2010) and Rascle and Ardhuin (2013) which observed the largest impact of the Ardhuin's parameterization in the swell-dominated regions such as the Central Tropical Pacific.

Table 6: *Analysis of percentage difference in NRMSE (calculated using Jason-3 observations) for the Significant Wave Height*

Percentage difference in NRMSE (%) is computed as: 100 *($NRMSE_{exp2}$ - $NRMSE_{exp1}$)/$NRMSE_{exp1}$, so that negative/positive values indicate an overall decrease/increase in the error for exp2.				
	Percentage difference in NRMSE (%) for SWH			
exp1/exp2	**G**	**NH**	**T**	**SH**
same wave model, different forcing				
CGWAM-era5/CGWAM-echres	-6.62	-6.56	-7.68	-5.95
CGWW3-era5/CGWW3-echres	-2.43	-3.79	-3.87	-0.57
same forcing, different wave model				
CGWAM-era5/CGWW3-era5	-7.64	-2.22	-11.07	-8.63
CGWAM-echres/CGWW3-echres	-3.51	0.67	-7.39	-3.40

To further exanimate the wave experiments, **Figure 13** presents scatter and frequency distribution plots of observed and modelled Significant Wave Height. In these plots, we consider the Jason-3 observational dataset that we used to compute the statistics at global scale shown in **Figure 10**. Identical results to those of **Figure 13** are obtained using the Sentinel-3a observational dataset (figure not shown). We can once again appreciate the positive impact of using high-resolution wind data which leads to a better fit to the observations and those situations where the magnitude of the observed SWH is largely underestimated by the model.

As final part of this assessment, it is interesting to evaluate the global distribution of the SWH bias. To do that, **Figure 14** and **Figure 15** show latitude and longitude maps where the mean bias is computed considering model and observational values on regular bins of 1 degree. **Figure 14** shows the maps for the statistics calculated using Jason-3 observations, while **Figure 15** is for Sentinel-3a measurements. In every figure, statistics are displayed for the 4 experiments: CGWAM-era5 (panel a), CGWAM-echres (panel b), CGWW3-era5 (panel c) and CGWW3-echres (panel d). For instance, comparing the maps of SWH bias (e.g. **Figure 14a** with **Figure 14b** and **Figure 14c** with **Figure 14d**) we can evaluate the impact of the high-resolution wind data at global scale. It is interesting to observe that, at high latitudes, statistics seem to be unaffected by the change in the wind forcing. Particularly in the Southern Hemisphere, although statistics might be affected by the bin size, validation using different observational dataset generates a similar large positive bias in both CGWAM and CGWW3. This may suggest potential deficiencies in the way the wave modelling works close to the ice edge. To conclude, we can say that **Figure 14** and **Figure 15** also provide a way to globally visualize the skill of wave model outputs as a result of the different setup of CGWAM and CGWW3. As discussed in section 3.1, bathymetry, propagation schemes and the formulation of the input and dissipation source term are the 3 primary sources which can raise differences in the modelled SWH at global scale. It is not feasible to discriminate and quantify the impact associated to each model configuration, but results encapsulated in **Figure 14** and **Figure 15** are consistent with the expected changes due to the Ardhuin’s source term parameterization which expects the main differences in the swell-dominated regions such as the Tropics.

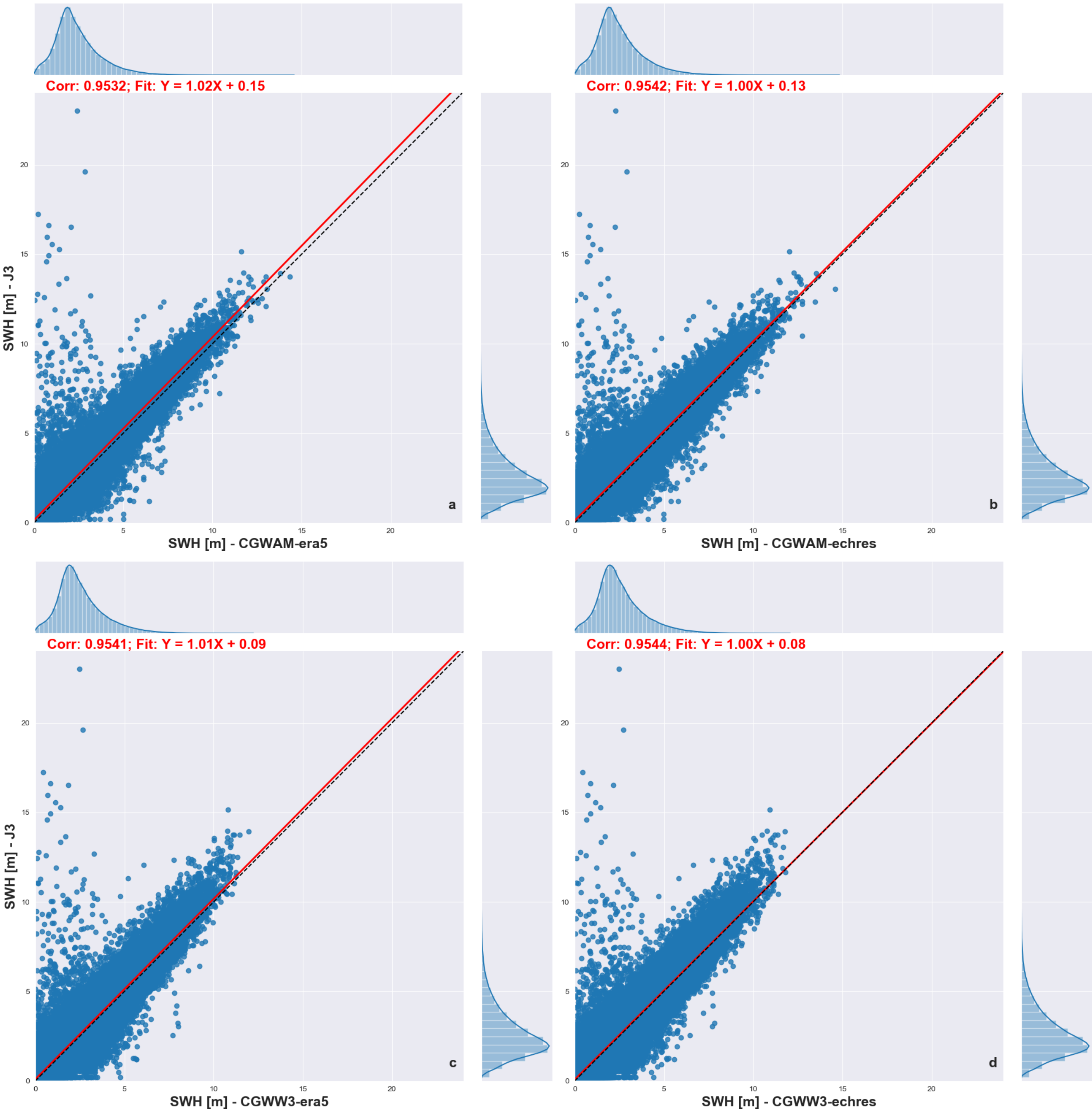

Figure 13: *Scatter and frequency distribution plots of observed (y-axis) and modelled (x-axis) Significant Wave Height (unit in meters). Observations are those acquired by Jason-3 across 1 July 2017 and 1 April 2018 (a total of 497973 observations) and the equivalent model value is displayed as follows: a) CGWAM-era5, b) CGWAM-echres, c) CGWW3-era5 and d) CGWW3-echres. In every panel, the red text indicates the Pearson correlation coefficient together with the result of the data linear fit which is also represented by the red line in the plot. The black dashed line, as a reference, shows the line y = x.*

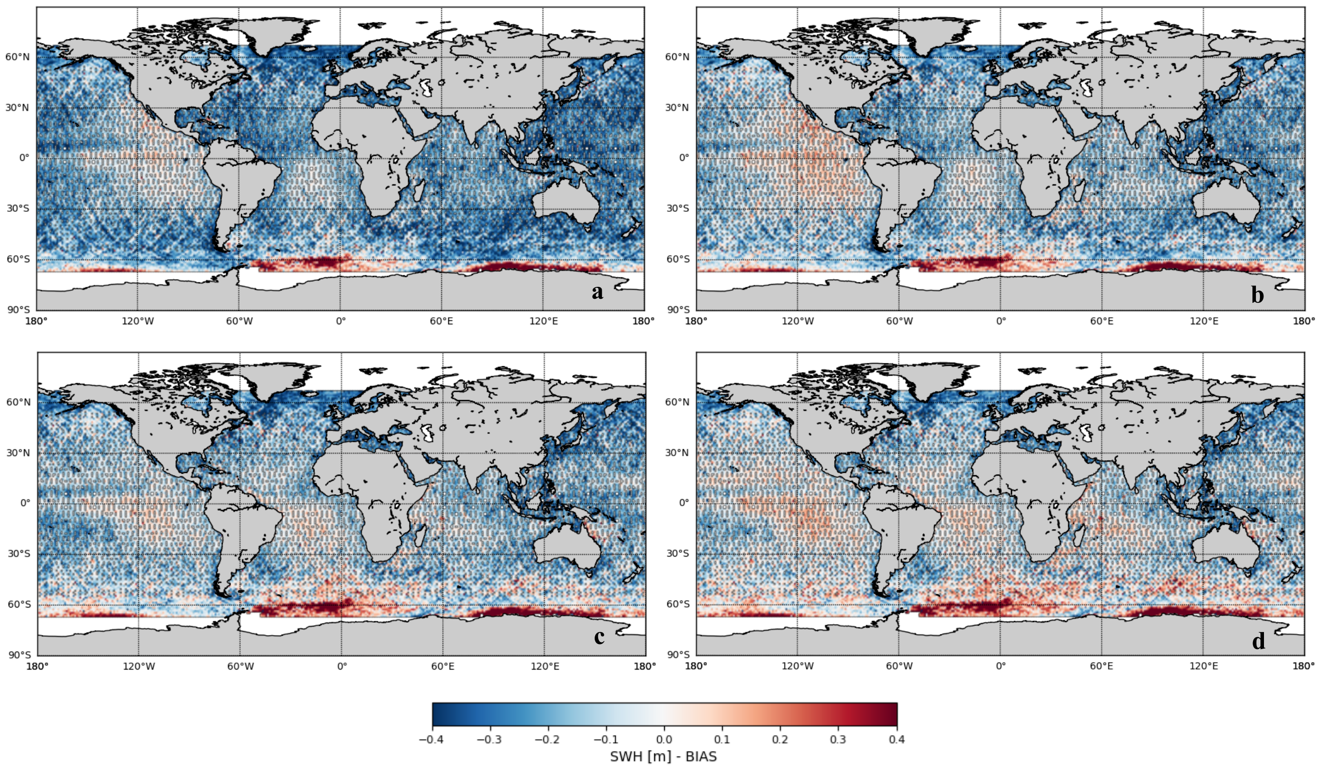

Figure 14: *Latitude-Longitude binned map (bin size is 1 degree) of mean bias for the Significant Wave Height (unit in meters): a) CGWAM-era5, b) CGWAM-echres, c) CGWW3-era5 and d) CGWW3-echres. Bias in every bin is computed considering observations acquired by Jason-3 from 1 July 2017 to 1 April 2018.*

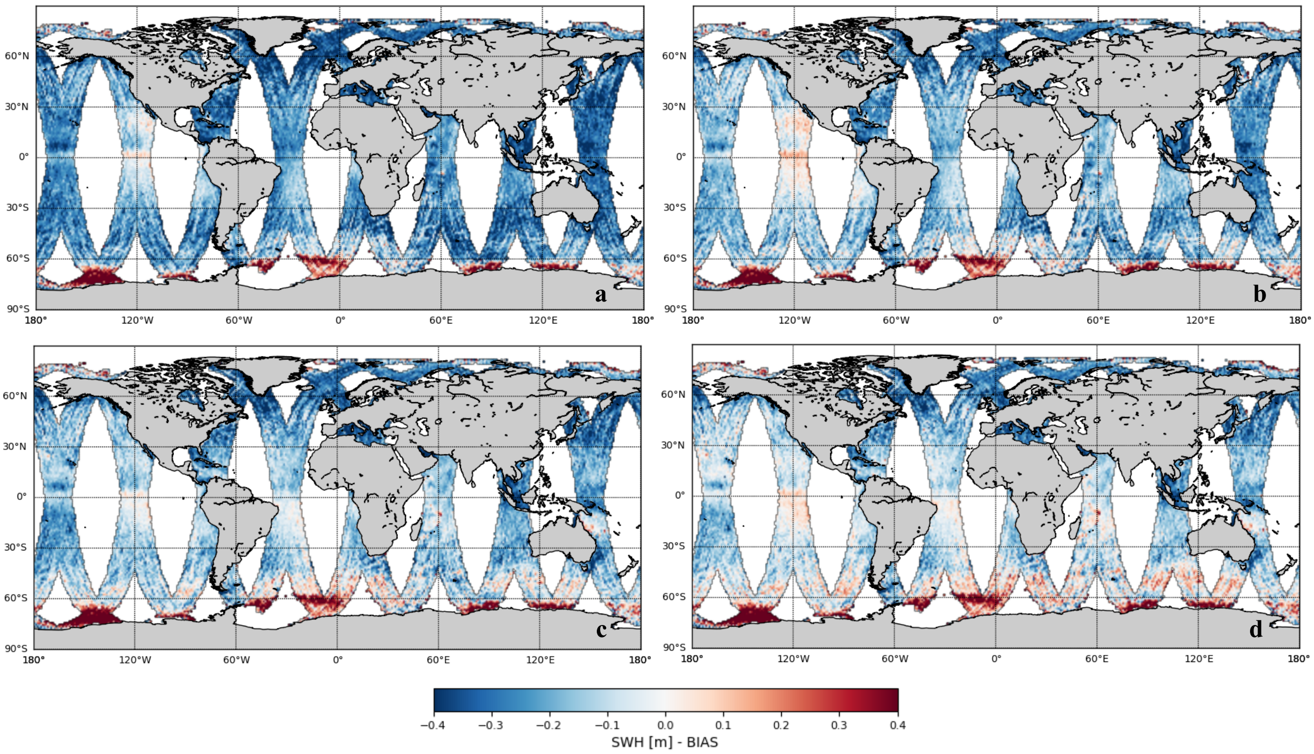

Figure 15: *As* ***Figure 14****, but the mean bias for the SWH is computed considering observations acquired by Sentinel-3a from 1 July 2017 to 1 April 2018.*

6. DISCUSSION AND CONCLUSIONS

We conducted an assessment study to evaluate the performance of two state of the art spectral wave models: the European model WAM (cycle 4.6.2) and the American model WW3 (version 5.16). WAM and WW3 were configured as close as possible and were run at global scale considering the horizontal resolution of 0.25 degree and the spectral resolution of 24 directions and 30 frequencies. We conventionally named the configuration of our wave experiments as CGWAM and CGWW3 (where 'CG' is the acronym to indicate 'CMCC Global'). CGWAM and CGWW3 present some basic differences in the bathymetry, propagation schemes and treatment of ice concentrations, but presumably the significant difference in the configuration is associated with the diverse parameterization of the wind input and swell dissipation source term: GCWAM and CGWW3 are respectively based on the parametrizations provided by Bidlot et al. (2007, 2012) and Ardhuin et al. (2010). Nevertheless, the overall setup for CGWAM and CGWW3 represents the optimal choice for global scale applications: on one side, configuration in CGWAM should be similar to that adopted in the operational wave model ECWAM-cycle 38R1; on the other side, CGWW3 is based on a configuration which was tested and validated in previous studies (e.g. Ardhuin et al., 2010). In our investigation, to evaluate the sensitivity of CGWAM and CGWW3 to the temporal and spatial resolution of the wind forcing, we used two different ECMWF datasets: the ERA5 reanalysis (3-hourly; 0.25 degree horizontal resolution) and the analysis of the operational ECMWF high resolution forecast system (6-hourly; 0.125 degree horizontal resolution). As a result, we run 4 distinct experiments which, according to the wind forcing, were named as: CGWAM-era5, CGWAM-echres, CGWW3-era4 and CGWW3-echres.

Firstly, we point out that the reliability and consistency of the configurations selected for CGWAM and CGWW3 were confirmed by intercomparing the modelled Significant Wave Height (SWH) against the outputs provided by the CMEMS global wave analysis and forecast system (MFWAM) and the ECMWF ERA5 ocean waves reanalysis (ECWAM). Secondly, as benchmark, we used the CMEMS near real time in-situ and satellite observations (Jason-3 and Sentinels-3a) to conduct a robust statistical validation for CGWAM and CGWW3. Although in different degrees, assessment performed using in-situ and satellite measurements provides similar results which can be summarized as follows:

- the spatial resolution and the quality of the input wind data appear essential to improve the accuracy of wave modelling. When the atmospheric forcing is provided by means of the analysis of the operational ECMWF high resolution forecast system, CGWAM (e.g. CGWAM-era5 against CGWAM-echres) and CGWW3 (e.g. CGWW3-era5 against CGWW3-echres) systematically show improvements in the statistics which were calculated to evaluate the model's performance. As a reference, when the validation is conducted by means of in-situ observations, decrease in Normalized Root Mean Square Error (NRMSE) for SWH at global scale is 4.7% and 2% respectively for CGWAM and CGWW3. The reduction is even larger, 9.4% (CGWAM) and 4.7% (CGWW3), if we consider buoys located in the Gulf of Mexico and the Caribbean. The equivalent statistics calculated using Jason-3 measurements confirm this result: NRMSE is globally reduced by 6.6% (CGWAM) and 2.4% (CGWW3) and, in Tropical regions, the decrease of NRMSE reaches 7.7% (CGWAM) and 3.9% (CGWW3).
- In the tested configurations, independently of the atmospheric forcing, CGWW3 looks more skillful than CGWAM. We believe that this outcome is primary due to the different formulation of the input and dissipation source terms. To quantify this benefit, statistics computed by means of Jason-3 observations indicate a reduction in NRMSE for SWH at global scale of 7.6% in the CGWW3-era5 experiment with respect to CGWAM-era5, and 3.5% when we compare CGWW3-echres with CGWAM-echres. In the same

way, the global statistics calculated using in-situ measurements show a decrease in NRMSE of 14.4% in the CGWW3-era5 and 11.9% in CGWW3-echres when they are compared with the equivalent CGWAM experiments.

- The assessment based on both satellite and in-situ observations reveals a systematic negative SWH bias of few centimeters which indistinctly affects CGWAM and CGWW3.

The systematic negative bias was further investigated clustering the observed SWH according to different thresholds and recomputing the corresponding model mean bias. Results of this analysis show that the wave models are characterized by small bias (about 1 cm) where the observed SWH is less than 2 m. However, as the measured SWH increases, so does the model bias and in case of large SWH (bigger than 4 m), the bias can be greater than 56 cm. We can conclude that the primary source of the bias is linked to the largest SWH that the models cannot properly reproduce. More investigations should be conducted to assesses the accuracy of the wind products, but we may speculate that the deficiency in the wave modelling is directly connected to the existing bias in the wind forcing. It is known that under intense storm conditions, the NWP systems may not be able to accurately predict the correct direction and intensity of the winds. In these circumstances, the NWP wind speed is often underestimated, and consequently the modelled SWH results to be lower than observations. Ultimately, the approximation of the non-linear wave-wave interaction may also play an important role in generating systematic errors under storm conditions. The use of different source term parameterizations, which are available in WW3 (e.g. the Generalized Multiple Discrete Interaction Approximation), should also be investigated.

It would be interesting to repeat this assessment when an official release of WAM 4.7 is available, so that it will be possible to configure CGWAM and CGWW3 using the same parameterization of the wind input and swell dissipation source term (Ardhuin et al., 2010). This will give the possibility to verify if the difference in the skills of the wave models found in this study will be reduced. Nevertheless, we can certainly state that to improve the accuracy of a “stand-alone” wave model at global scale, the use of high quality NWP wind data is crucial. To conclude, we should also point out that wave forecasting can benefit from additional modelling features such as the higher horizontal and spectral resolution, the assimilation of wave data and the use of ocean currents as input forcing. As future work, it would be interesting to assess the measure in which these additional characteristics impact on the accuracy of wave modelling at global scale. In a broader picture, we also need to consider the increasing interest in the development of coupled earth system models which include waves. The future efforts in global wave modelling should be focused on the wave–current interactions with the ocean and the wave growth from the atmosphere wind stress which are the processes primarily investigated in recent studies by the scientific community (Ardhuin et al., 2017; Babanin et al., 2012; Chen and Curcic, 2015; Clementi et al., 2017; Øyvind et al., 2015; Staneva et al., 2017; Wahle et al., 2017; among many others).

APPENDIX

A git repository (https://github.com/CMCC-Foundation/WAVE) was created to collect the source code of the wave models (WAM and WW3) together with the configuration files (namelist) and the scripts used to run the experiments. Python diagnostic tools which were developed for this study are also available. A general overview of WAM and WW3 is provided in the README file of the repository.

As a reference, we briefly describe the computational resources which we used to run CGWAM and CGWW3 (configured as described in **Table 1**). In our cluster (Athena) every computational node is made of 16 processors for a total memory of 64G. Table below provides an indication of the computational resource and time to run our experiments.

Run length (start/end date): 15 days	**CGWW3 Regular grid**	**CGWAM Regular grid**	**CGWAM Reduced grid**
Computational recourses	30 nodes * 14 procs (420 CPUs)	25 nodes * 16 procs (400 CPUs)	20 nodes *16 procs (320 CPUs)
Computational time	1h 10m	1h	10m

The model grid adopted in this study is a regular latitude-longitude grid (from 80S to 89N) with step of 0.25° and made of about 976000 ocean grid points. As shown in the table, CGWAM and CGWW3 spend similar computational time to perform a 15-day simulation, but to do that CGWAM requires a smaller number of CPUs. Estimates provided in table depend also on the spectral resolution of the wave models which in our case is made of 24 directions and 30 frequencies. CGWAM also allows to run the wave model using a “reduced grid” (number of longitudinal grid points decreases with increasing latitude). By means of the “reduced grid”, the model takes less time to run with reduced computational resources (points over ocean are about 470000). As a test, we run CGWAM using the reduce grid approach, but with a latitude-longitude step of 0.1 degree: the cost to perform a simulation long 1 day at this higher resolution is estimated as 10 minutes when 80 nodes by 6 processors are employed (a total of 480 cpus). In this respect, WW3 does not have a similar facility to generate a global grid. Probably, in WW3, the best approach to improve the horizontal model resolution at a reasonable computational cost would be the use of the Spherical Multiple-Cell grid (Saulter et al., 2016) which has not been tested in this study. On the contrary, as a test, we tried to run CGWW3 using a Tripolar ORCA grid (0.25 resolution), but, during the execution of the wave model, we experimented random instability without identifying the source of the problem.

ACKNOWLEDGEMENTS

Thanks to Arno Behrens (Helmholtz-Zentrum Geesthacht) for providing the WAM source code and for the helpful guidance with the configuration and to Jean Bidlot (ECMWF) for the supportive discussions about source term parameterizations in WAM and WW3.

BIBLIOGRAPHY

Alari V., Staneva J., Breivik O, Bidlot J.R., Mogensen K. and Janssen PAEM, 2016. Response of water temperature to surface wave effects in the Baltic Sea: simulations with the coupled NEMO-WAM model. Ocean Dynamics.

Ardhuin, F., Rogers, E., Babanin, A., Filipot, J.-F., Magne, R., Roland, A., Van Der Westhuysen, A., Queffeulou, P., Lefevre, J.-M., Aouf, L., Collard, F., 2010. Semi-empirical dissipation source functions for ocean waves: part I, defini- tion, calibration and validation. J. Phys. Oceanogr. 40, 1917–1941.

Ardhuin F., Gille S.T., Menemenlis D., Rocha C.B., Rascle N., Chapron B., Gula J., Molemaker J., 2017. Small-scale open-ocean currents have large effects on wind-wave heights. Journal of Geophysical Research-oceans, 122(6), 4500-4517.

Babanin, A.V., Onorato, M., Qiao, F., 2012. Surface waves and wave-coupled effects in lower atmosphere and upper ocean. J. Geophys. Res. 117, 1–10.

Battjes, J. A. and J. P. F. M. Janssen, 1978: Energy loss and set-up due to breaking of random waves. in Proc. 16th Int. Conf. Coastal Eng., pp. 569–587. ASCE.

Bidlot, J.-R., Janssen, P. and Abdalla, S., 2007. A Revised Formulation of Ocean Wave Dis- sipation and Its Model Impact. Technical Report 509.

Bidlot, J.-R., J.-G. Li, P. Wittmann, M. Faucher, H. Chen, J.-M, Lefevre,T. Bruns, D. Greenslade, F. Ardhuin, N. Kohno, S. Park and M. Gomez, 2007a: Inter-Comparison of Operational Wave Forecasting Systems. In Proc. 10th International Workshop on Wave Hindcasting and Forecasting and Coastal Hazard Symposium, North Shore, Oahu, Hawaii, November 11-16.

Bidlot, J.-R., 2012: Present status of wave forecasting at E.C.M.W.F. in Proceedings of ECMWF workshop on ocean wave forecasting, June.

Cavaleri, L. and Bertotti, L., 2004. Accuracy of the modelled wind and wave fields in enclosed seas, Tellus A, 56, 167–175.

Cavaleri, L., Alves, J.H.G.M., Ardhuin, F., Babanin, a., Banner, M., Belibas- sakis, K., Benoit, M., Donelan, M., Groeneweg, J., Herbers, T.H.C., Hwang, P., Janssen, P.a.E.M., Janssen, T., Lavrenov, I.V., Magne, R., Monbaliu, J., Onorato, M., Polnikov, V., Resio, D., Rogers, W.E., Sheremet, a., McKee Smith, J., Tolman, H.L., van Vledder, G., Wolf, J., Young, I., 2007. Wave modelling - the state of the art. Prog. Oceanogr. 75 (4), 603–674.

Chen, S.S., Curcic, M., 2015. Ocean surface waves in hurricane Ike (2008) and super-storm Sandy (2012): coupled model predictions and observations. Ocean Modelling.

Clementi, Emmanuela, Paolo Oddo, Massimiliano Drudi, Nadia Pinardi, Gerassimos Korres, and Elssandro Grandi, 2017. Coupling hydrodynamic and wave models: first step and sensitivity experiments in the Mediterranean Sea. Ocean Dynam., 67, 1293-1312.

CMEMS Product User Manual, documentation available online at: http://resources.marine.copernicus.eu/documents/PUM/CMEMS-GLO-PUM-001-027.pdf, last access October 2019.

CMEMS Quality Information Document, documentation available online at: http://resources.marine.copernicus.eu/documents/QUID/CMEMS-GLO-QUID-001-027.pdf, last access October 2019.

ECWAM-Cy41R2, documentation available online at: https://www.ecmwf.int/sites/default/files/elibrary/2016/16651-part-vii-ecmwf-wave-model.pdf, last access October 2019

Hasselmann, S., K. Hasselmann, J. H. Allender and T. P. Barnett, 1985. Computations and parameterizations of the nonlinear energy transfer in a gravity-wave spectrum, Part II: Parameterizations of the nonlinear energy transfer for application in wave models. J. Phys. Oceanogr., 15, 1,378–1,391.

Holthuijsen, L.H., Powell, M.D. and Pietrzak J.D., 2012. Wind and waves in extreme hurricanes. Journal of Geophysical Research, 117.

Janssen, P. A. and Bidlot, J.-R., 2018. Progress in operational wave forecasting, Procedia IUTAM, 26, 14–29.

Komen, G. J., Cavaleri, L., and Donelan, M., 1994. Dynamics and modelling of ocean waves, Cambridge university press.

Øyvind Breivik, Kristian Mogensen, Jean-Raymond Bidlot, Magdalena Alonso Balmaseda, Peter A.E.M. Janssen, 2015. Surface Wave Effects in the NEMO Ocean Model: Forced and Coupled Experiments. Journal of Geophysical Research.

Rascle, N., F. Ardhuin, P. Queffeulou, and D. Croiz´e-Fillon, 2008: A global wave parameter database for geophysical applications. part 1: wave-current-turbulence interaction parameters for the open ocean based on traditional parameterizations. Ocean Modelling, 25, 154–171.

Rascle, N. and F. Ardhuin, 2013: A global wave parameter database for geophysical applications. part 2: Model validation with improved source term parameterization. Ocean Modelling, 70, 174-188.

Saulter, A. Bunney, C. and Jian-Guo Li, 2016: Application of a refined grid global model for operational wave forecasting. UK Met Office Forecasting Research Technical Report No: 614.

Staneva, Joanna, Victor Alari, _yvind Breivik, Jean-Raymond Bidlot, and Kristian Mogensen, 2017. Effects of wave-induced forcing on a circulation model of the North Sea. Ocean Dynam., 67, 81-101.

Tolman, H. L., 2002. Alleviating the Garden Sprinkler Effect in wind wave models. Ocean Modelling, 4, 269-289.

Tolman, H. L., 2003. Treatment of unresolved islands and ice in wind wave models. Ocean Modelling, 5, 219–231.

Wahle, Kathrin, Joanna Staneva, Wolfgang Koch, Luciana Fenoglio-Marc, Ha T.M. Ho-Hagemann, and Emil Stanev, 2017. An atmosphere-wave regional coupled model: improving predictions of wave heights and surface waves in the Southern North Sea. Ocean Sci., 13, 289-301.

Wessel, P. and W. Smith, 1996. A global self-consistent hierarchical highresolution shoreline database. J. Geophys. Res., 101(B4), 8741 – 8743.

WAMDI Group, 1988. The WAM model – A third generation ocean wave prediction model, J. Phys. Oceanogr., 18, 1775–1810.

WW3DG: The WAVEWATCH III Development Group, 2016. User Manual and System Documentation of WAVEWATCH III Version 5.16. Tech. Note 329, NOAA/NWS/NCEP/MMAB, College Park, MD, USA.