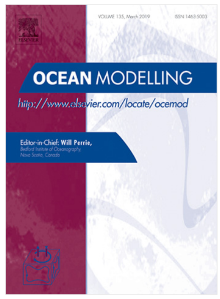
Journal Pre-proof



Performance evaluation of a global CMIP6 single forcing, multi wave model ensemble of wave climate simulations

Gil Lemos, Alvaro Semedo, Rajesh Kumar, Mikhail Dobrynin, Adem Akpinar, Bahareh Kamranzad, Jean Bidlot, Hector Lobeto

PII: S1463-5003(23)00078-1
DOI: https://doi.org/10.1016/j.ocemod.2023.102237
Reference: OCEMOD 102237

To appear in: *Ocean Modelling*

Received date : 15 January 2023
Revised date : 26 June 2023
Accepted date : 27 June 2023

Please cite this article as: G. Lemos, A. Semedo, R. Kumar et al., Performance evaluation of a global CMIP6 single forcing, multi wave model ensemble of wave climate simulations. *Ocean Modelling* (2023), doi: https://doi.org/10.1016/j.ocemod.2023.102237.

Performance Evaluation of a Global CMIP6 Single Forcing, Multi Wave Model Ensemble of Wave Climate Simulations

Gil Lemos[1]*

Alvaro Semedo[2]

Rajesh Kumar[3]

Mikhail Dobrynin[4]

Adem Akpinar[5]

Bahareh Kamranzad[6]

Jean Bidlot[7]

Hector Lobeto[8]

[1] Universidade de Lisboa, Faculdade de Ciências, Instituto Dom Luiz, Lisboa, Portugal

[2] IHE Delft, Department of Coastal, Urban Risk and Resilience, Westvest 7, 2611 AX Delft, The Netherlands

[3] Centre for Climate Research Singapore, 36 Kim Chuan Rd, Singapore 537054, Singapore

[4] Helmholtz-Zentrum Geesthacht Centre for Materials and Coastal Research, Geesthacht, Germany

[5] Department of Civil Engineering, Bursa Uludag University, Bursa, Turkey

[6] Department of Civil and Environmental Engineering, University of Strathclyde, Glasgow, G11XJ, United Kingdom

[7] European Centre for Medium-range Weather Forecasts (ECMWF), Reading RG2 9AX, UK

[8] IHCantabria - Instituto de Hidráulica Ambiental de la Universidad de Cantabria, Santander, Spain

* Corresponding author: grlemos@fc.ul.pt

Keywords: CMIP6, ensemble, wave climate, simulations, evaluation, uncertainty

Abstract

A performance evaluation is conducted for a state-of-the-art Coupled Model Intercomparison Project Phase 6 (CMIP6)-derived ensemble of global wave climate simulations. A single-model (forcing), single-scenario approach is considered to build the ensemble, where the differentiating factor between each member is the wave model or physics parameterization used to simulate waves. The 7-member ensemble is evaluated for the 1995-2014 historical period, highlighting the impact of the multiple source terms on its robustness. The ensemble's ability to accurately represent the present wave climate is assessed through an extensive comparison with long-term ERA5 reanalysis and *in-situ* observational data. Relevant aspects such as the depiction of extremes and natural wave climate variability are analyzed, and inter-member uncertainties are quantified. Overall, the results indicate that the ensemble is able to accurately simulate the global wave climate, regarding the significant wave height (H_S), mean and peak wave periods (T_m and T_p, respectively) and mean wave direction (MWD). However, we show that using multiple wave models and parameterizations should be cautiously considered when building ensembles, even under the same forcing conditions. Model-parameterization-induced ensemble spreads during the historical period are found to be high, compromising the robustness of projected changes in wave parameters towards the end of the 21st century across several areas of the global ocean.

1. Introduction

Ocean surface gravity waves (also commonly named "wind waves") are generated by the action of the wind over the water (Jeffreys, 1924; 1925). These (henceforth just "waves") are almost always present at the sea surface, in the form of seas, swells, or a combination of them, being a clear part of the climate system (Cavaleri *et al.*, 2012; Babanin *et al.*, 2012) and responsible for modulating the exchange of radiation, heat, mass and momentum between the atmosphere and the ocean (Sullivan *et al.*, 2008; Hogstrom *et al.*, 2009, 2011; Semedo *et al.*, 2009; Rutgersson *et al.*, 2010).

Waves play an important role in engineering and environmental issues, as well as in human activities, with direct impacts on coastal dynamics (*e.g.*, Cazenave and Cozannet, 2014; Melet *et al.*, 2018; Shih *et al.*, 1995; Ruggiero *et al.*, 2001), shoreline stability (Harley *et al.*, 2017; Barnard *et al.*, 2015; 2017), coastal flooding and sea level extremes (De Leo *et al.*, 2019; Dietrich *et al.*, 2011; Vitousek *et al.*, 2017; Vousdoukas *et al.*, 2018; Kirezci *et al.*, 2020; Almar *et al.*, 2021), and ship routing and design standards (Bitner-Gregersen *et al.*, 2015; Bitner-Gregersen and Gramstad, 2018). Moreover, waves influence the entire climate system due to their complex feedbacks with the atmosphere, sea ice and the underlying ocean (Cavaleri *et al.*, 2012). For that matter, not only is the monitoring of the present wave climate of paramount importance (Young, 1999; Caires and Swail, 2004; Young *et al.*, 2011; Semedo *et al.*, 2008; 2011; 2014; Aarnes *et al.*, 2012; 2015), but also the accurate projection of global future wave conditions (Morim *et al.*, 2019, 2023; Lobeto *et al.*, 2021a,b).

Sea state observations are required to accurately describe the historical wave climate, but long-term measurements are relatively limited. *In-situ* instruments, such as moored buoys, have been used over the last five decades by many countries as part of their operational observing capabilities. Some of these buoys can currently provide approximately 45 years of (almost) continuous observations (Bidlot *et al.*, 2002). While providing some of the most comprehensive wave datasets, often assumed as "ground truth" (*e.g.*, Bidlot, 2020; Menendez, 2008; Semedo *et al.*, 2014), the most significant disadvantage of *in-situ* observations is, nevertheless, their uneven global positioning, found disproportionately near the coasts of industrialized countries, mainly in the Northern Hemisphere (NH). In the absence of observations, wave modelling efforts like reanalyzes or hindcasts (*e.g.*, ERA5; Hersbach *et al.*, 2020; Bidlot *et al.*, 2019) provide relatively accurate depictions of the global and local wave climates, being currently the only available time- and space-continuous sources of a full spectral description of the ocean surface. Despite the ever-greater accuracy of these modelling products, they rely on forcing winds from atmospheric reanalyzes, which often exhibit well-documented biases and long-term inconsistencies (Ramon *et al.*, 2019; Torralba *et al.*, 2017). In fact, despite the wave's role in the climate system, no fully coupled ocean-wave-atmosphere climate model exists yet, although some attempts have been conducted (*e.g.*, Lionello *et al.* 1998; Rutgersson *et al.* 2010).

Understanding the future evolution of the global wave climate poses one of the greatest challenges in climate modelling. At the same time, it became an important issue for decision and policy-makers in climate change adaptation and mitigation strategies (Magnan *et al.*, 2016; Jones *et al.*, 2014). Future wave climate projections rely on wind and sea ice simulations from global climate models (GCMs), used to force dynamic or statistical wave models (Stopa *et al.*, 2019). Several studies exploring the impact of climate change in future global wave climate have been conducted recently, using forcing GCM outputs from the World Climate Research Program (WCRP) Coupled Model

Intercomparison Project phases 3 (CMIP3) and 5 (CMIP5), namely Mori *et al.* (2010), Dobrynin *et al.* (2012), Fan *et al.* (2013), Hemer *et al.* (2013a), Semedo *et al.* (2013), Wang *et al.* (2015), Dobrynin *et al.* (2015), Erikson *et al.* (2015), Kamranzad *et al.* (2015), Hemer and Trenham (2016), Camus *et al.* (2017), Casas-Prat *et al.* (2018), Kamranzad and Mori (2018), Morim *et al.* (2018, 2019), Kamranzad and Mori (2019) and Lemos *et al.* (2019, 2020a, 2020b, 2021a, 2021b) and Lobeto *et al.* (2021a, 2021b, 2022). While the first studies were based on a single GCM forcing climate simulation (*e.g.*, Mori *et al.*, 2010; Hemer *et al.*, 2013a; Semedo *et al.*, 2013), the use of ensembles has been widely adopted in more recent studies. The primary goal of the ensemble approach is to better quantify the uncertainties associated with individual simulations (Hawkins and Sutton, 2009; Knutti and Sedlacek, 2010; Rauser *et al.*, 2015) for a more realistic depiction of the variability, trends and extremes of past and future projected wave climates. These uncertainties arise from various sources, namely the use of different GCMs, scenarios, wave models, physical parameterizations, the inaccurate depiction of small-scale processes not yet fully understood, or processes not resolved due to computational constrains (Stocker *et al.*, 2013). Cascading uncertainties have often been a limiting factor in climate studies, particularly at regional scales (Foley, 2010; Falloon *et al.*, 2014; Payne *et al.*, 2015).

Most wave climate ensembles rely on a multi-forcing strategy, *i.e.*, different GCMs were used to force dynamical or statistical wave model(s). Recently, Morim *et al.* (2019) compiled the largest set (to date) of individual studies to quantify the uncertainties associated with GCM wind forcing and emission scenarios. It was concluded that uncertainty in current wave climate projections is mostly GCM-driven, in such a way that considering multiple studies at once, robust projected changes in wave parameters (*i.e.*, exceeding the natural historical variability) are only detectable for the RCP8.5 high emissions scenario (Riahi *et al.*, 2011). This study, however, did not investigate to what measure the use of different wave models and parameterizations while generating wave climate projections impacts their uncertainty range and robustness. In fact, this relevant uncertainty source has often been overlooked in the scientific literature (*e.g.*, Erikson *et al.*, 2015; Hemer and Trenham, 2016; Bricheno and Wolf, 2018; Morim *et al.*, 2019; Lemos *et al.*, 2020b). Kumar *et al.* (2022), nevertheless, addressed it, in an attempt to quantify the uncertainties in CMIP6 wave climate projections towards the end of the 21st century using a 4-member ensemble, being the parameterizations (source terms; STs) within the WaveWatchIII (WW3; Tolman *et al.*, 2009; WW3DG, 2019) wave model the differentiating factor. Despite keeping the GCM forcing constant, it was concluded that the uncertainties induced by different STs are enough to seriously affect the robustness of the projections in several areas of the global ocean, even considering a high-emission scenario (SSP5-8.5; O'Neill *et al.*, 2016).

To accurately quantify the impact of climate change, as the differences between the future projected and historical climates, the ability of the ensemble to reproduce the baseline (present) wave climate conditions (mean conditions, intra- and inter-annual variabilities and extremes) must be previously evaluated. The accurate historical climate representation is key to increasing user confidence in the associated future projections. Therefore, a thorough evaluation of the ensemble's performance skills against long-term historical observations or reanalyzes/hindcasts is required (*e.g.*, Semedo *et al.*, 2018b).

In the present study, a unique type of ensemble is presented and evaluated. In our approach, a single CMIP6 GCM (EC-Earth3; Döscher *et al.*, 2022) is used to force seven dynamic wave climate simulations. The differentiating factor between each ensemble

Journal Pre-proof

member (individual simulation) is the wave-model-parameterization pair used to generate the wave climate simulations. In total, three different wave models are used: WW3, SWAN (Booij *et al.*, 1996) and WAM (WAMDI Group, 1988) to produce seven different simulations with multiple STs. The ensemble used here is therefore a "single forcing, multi wave model" one, built to investigate the (usually discarded) impact of multiple parameterizations on the wave climate (both historical and future projected ones). To do so in an effective way, the remaining sources of uncertainty (*e.g.*, adopting a multi-forcing strategies, different initializations, or even multiple future emission scenarios) were limited. Near-surface wind speeds (U_{10}) and sea ice cover (SIC) are used as forcing for the wave models (except SWAN, for which only U_{10} is required), both during the 1995-2014 historical period (henceforth "PC20") and 2081-2100 future projections (not analyzed here). The PC20 ensemble is extensively evaluated through comparison with an extensive *in-situ* observational set (buoys and platforms), and with the European Centre for Medium-range Weather Forecasts (ECMWF) ERA5 reanalysis (Hersbach *et al.*, 2020). Our main goal is not to uniquely present a new ensemble of wave climate simulations and projections, or to focus on an optimal output, but instead to assess the uncertainty generated by an ensemble containing several wave models and parameterizations, as in Morim *et al.* (2018, 2019). For the same reason, although quick progress has been made to improve the overall quality of wave modelling results, we use parameterizations that can be considered outdated, for example, ST1 (Komen *et al.*, 1994), ST2 (Tolman and Chalikov, 1996) and ST3 (Janssen, 2004; Bidlot *et al.*, 2007). We aim to demonstrate to which extent there is a negative impact when pairing older parameterizations with more recent ones in a single ensemble, in terms of uncertainty.

The remainder of the paper is structured as follows. In section 2, the EC-Earth3 GCM, the wave models, the reanalysis and the observational data are described, as well as the general methodology for the evaluation process. In section 3, the performance skills of the PC20 ensemble are assessed in depth, focusing on the representation of means, extremes, short- and long-term variabilities and uncertainties along the historical time-slice. A discussion of the obtained results, together with the concluding remarks, are offered in section 4.

2. Data and methods

2.1. The EC-Earth3 GCM

The EC-Earth is a widely used GCM in both global and regional climate assessments, collaboratively developed by the European Consortium (Döscher *et al.*, 2022). Here, the CMIP6 generation of the model (EC-Earth3) is used, in version 3.3. The EC-Earth3 GCM provides a description of the atmosphere (and its composition), ocean, sea ice, land surface, dynamic vegetation, ocean biogeochemistry and Greenland ice sheet, using the ECMWF Integrated Forecast System (IFS) model cycle CY36R4, coupled with the Nucleus for European Modelling of the Ocean (NEMO) version 3.6, the sea ice model LIM3 and the Pelagic Interactions Scheme for Carbon and Ecosystem Studies (PISCES) biogeochemical model. Terrestrial parameters such as land use, dynamical vegetation and biogeochemistry are given by the Lund-Potsdam-Jena General Ecosystem Simulator (LPJ-GUESS). Additional details are available in Döscher *et al.* (2022).

In the context of CMIP5, the EC-Earth GCM was shown to provide one of the most accurate representations of the historical U_{10} and SIC amongst its remaining counterparts (Shu *et al.*, 2015; Casas-Prat *et al.*, 2018). More recently, within CMIP6, an evaluation for wave climate modelling purposes conducted by Meucci et al. (2023) showed that EC-Earth3 ranks as one of the best GCMs to represent sea level pressure and U_{10} values above the global ocean. Nevertheless, positive U_{10} biases were still identified in the Northern Hemisphere (NH) mid-latitudes, related to an equatorward storm track bias (Harvey et al., 2020; Priestley et al., 2020), and a relatively poor performance for SIC was detected in the Southern Hemisphere (SH).

In this study, one realization of the EC-Earth3 was considered, the r1i1p1f1 ("realization" 1, "initialization" 1, "physics" 1, "forcing" 1) one, to force all the wave climate simulations. This approach eliminates the uncertainty related to different forcings on the GCM side, allowing the isolation of sources related to wave model physics and parameterizations. The spatial domain ranges from 80.36ºS to 80.64ºN and 180ºW to 179.296875ºE in a 0.7º x 0.703125º (latitude x longitude) horizontal resolution grid for all variables (U_{10}, defined by its longitudinal and meridional components uas and vas, and SIC, interpolated from a non-structured grid). The time resolution is 3 hours. The full simulation period corresponds to 1984-2014 and 2070-2100, under the SSP5-8.5 scenario.

2.2. Wave models and parameterizations

2.2.1. WW3

The WW3 is a third-generation spectral wave model vastly used for operational wave forecasting, research, and engineering applications. Here, WW3 version 6.07 (WW3DG, 2019) is used to generate four of the seven global wave climate simulations that compose the ensemble. Within the model, physics and numerical schemes are defined by switches (Tolman, 2009). The switches activated for the four WW3 runs considered here are as follows:

- Third-order Ultimate Quickest (UQ) propagation scheme along with the averaging technique (PR3) for garden sprinkler reduction (Tolman, 2002);

- Discrete interaction approximation (DIA; Hasselmann *et al.*, 1985) for nonlinear wave-wave interactions (switch NL1);
- Linear input (switch LN1) from the parameterization of Cavaleri and Malanotte-Rizzoli (1981), along with a low-frequency filter from Tolman (1992), for consistent spin-up from calm conditions and improving initial wave growth.
- Bottom friction (switch BT1) from the Joint North Sea Wave Project (JONSWAP; Hasselmann *et al.*, 1973);
- Depth-induced wave breaking, accounted using the Battjes and Janssen (1978) formulation (switch DB1);
- Miche-style shallow water limiter (switch MLIM) for maximum wave height;
- Deactivated reflections by shorelines and icebergs (switch REF0) and no bottom or sea ice scattering (switches BS0 and IS0);
- Ice-blocking scheme (switch IC0) considering all grid-points with SIC over 50% as land.

Each of the four WW3 ensemble members correspond to a different input-dissipation parameterization (ST package), namely the ST2 (Tolman and Chalikov, 1996), ST3 (Bidlot *et al.*, 2007 and Janssen, 2004; also named "BJA"), ST4 (Ardhuin *et al.*, 2010) and ST6 (Zeiger *et al.*, 2015, Rogers *et al.*, 2012; Babanin, 2011). Generally, the default parameter settings of each ST package are used. In ST4, coefficients corresponding to the TEST471 option are selected, with β_{max} = 1.43, which generally provides the best results at global scale (WW3DG, 2019). In ST6, switch FLX4 is activated using the air-sea coupling factor CDFAC = 1. It should be highlighted that the ST6 parameterization in WW3 v6.07 suffered a re-calibration, following Rogers et al. (2017) and Liu et al. (2019), updating the U_{10} scaling factor to 32. Additional details can be found in Table SM1 in the Supplementary Material and in Kumar *et al.* (2022).

The bathymetry is based on ETOPO-1 (Amante and Eakins, 2009) and the Global Self-Consistent Hierarchical High-Resolution Shoreline (GSHHS) v1.10 Database. Three files were created: bathymetry, mask, and obstruction grid accounting for wave attenuation by unresolved islands, using the *gridgen* software package (Chawla and Tolman 2007; 2008). The global output time step in WW3 was set to 3 hours, using a spectral resolution of 29 frequencies, logarithmically ranging from 0.0350 Hz to 0.5047 Hz, and 24 directional bins of 15º. The domain and horizontal resolution of the wave fields were kept the same as in the EC-Earth3 forcing winds. Bathymetry, time steps, spectral characteristics, domain and resolution were kept the same for the remaining ensemble members produced using the SWAN and WAM wave models, to limit additional sources of uncertainty. The remaining model configurations were kept constant whenever possible.

2.2.2. SWAN

The Simulating Waves Nearshore (SWAN; Booij *et al.*, 1999; Ris *et al.*, 1999) is a third-generation spectral wave model used for several operational, research and engineering applications. Here, the SWAN version 41.20AB is used to generate two of the seven wave climate simulations. Similarly to the WW3 runs, each of the SWAN members correspond to a different parameterization within the model, namely the ST1, as the recommended SWAN default setting (SWAN Team, 2022) and ST6, with a degree of equivalence to the WW3-ST6 (Donelan *et al.*, 2006; Rogers *et al.*, 2012). In fact, although SWAN is more frequently employed to simulate waves across local to regional domains, it shares most of the physical processes present in other models, as WW3 and WAM (Table SM1). Therefore, SWAN has been considered suitable to simulate waves at global scale (*e.g.*, Mori *et al.*, 2010; Liang *et al.*, 2019; Li and Zhang, 2020). Within the model, run in non-stationary model, general configurations are considered as follows:

- Garden sprinkler effect reduction according to Tolman (2002);
- DIA according to Hasselmann et al. (1985);
- Bottom friction formulation according to the JONSWAP (Hasselmann et al., 1973), but considering C_{fjon} = 0.067 m^2s^{-3} as in Zijlema et al. (2012);
- Depth-induced wave breaking as described in Battjes and Janssen (1978);
- Courant-type limiter, which deactivates quadruplets permanently when the Ursell number exceeds 10 (excluding cases when the fraction of breaking waves exceeds 1 under decreasing action density);
- Third-order upwind scheme according to Stelling and Leendertse (1992) with a diffusive correction for the garden sprinker effect as in Booij and Holthuijsen (1987).

In ST1, the wind and whitecapping formulations follow Komen *et al.* (1994) and Rogers *et al.* (2003). In ST6, some differences to the WW3-ST6 run should be highlighted, namely the inclusion of the new "SSWELL ARDHUIN" option for non-breaking dissipation from Ardhuin et al. (2010) as well as a U_{10} scaling factor of 28 (Hwang, 2011). Additional details can be found in Table SM1.

2.2.3. WAM

The third-generation WAM wave model (WAMDI Group, 1988) version 4.6 is used to produce one of the seven wave climate simulations that compose the ensemble. Specifically, the default WAM settings of physical parametrizations from ECMWF CY45R1 (WAM Cycle 4.6.2.2; ECMWF, 2018) are considered, defined by the switch IPHYS = 0 (overall similar to the WW3-ST3 run), as follows:

- Wind input and wave growth according to Miles (1957) and Janssen (1991);
- DIA according to Hasselmann et al. (1985) and Komen et al. (1994);
- Bottom friction formulation as in Komen et al. (1994);
- Whitecapping dissipation according to Hasselmann (1974) and Janssen (1989a);
- Shallow-water mode.

This simulation, forced by the same 3-hourly EC-Earth3 winds and daily SIC as previous ensemble members, also preserves all remaining setup characteristics, including bathymetry (despite previously converted into a WAM format regular grid), time steps and spectral resolution. Additional details can be found in Table SM1.

2.3. <u>The ERA5 reanalysis</u>

The ERA5 reanalysis provides a comprehensive, high-resolution record of the global atmosphere, land surface, and ocean wind waves from 1950 onwards, continuing to be extended in almost real-time. It is produced using the IFS cycle CY41R2 (ECMWF, 2016), used for the operational forecast from March to November 2016. ERA5 uses an advanced data assimilation system (4D-Var scheme). The horizontal resolution of the atmospheric model in ERA5 is about 30 km (0.25º x 0.25º), being the resolution of the wave parameters approximately 40 km (0.36º x 0.36º). The time resolution is 1 hour. The wave component in ERA5 is produced with a specific configuration of the WAM model named as "ECWAM", based on WAM Cycle 4 (Bidlot, 2007), yet providing a better representation of long-period swells and dissipation levels due to white-capping, as described in Bidlot et al. (2012). Extra output parameters were also introduced to better characterize freak waves, based on the work from Janssen and Bidlot (2009). ERA5 wave spectral domain ranges for 30 logarithmically spaced frequency bins, from 0.03453 Hz to 0.5478 Hz, and 24 directional bins of 15º. The bathymetry in ERA5 is based on the ETOPO2 (NGDC, 2006) dataset. Altimeter wave height wave has been assimilated by the wave model component of the system. Additional details regarding the ERA5 reanalysis can be found in Hersbach *et al.* (2020). Here, the ERA5 is used upon interpolation into the wave climate simulations' grid.

2.4. *In-situ* data

An extensive *in-situ* observational dataset (from buoy and oil platform observations) is used to complement the ensemble performance evaluation. The original ECMWF *in-situ* observational data set, obtained via the WMO Global Telecommunication System (GTS), has regularly been used to evaluate the operational wave forecasts (Bidlot *et al.*, 2002, 2007; Bidlot, 2017), was complemented with *in-situ* wave and wind measurements from Australia, Portugal (mainland and Azores), Baltic Sea and Brazil. The *in-situ* observations from Australia were supplied by Australia's Integrated Marine Observing System (IMOS; enabled by the National Collaborative Research Infrastructure Strategy – NCRIS). The *in-situ* data from Portugal mainland and the Azores were supplied by the Portuguese Hydrographic Institute and by the CLIMAAT (Portuguese acronym, as Clima e Meteorologia dos Arquipélagos Atlânticos) project, respectively. On the other hand, the observations from the Baltic Sea were supplied by the CMEMS (Copernicus Marine Environment Monitoring Service) and the BOOS (Baltic Operational Oceanographic System) online platforms, and the *in-situ* data from Brazil were obtained from the PNBOIA (Portuguese acronym, as Programa Nacional de Bóias).

A quality control assessment was performed for all *in-situ* observations. From the raw dataset, in the first stage, only the *in-situ* instruments with unchanged geographical positioning by more than 1º latitude or longitude from their nominal locations were selected. If this limit was exceeded during a short time (random errors), nevertheless, the observations outside the interval were still considered valid. If the geographical position changed consistently to a different location, observations were still considered valid, yet separately for both locations. All *in-situ* measuring instruments with a reported significant wave height resolution above 0.1 m, a mean or peak wave period resolution above 1 s, and a wind speed resolution above 1 m/s, were automatically excluded. Finally, *in-situ* locations with less than 10 years of measurements or more than 30% of invalid data were removed from the analysis. Upon the selection process, a total of 260 (194) *in-situ* locations remained for the significant wave height (peak wave period) parameter. Their geographical distribution is shown in Fig. 1.

2.5. Methodology

The ensemble in the present study is composed of seven members, being the differentiating factor the wave model and/or the physics parameterization (ST) used to generate each wave climate simulation. All spatial and temporal resolutions between the forcing fields and final outputs are the same. Other inputs, such as bathymetry and land mask, were also preserved between ensemble members, even when considering different wave models. Here, we aimed to restrict the ensemble uncertainty sources (the "degrees of freedom") to represent only the impact of varying wave model architectures and STs.

All the remaining sources, better illustrated in Morim *et al.* (2019), are kept constant to the maximum possible extent.

A set of four wave parameters is analyzed, comprising long-term climate simulations of significant wave height (H_S), mean energy wave period ($T_{m-1,0}$ or simply T_m), peak wave period (T_p) and mean wave direction (MWD). The ensemble mean considers a democratic approach: the unweighted mean of the seven ensemble members (as in Semedo *et al.*, 2018b; Lemos *et al.*, 2019; 2020a; 2020b; 2021a; 2021b; Kumar *et al.*, 2022). For convenience, when referring to individual ensemble members, the notation PC20-i (where i = 1 to 7) is used. The first four members (i = 1 to i = 4) correspond to the WW3 wave climate simulations under ST2, ST3, ST4 and ST6 parameterizations, respectively. The remaining members (i = 5 to i = 7) refer to the SWAN (ST1 and ST6) and WAM simulations, respectively. The 3-hourly wind and wave parameters were processed for both an annual and seasonal (December to February – DJF and June to August – JJA) analysis.

The performance evaluation is carried out at both global and regional scales, considering 13 different sub-areas, chosen according to Alves (2006). These are detailed in Fig. 1 and Table SM2. The evaluation metrics considered here include the Bias (Eq. 1), the normalized bias (NBias; Eq. 2), the root mean squared error (RMSE; Eq. 3), the correlation coefficient (R; Eq. 4), the normalized RMSE, or scatter index (SI; Eq. 5), the slope associated with the linear regression between simulated and reference fields (SL), and the non-dimensional arcsin–Mielke score, or M-score (Watterson, 1996; Watterson *et al.*, 2014; Semedo *et al.*, 2018b; Lemos *et al.*, 2020a; Eq. 6), the mean annual variability index (MAV; Stopa *et al.*, 2014; 2018; Eq. 7) and the inter-annual variability index (IAV; Stopa *et al.*, 2018; Lemos *et al.*, 2019; Eq. 8).

$$\mathrm{Bias} = \overline{\mathrm{PC20}} - \overline{\mathrm{REF}} \quad (1)$$

$$\mathrm{NBias} = \frac{\overline{\mathrm{PC20}} - \overline{\mathrm{REF}}}{\overline{\mathrm{REF}}} * 100\% \quad (2)$$

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{N}(\mathrm{PC20_i} - \mathrm{REF_i})^2}{N}} \quad (3)$$

$$r = \frac{\sum_{i=1}^{N}(\mathrm{REF_i} - \overline{\mathrm{REF}})(\mathrm{PC20_i} - \overline{\mathrm{PC20}})}{\sqrt{\sum_{i=1}^{N}(\mathrm{REF_i} - \overline{\mathrm{REF}})^2}\sqrt{\sum_{i=1}^{N}(\mathrm{PC20_i} - \overline{\mathrm{PC20}})^2}} \quad (4)$$

$$\mathrm{SI} = \frac{\sqrt{\frac{\sum_{i=1}^{N}(\mathrm{PC20}_i - \mathrm{REF}_i)^2}{N}}}{\overline{\mathrm{REF}}} \tag{5}$$

$$\mathrm{M} = \frac{2}{\pi}\arcsin\left(1 - \frac{\mathrm{MSE}}{\mathrm{V_{PC20}} + \mathrm{V_{REF}} + (\mathrm{G_{PC20}} + \mathrm{G_{REF}})^2}\right) * 1000 \tag{6}$$

$$\mathrm{MAV} = \frac{1}{Y}\frac{\sum_{j=1}^{Y}\sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(X_{ji} - \left(\frac{1}{N}\sum_{i=1}^{N}X_{ji}\right)\right)^2}}{\frac{1}{N}\sum_{i=1}^{N}X_{ji}} = \overline{\left(\frac{\sigma_X}{\overline{X}}\right)} \tag{7}$$

$$\mathrm{IAV} = \mathrm{MAV}, \text{ but considering an inter-annual scale (\textit{i.e.}, excluding i)} \tag{8}$$

In Eqs. (1) to (8), PC20 refers to the wave climate simulations, REF to the reference data (ERA5 reanalysis or *in-situ* observations), and X to situations where both are used. Throughout the formulas, N corresponds to the number of outputs considered. In Eq. (6), MSE is the mean squared error, V the spatial variance and G the spatial mean. The M-score performance measure ranges from a hypothetic zero for no skill (MSE = ∞), to a hypothetic maximum score of 1000 (MSE = 0). In Eqs. (3-5) and (7-8), i corresponds to the data index, here as multi-year daily means, computed prior to the evaluation process, in order to avoid the constraints of non-synchrony between the wave climate simulations and the reference data. In Eq. (7), j corresponds to the Julian year's index, being Y the total number of Julian years considered (here set as 20). The MAV (IAV) corresponds to the average of the intra-annual (inter-annual) standard deviation normalized by the yearly (full) mean, providing an indication of the dataset's spread and ability to simulate extremes.

For the comparison between PC20, ERA5 and *in-situ* observations, all multi-year daily averages were collocated through bilinear interpolation to the *in-situ* locations. In the higher latitudes, SIC extent variability can dramatically affect the quality of the mean wave fields, due to a considerable reduction of the available outputs at each grid-point (considered as land when SIC exceeds 50%). Therefore, here, possible inadequate sampling issues at the higher latitudes were dealt with by using one of the approaches proposed by Tuomi *et al.* (2011): grid-points coded as land during 30% or more of the analyzed period were ruled out of the statistics, leaving only the remaining grid-points to be treated as open water.

3. Results

The normalized biases (in %) between the annual, DJF and JJA EC-Earth3 and ERA5 mean U_{10} values are shown in Fig. 2. At an annual scale (corresponding to the entire 1995-2014 period; Fig. 2a), the EC-Earth3 performs better in the extratropical areas than in the tropical, where mostly underestimations are visible, surpassing -20% in the Atlantic region between Brazil and the Gulf of Guinea, and in the tropical North Pacific, south of Hawaii. Local overestimations of up to 20% are visible near the Maritime Continent. In the remaining areas of the global ocean, normalized biases are generally low, below 12%. It should be noted that differences are even lower throughout the Southern Ocean (the single largest global wave generation area), ranging mostly between -4% and 4%. During DJF (Fig. 2b), the patterns are similar to the ones in Fig. 2a, with exacerbated differences in the tropical areas, ranging between -52% and 28% in the Atlantic, -36% and 44% in the Pacific and -20% and 28% in the Indian basins. These are essentially related to the positioning of the Intertropical Convergence Zone (ITCZ) in EC-Earth3, showing a slightly positive latitudinal displacement during DJF, when compared to ERA5. In JJA (Fig. 2c), normalized biases are usually higher, above 4% in most of the global ocean. During this season, while some of the greatest differences are still visible along the tropical areas (mostly negative, down to -36%), positive ones are detectable in the higher latitudes of both hemispheres (up to 36% in the SH and 44% in the NH). Such behavior might be related to a worse representation of the polar vortexes by EC-Earth3 during JJA (Döscher *et al.*, 2022). Nevertheless, Fig. 2 demonstrates that the EC-Earth3 is able to represent the near-surface wind speeds at a global scale with relatively high accuracy. It should be noted that the modelling frameworks of EC-Earth3 and ERA5 are relatively similar (*e.g.*, both using IFS), which could contribute to an enhanced performance against this reanalysis. Overall, EC-Earth3 is considered appropriate to provide forcing to the seven wave climate simulations.

Fig. 3 shows the annual mean (left) and 95% percentile (right) H_S normalized biases between each member of the PC20 ensemble and ERA5. Ensemble members are ordered vertically from PC20-1 to PC20-7. Substantial differences between each member are visible, in both the representation of the annual mean H_S values, and the extremes. Normalized biases range from mostly negative at a global scale (PC20-1 – WW3-ST2 and -6 – SWAN-ST6, in Figs. 3a,b,k,l), to mostly positive (PC20-5 – SWAN-ST1 and -7 – WAM4.6, in Figs. 3i,j,m,n, but also visible for the extreme H_S values for the PC20-2 – WW3-ST3, in Fig. 3d). For the H_S annual mean, the WW3-ST4 (PC20-3) corresponds to the model-parameterization pair that yields overall lower global biases, averaging at 0.98% (Table SM3) and not exceeding 20% (Fig. 3e). For the 95% percentile H_S, PC20-4 (WW3-ST6) shows the best performance, with differences averaging globally at 0.27% (Fig. 3h and Table SM3). Interestingly, the opposite is visible for the SWAN-ST6 pair, averaging at -25.2% and -20.1%, respectively. The impact of model parameterizations in the accuracy of swell propagation is clearly noticeable: for PC20-5 (SWAN-ST1) and -7 (WAM4.6), the highest (positive) normalized biases are found in the tropical latitudes, with lower values in the extratropics, revealing an overestimation of long swell energy content. For PC20-1 (WW3-ST2) and -6 (SWAN-ST6) a similar pattern is observed, however in a global underestimation. Throughout the WW3 simulations (PC20-1 to -4), the agreement between each member and ERA5 tends to increase. This feature highlights the latest efforts in creating ever-more accurate wave model parameterizations, such as the ST4, with an improved swell attenuation scheme (Ardhuin *et al.*, 2010), and ST6, containing both physical and observation-based source terms (Liu *et al.*, 2019; 2021). Note that the widespread H_S underestimation in PC20-1 (WW3-ST2; Fig. 3a,b) is related to a known overestimation of swell dissipation in ST2 which, as a result, underestimates deep-ocean wave growth under stable atmospheric conditions (Tolman, 2002). On the other hand, the overestimations visible for PC20-2 (WW3-ST3; Fig. 3c,d) are mainly related to dissipation constrains depending on swell height, influencing dissipation at the wind-sea peak. Note that, while similar to the WAM4 parameterization (here represented in PC20-7, WAM4.6), the WW3-ST3 run with "BJA" dissipation terms shows a generally better performance than the former, a result which is also described in the WW3 v6.07 manual (WW3DG, 2019).

An optimal balance between the correct description of energy input from the overlaying winds at the wave generation areas, its conversion into swell, and the correct dissipation upon propagation, is not yet obtained, as it is visible in Fig. 3. Even for the WW3-ST4 pair (PC20-3; Figs. 3e,f), with better global performance, slightly positive (negative) biases are visible in the extratropical (tropical) latitudes. Such differences reveal that inaccuracies may still be present in processes such as swell attenuation (wave growth and dissipation due to white-capping) mostly in the low (mid-to-high) latitudes, in comparison with ERA5. Note that ERA5 presented a very reasonable H_S agreement with observations in tropical areas, as it was shown in Bidlot *et al.* (2019). For DJF and JJA mean and 95% percentile H_S (Figs. SM1 and SM2 in the Supplementary Material – SM), despite an expected seasonal shift in the main and extreme patterns from each ensemble member, the overall bias behavior remains similar. Therefore, it is only fair to assume that the main features shown in Fig. 3 are preserved throughout the year.

Fig. 4 is similar to Fig. 3, but for the T_m parameter. As shown for H_S, ensemble member performance varies considerably depending on the model-parameterization pair. Although, in general, a slight overestimation of the mean and extreme (95% percentile) T_m values is visible, mostly below 28%, for PC20-1 (WW3-ST2) and -6 (SWAN-ST6) a widespread underestimation occurs, especially along the subtropics (down to -20%). Among the seven ensemble members, PC20-4 (WW3-ST6) shows the best agreement with ERA5, globally differing, on average, 0.83% (2.11%) considering the mean (95% percentile) T_m (Table SM4). Seasonally, while the normalized bias patterns are similar to the annual ones during DJF (despite slightly higher values in the tropical areas, as visible in Fig. SM3 in the SM), during JJA (Fig. SM4 in the SM), differences are especially relevant in the NH. In fact, during the Austral winter (JJA), the increase in wave storminess in the Southern Ocean (Lobeto *et al.*, 2022) allows for the generation of longer and more energetic swells that deflect to the left (due to the Coriolis force) and propagate northwards, easily surpassing the equator line (Lemos *et al.*, 2021b). Accurate modelling of swell attenuation rates is especially challenging, and therefore, for most ensemble members, normalized biases attain higher values at longer-swell-arriving locations during JJA.

Fig. 5 depicts the annual mean MWD absolute biases (in º) between each PC20 ensemble member and ERA5. MWD means are obtained following the appropriate formula for directional means, *i.e.*, by computing the arctangent of the quotient between components. Each member from PC20-1 to -7 is presented sequentially (Fig. 5a-g). MWD biases are usually higher in high intra-annual variability areas, such as in the subtropics, where a clear influence of the ITCZ positioning is visible, and along near-polar areas, but especially in the NH, possibly due to a more challenging representation of SIC variations. For PC20-5 (SWAN-ST1; Fig. 5e) and -7 (WAM4.6; Fig. 5g), the enhanced swell propagation from the Southern Ocean compared to ERA5 (also revealed by the overestimations found for H_S and T_m in Figs. 3 and 4) is visible through large areas of positive (clockwise) biases, especially in the central Pacific (Table SM5). For the remaining ensemble members, the overall agreement is good, with differences below 36º in most of the global ocean. Along the extratropical latitudes of both hemispheres, biases below 12º are dominant. The seasonal behavior of the MWD biases (in Figs. SM5 and SM6 of the SM for DJF and JJA, respectively) is strongly related to seasonal atmospheric phenomena, such as the main position of the main atmospheric synoptic circulation systems, ITCZ, and the aforementioned swell propagation issues. Higher seasonal biases are dominant in the SH during DJF (except for PC20-5 and -6) and in the NH during JJA. Seasonal biases remain relatively low in most of the global ocean, nevertheless.

Figs. 6 and 7 show the (left) H_S and (right) T_p merged scatter-QQ-plots and the intra-annual cycles, respectively, for the performance evaluation between PC20-1 to -7 and the reference datasets, here exclusively at the *in-situ* locations. It should be noted that until recently, the WMO GTS data only reported H_S and T_p for most *in-situ* locations. Therefore, the majority of the wave period observations in the dataset correspond to T_p instead of T_m. To avoid the effects of the non-synchronized climates between model simulations and reference datasets, multi-year annual means were considered in both figures, and only *in-situ* locations with at least 10 years of continuous observations were selected. The global ocean is divided into areas to evaluate regional performance. Only areas with at least 10 *in-situ* locations available were selected. TWSP and ETSP are shown together to enhance the robustness of the results, given the low number of locations available for ETSP.

Across the ETNA area, 89 (24) locations were selected for H_S (T_p). At these locations, Fig. 6a shows compatible results to those in Fig. 3: while PC20-1 (WW3-ST2) and -6 (SWAN-ST6) show a consistent H_S underestimation, with mean biases of -0.29 m and -0.27 m, respectively, PC20-5 (SWAN-ST1) and -7 (WAM4.6) show the greatest overestimations, with mean biases of 0.41 m and 0.37 m (Table 1). Biases for the 99% percentile range from -0.80 m (PC20-6) to 0.64 m (PC20-7). Such features are noticeable throughout the entire year, as shown in Fig. 7a. Nevertheless, differences attain greater values during the boreal winter season, when the uncertainty range between ensemble members exceeds 1 m. Table 2 shows that, for some members, the performance of extreme H_S is better than the average. In fact, for PC20-2, -3 and -5, biases tend to decrease above the 90% percentile. Nevertheless, the best overall agreement is found for the PC20-3 (WW3-ST4) and -6 members (as in Fig. 3), with mean (extreme) biases of 0.12 m and -0.02 m (0.01 m and -0.09 m). The remaining metrics show relatively similar values for all model-parameterization pairs, with RMSEs, Rs and SIs ranging between 0.46 m (PC20-3) and 0.60 m (PC20-7), 0.87 (PC20-1) and 0.88 (PC20-7) and 0.24 (PC20-3) and 0.32 (PC20-7). In terms of T_p (Fig. 6b), PC20-1 and -6 show consistent underestimations, averaging at -1.01 s and -0.31 s, respectively (Table 3). The greatest overestimations are visible for PC20-5, -7 and -2, at 1.49 s, 0.78 s and 0.47 s on average, respectively. For these members, biases for the 99% percentile onwards surpass 2 s (Table 4). Throughout the year, T_p differences are greater during the boreal summer season, when the inter-member uncertainty range exceeds 3 s (mostly due to the PC20-2 and -5 members; Fig. 7a). Similarly to H_S, PC20-3 and -4 (WW3-ST6) show the best agreement with *in-situ* observations, with mean (extreme) biases of 0.11 s and 0.06 s (below 1.4 s). Overall, RMSEs and SIs range from 0.95 s and 0.12 (PC20-3) to 1.83 s and 0.22 (PC20-5), being the R values generally lower than for H_S.

For the TNAO area, a set of 42 (39) *in-situ* locations were used to locally evaluate H_S (T_p). In Fig. 6c, it can be seen that H_S is mostly underestimated at these locations, with only PC20-5 (SWAN-ST1) and -7 (WAM4.6) showing a consistent overestimation (Table 1), nevertheless, starting from the 30% percentile (Table 3). Note that across TNAO, ERA5 also shows a slight underestimation compared to observations and therefore, while PC20-2 (WW3-ST3) and -3 (WW3-ST4) biases are very close to zero compared to the reanalysis, deviations assume greater values facing the *in-situ* observations (-0.07 m and -0.04 m, respectively; Table 1). However, overall, the bias range (also a proxy to the overall ensemble uncertainty range) is tighter compared to ETNA, from -0.36 m (-0.46 m) to 0.09 m (0.54 m) for the mean (95% percentile) H_S (Tables 1 ad 2). Throughout the year, while PC20-1 (WW3-ST2), -4 (WW3-ST6) and -6 (SWAN-ST6) show consistent underestimations, more evident during the boreal summer (between -0.2 m and -0.3 m), the remaining members' performance varies between extreme seasons, being most differences positive (negative) during the boreal winter (summer), as visible in Fig. 7c. At TNAO, H_S RMSEs, Rs and SIs vary between 0.34 m and 0.50 m, 0.79 and 0.83 and 0.27 and 0.39, respectively. For T_p, Fig. 6d shows a more consistent representation between ensemble members, with mean biases between -0.17 s and 0.24 s, apart from PC20-1 (-0.97 s) and -6 (-1.02 s; Table 3). Deviation patterns are also relatively constant throughout the year (Fig. 7d). RMSEs, Rs and SIs range between 1.02 s and 1.55 s, 0.75 and 0.84 and 0.14 and 0.21.

At ETNP, 84 (79) *in-situ* locations matched the required criteria for H_S (T_p). The performance of individual ensemble members in this area varies considerably, even within each model-parameterization pair. While the H_S performance in Fig. 6e is similar to ETNA's one (Fig. 6a), at ETNP the bias range is greater and inter-member uncertainty is dominated by the SWAN simulations, from -0.38 m (PC20-6; SWAN-ST6) to 0.70 m (PC20-5; SWAN-ST1) for the mean H_S (Table 1). For PC20-5 and -7 (WAM4.6), differences peak between the 10% and the 50% percentiles (Table 2). R coefficients peak for the WW3 simulations, at 0.91, the highest value found for all analyzed areas. RMSEs vary between 0.38 m and 0.78 m and SIs between 0.19 and 0.38. Within the average year, in Fig. 7e, the behavior of the ensemble members is relatively consistent, despite a slight best (worst) performance for PC20-2 and -3 (PC20-5) during the boreal summer. T_p values are generally higher across ETNP than in the remaining areas of the global ocean, partially due to the arrival of long swells generated in the Southern Ocean (Fig. 6f). Nevertheless, most ensemble members reveal a consistent overestimation (except PC20-1; Tables 3 and 4), up to 2.85 s for the mean T_p (PC20-5). R values are slightly lower than for H_S, within 0.68–0.86. RMSEs and SIs range between 1.18 m and 3.37 m, and 0.11 and 0.31, respectively. The mean yearly cycles in Fig. 7f show that, at ETNP, most ensemble members perform worse during summer. In fact, while the mean observed T_p is close to 10 s from May to September, overestimations of up to 4 s (~40%) are visible during this period.

Across the TWSP / ETSP areas, 22 (29) *in-situ* locations were selected for H_S (T_p). Fig. 6g shows that H_S is mostly overestimated by the ensemble members, with only a very slight underestimation (on average) by PC20-1 (WW3-ST2), of -0.02 m (Table 1). This is, in fact, the best performing model-parameterization pair at these areas. The remaining (positive) biases for the mean H_S reach 0.63 m (PC20-7; WAM4.6). For the extreme H_S, nevertheless, both PC20-1 and -6 (SWAN-ST6) show underestimations, down to -0.70 m (Table 2). RMSEs, Rs and SIs range within 0.43–0.67 m, 0.76–0.82 and 0.35–0.57 (highest obtained values), respectively. The generalized H_S overestimation is visible throughout the year, especially for PC20-5 (SWAN-ST1) and -7 (WAM4.6). Ensemble performance (inter-member uncertainty range) is slightly better (lower) during the austral winter (Fig. 7g). In terms of T_p (Fig. 6h), mean biases vary between -1.09 s and 1.90 s for PC20-1 and -5, respectively (Table 3). Extreme differences are usually below 2 s (Table 4). The remaining metrics show T_p performance to be slightly better than the H_S one, with Rs ranging between 0.75 s and 0.88 s and SIs between 0.11 and 0.24. RMSEs vary within 1.26–2.17 s. Similar to H_S, Fig. 7h shows that along the average year, T_p biases and inter-member uncertainty are reduced during the austral winter in TWSP / ETSP.

Fig. 8 displays the average intra-annual H_S cycles for each of the 13 regional areas, considering all grid-points available across each one, for both the PC20 ensemble members, and ERA5. At ETNA, TNAO and ETNP, results are somewhat similar to those of Fig. 7. Yet, at ETNA (Fig. 8a), most members show a slight underestimation *versus* ERA5, and at ETNP (Fig. 8b), the performance of PC20-5 (SWAN-ST1) and -7 (WAM4.6) is considerably better than in Fig. 7e. While the intra-annual H_S cycles are generally well represented by all ensemble members, the agreement with ERA5 is maximized at the extratropical areas of the NH (Figs 8a,b). At ETSA, ETSP and ETSI (Figs. 8c,d,e), PC20-1 (WW3-ST2) and -6 (SWAN-ST6) show a worse, isolated performance, considerably increasing inter-member uncertainty. In fact, in these areas (as well as in TENP and TESP), parameterization-driven uncertainty ranges consistently between 1 m and 1.5 m. Fig. 9 is similar to Fig. 8, but for T_m. For this parameter, performance is more consistent between areas, with an overall underestimation by PC20-1 and -6, and overestimation by PC20-5, -7, and often -2. Inter-member uncertainty varies between 2 s and 3 s. Complementarily, the evolution of the global and regional monthly H_S and T_m means during the historical 1995-2014 period is shown in Figs. SM7 and SM8 in the SM. At a global scale (Figs. SM7n and SM8n), the differences between members and ERA5 are similar to the regionally described in Figs. 8 and 9. No major trends are identifiable during this period for both the PC20 simulations and ERA5.

Fig. 10 presents the H_S MAV (Eq. 7) normalized differences (in %), between each PC20 ensemble member and ERA5. It is noticeable that most model-parameterization pairs tend to overestimate intra-annual variability, especially in the tropical and subtropical areas of the NH (mostly below 18%), potentially due the combined misrepresentation of local tropical phenomena (such as the positioning and strength of the ITCZ; Fig. 2) and the highly seasonal mid-latitude storm belt. An exception is PC20-5 (SWAN-ST1; Fig. 10e) and partially PC20-7 (WAM4.6; Fig. 10g), for which a slight but generalized variability underestimation is visible. PC20 MAVs show a better agreement with ERA5 across the SH, with differences generally below 6%, especially in the Southern Ocean, possibly due to lower seasonal variability resulting from the almost permanent zonal winds. While results are similar for T_m (Fig. SM9 in the SM), MAV differences for this parameter are more circumscribed to the tropical areas.

Fig. 11 is similar to Fig. 10, but for the differences between the H_S IAVs (in %). While most ensemble members showed an overestimation of the MAVs, in this case, slight underestimations are dominant, mainly between -0.5% and -3.5%. Most ensemble members depict areas of positive differences, however, in regions dominated by tropical cyclone activity, namely across the western tropical Pacific and in the Gulf of Mexico. H_S IAVs also tend to diverge in the higher latitudes, possibly due to long-term differences in sea ice area extent between EC-Earth3 and ERA5 (except for the PC20-7 and PC20-6; Figs. 11e,f). Considering T_m (Fig. SM10 in the SM), differences are generally of lower magnitude, however, following similar overall patterns as for H_S.

The boxplots of the ensemble members' H_S and T_m M-scores (Eq. 6), computed for the global ocean and for each of the 13 regional areas, are shown in Fig. 12, considering the annual (grey), DJF (blue) and JJA (red) mean fields. The highest mean H_S M-scores, generally with the lowest uncertainty ranges between model-parameterization pairs, are visible for the extratropical latitudes of both hemispheres, peaking at ETNA (annually and during DJF) and ETSP (during JJA; Fig. 12a). Between members, the highest (lowest) extratropical annual M-score is obtained for the ETNP (ETSP) area at 928 (511). Seasonally, ETNP (ETSI) presents the highest (lowest) M-score, at 895 (509) during DJF, whereas ETSP presents both during JJA, from 464 to 908 (Table SM6). Interestingly, extreme seasonal M-scores are found in the same hemisphere or even in the same area, highlighting the potential differences induced by model-parameterization pairs in the description of the seasonal H_S climate. Overall, the lowest extratropical scores are obtained for the PC20-6 (SWAN-ST6), and the highest occur for the PC20-4/2 (WW3-ST4/ST2) members. Across the tropical areas, H_S M-scores are generally lower, especially at TWSP during JJA, ranging between 119 (PC20-5; SWAN-ST1) and 571 (WW3-ST3). The remaining tropical areas show H_S M-scores between 198 and 952 (Table SM6). For T_m, the regional behavior differs: while lower scores are generally observable for some of the tropical areas (TSAO, TENP and TESP; Fig. 12b), others show performances comparable to the extratropical latitudes (TNAO, TWNP, TWSP, TNIO and TSIO). On the other hand, ETSP and ETSI show only reasonable overall T_m M-scores, mostly between 400 and 700 (Table SM7). At a global scale, nevertheless, both H_S and T_m show a good agreement with ERA5, with M-scores between 713 and 940, and 618 and 911, respectively. Between ensemble members, scores are consistently higher for PC20-2 to 4 (WW3-ST3, -ST4 and -ST6).

4. Discussion and conclusions

In this paper, a performance evaluation was conducted for a 7-member CMIP6 single-forcing, multi-model ensemble of wave climate simulations. The ensemble was built using three different wave models, to investigate the influence of different model-parameterization pairs on the description of the present global wave climate, and on the future projections towards the end of the 21st century (not shown). This uncertainty source is often overlooked in wave climate studies using large, multi-model ensembles, and an accurate quantification of its impacts on the overall ensemble spreads had not yet been conducted. Large uncertainty ranges within ensembles are one major constraint in the correct attribution of future climate change signals (Wallace *et al.*, 2015; Dobrynin *et al.*, 2015). Here, we aimed to characterize the ensemble performance in representing the global and regional wave climates, using the ERA5 reanalysis and an extended, quality-controlled set of *in-situ* observations as references to conduct the analysis. Simultaneously, we focused on the model-parameterization-induced spreads within the ensemble. Note that the 7-member ensemble used in this study contains several parameterizations that could be considered outdated by the present-day wave modelling standards. These were purposedly included to account for the uncertainty generated by an ensemble containing multiple model-parameterization configurations, even outdated ones, as in Morim *et al.* (2018, 2019).

Regarding the forcing EC-Earth3 wind speeds, it was shown in Fig. 2 that the greatest differences are located in the equatorial areas at the annual and seasonal (DJF) scales. During JJA, relatively large areas of mainly positive differences were also shown to be detected in the higher latitudes, especially in the Atlantic and Pacific basins. Overall, although the normalized U_{10} biases were shown to be mostly below 36%, these differences could be responsible for both local and remote misrepresentation of the wave fields on all ensemble members.

The global normalized (Figs. 3 and 4) and absolute (Fig. 5) biases between each ensemble member and ERA5 for the mean H_S, T_m and MWD, and 95% percentile H_S and T_m, *versus* ERA5, revealed considerably distinct patterns for each model-parameterization pair. Overall, the consistently best-performing ensemble members were shown to be PC20-3 (WW3-ST4) and -4 (WW3-ST6). While most members tended to overestimate H_S and T_m at a global scale, especially in the extratropical latitudes, PC20-1 (WW3-ST2) and -6 (SWAN-ST6) showed a consistent opposite behavior. Within the WW3 simulations (PC20-1 to -4), despite the different STs, uncertainty was shown to remain relatively contained. However, the integration of the remaining simulations led to a considerable decrease in the ensemble's robustness. SWAN runs (PC20-5 and -6), in particular, not only showed systematically different behaviors between each other, but also in comparison to other model-parameterization pairs. Despite sharing a similar configuration, SWAN-ST6 (PC20-6) and WW3-ST6 (PC20-4) revealed a distinct representation of the wave climate in Figs. 3 to 9, 12 and Tables 1 to 4, especially in the extratropical areas. As it was shown in Sections 2.2.1. and 2.2.2., as well as in Table SM1, the implementation of the ST6 parameterization in WW3 and SWAN revealed slight dissimilarities (*e.g.*, U_{10} scaling factors and swell dissipation terms) which may have contributed to the distinct representations of the global wave climate. A similar contrast was shown to be visible for the WW3-ST3 (PC20-2) and WAM4.6 (PC20-7) simulations, which produced slightly different global outputs despite their numerical similarities.

Fig. 13 reveals the present climate normalized ensemble inter-member uncertainty range (NUR) considering the full (7-member) ensemble (top), the WW3 subset (middle), and the SWAN subset (bottom), respectively. In the context of climate projections, the NUR represents the minimum ensemble/subset projected change necessary to exceed the present climate ensemble spread. Fig. SM13 is similar to Fig. 13, but for T_m. Both figures show that for the WW3 subset of the ensemble (Fig. 13c,d and SM13c,d), the NUR reaches up to 20% in the extratropical latitudes, and up to 50% (H_S) and 30% (T_m) in the tropical areas. On the other hand, the two SWAN simulations induce spreads within 30%–40% (60%–70%) at the extratropical (tropical) latitudes, for the mean H_S (Fig. 13e), and up to 30% across most of the global ocean for T_m (Fig. SM13e). Considering the full ensemble, the NUR attains values above 70% in the tropical Atlantic, Pacific and Indian basins, for both the mean and 95% percentile H_S (Fig. 13a,b), remaining above 30% in the remaining global ocean. For T_m, these values range between 30% and 40% for the mean and extremes in most locations (Fig. SM13a,b). The seasonal NURs, in Figs. SM11 (DJF) and SM12 (JJA) for H_S, and SM14 (DJF) and SM15 (JJA) for T_m, are consistent with the ones at an annual scale, despite slightly higher values in the respective summer hemisphere. Note that, overall, the NURs found for both the H_S and T_m surpass even the highest emission scenario projections obtained for these parameters towards the end of the 21st century, in recent scientific literature (*e.g.*, Hemer *et al.*, 2013a; Semedo *et al.*, 2013; Wang *et al.*, 2015; Lemos *et al.*, 2020b; albeit for CMIP3 and CMIP5). Ensemble spreads of such magnitudes can lead to serious robustness issues within future projected changes in wave climate. It should be highlighted that a single-forcing EC-Earth3 simulation was used here, and therefore, a multi-forcing approach under similar conditions could potentially lead to even greater NURs.

The comparison between the PC20 ensemble members and *in-situ* observations, in Fig. 6, revealed a reasonable agreement for all model-parameterization pairs across five different regional areas, for both H_S and T_p. Overall, the main behavior of each member was shown to be similar to those represented in Figs. 3 and 4 (for T_m, nevertheless). Biases were shown to generally increase towards the higher quantiles and assume positive values (Tables 2 and 4). Exceptions include PC20-2 (WW3-ST3) and -3 (WW3-ST4) H_S and PC20-1 (WW3-ST2) T_p across ETNA, and most members across ETNP (T_p). For H_S (T_p), the lowest RMSEs and SIs combined with the highest Rs were found for the ETNP (TWSP/ETSP) area, despite the higher mean biases when compared to ETNA and TNAO (Tables 1 and 3). Regarding the mean annual cycles, the PC20 ensemble was shown to be in better agreement with observations for H_S than for T_p, especially across ETNP and TWSP/ETSP. In these areas, both PC20-5 (SWAN-ST1) and -7 (WAM4.6) struggled to depict a correct T_p intra-annual climatology. A similar misrepresentation was visible for PC20-5 across ETNA. It should be noted, however, that in Fig. 9, the mean intra-annual cycles for T_m show a relatively accurate depiction from all ensemble members, despite the consistent biases compared to ERA5.

The H_S M-scores shown in Fig. 12a revealed a better overall agreement between ensemble members and ERA5 across the extratropical areas, with average values ranging between 700 and 900. In the tropical regions, not only was the inter-member M-score range shown to be greater, revealing less consistency in the overall performance, but the mean values were also shown to be lower, mostly between 500 and 800, and down to the 200-400 range for TENP and TESP. These areas were also shown to be the most challenging for T_m, with a mean M-score of approximately 200 for TESP. The highly variable sea state conditions across the eastern Pacific basin, dominated by both the long swells from the Southern Ocean (Lemos *et al.*, 2021b) and local tropical phenomena, contribute to lower modelling performance across TENP and TESP, also noted by Semedo *et al.* (2018a).

Finally, Fig. 14 shows the normalized biases (in %) from the comparison between the democratically built PC20 ensemble H_S, T_m and MWD annual means and extremes (for H_S and T_m), and ERA5, similar to the initially presented in Figs. 3, 4 and 5 for each ensemble member. For the three wave parameters, it is clear that the performance of the PC20 ensemble as a whole is far better than the ones from each model-parameterization pair. In fact, Figs. 14a,b show that for the annual mean (95% percentile) H_S, differences range from -20% to 12% at a global scale, except in the Maritime Continent (higher latitudes of the Southern Ocean – due to undersampling issues caused by sea ice cover), where slightly greater positive differences can be found. Similar normalized biases can be found for T_m, ranging between -12% and 20%, whereas for the MWD differences are only evident at the tropical and subtropical latitudes of the NH (areas dominated by local tropical phenomena). For the three parameters, normalized and absolute biases attain slightly higher values during the extreme seasons (Figs. SM16 and SM17 in the SM), ranging nevertheless between -28% and 20% for H_S, -12% and 28% for T_m and generally below 36º for MWD.

The performance assessment carried out in this study, with specific focus on wave model and physical parameterization uncertainty sources, led to two major conclusions. The first being that all PC20 ensemble members are able to reasonably represent the reference wave climate (both reanalyzed and observed), especially PC20-3 (WW3-ST4) and -4 (WW3-ST6), for which the overall accuracy was shown to be the highest. Finally, as an ensemble, PC20 was shown to perform better than each of its individual members. Secondly, however, despite the increased agreement with observations, changing the wave-model-parameterization combinations within PC20 ensemble members was shown to be enough to produce considerable spreads for the analyzed variables. The impact of this specific uncertainty source in the future wave climate projection ensembles requires further investigation. Nevertheless, it should be highlighted that substantial progress has been recently achieved in improving the global and regional wave climate description by wave models. A dedicated focus on reducing the wave-model-parameterization source of uncertainty in future assessments is paramount for modelling teams, and preference should be given to more recent and balanced parameterizations.

Acknowledgements

This work was funded by the Portuguese Fundação para a Ciência e a Tecnologia (FCT) I.P./MCTES through national funds (PIDDAC) – UIDB/50019/2020.

5. References

Aarnes, O. J., Breivik, Ø., Reistad, M. (2012). Wave extremes in the Northeast Atlantic. J. Clim., 25, 1529– 1543. DOI: 10.1175/JCLI-D-11-00132.1.

Aarnes, O, J., Abdalla, S., Bidlot, J.-R., Breivik, Ø. (2015). Marine Wind and Wave Height Trends at Different ERA-Interim Forecast Ranges. J. Clim, 28, 819–837.

Almar, R., Ranasinghe, R., Bergsma, E. W. J., Diaz, H., Melet, A., Papa, F., Vousdoukas, M., Athanasiou, P., Dada, O., Almeida, L. P., Kestenare, E. (2021). A global analysis of extreme coastal water levels with implications for potential coastal overtopping. Nature Communications, 12, 3775. DOI: https://doi.org/10.1038/s41467-021-24008-9.

Alves, J. H. G. M. (2006). Numerical modelling of ocean swell contributions to the global wind-wave climate. Ocean Modell., 11, 98–122.

Amante, C., Eakins, B. (2009). ETOPO1 1 Arc-minute Global Relief Model: Procedures, Data sources and Analysis. NOAA Technical Memorandum NESDIS (NGDC-24), 1-25.

Ardhuin, F., Rogers, E., Babanin, A. V., Filipot, J.-F., Magne, R., Roland, A., Van der Westhuysen, A., Queffeulou, P., Lefevre, J.-M., Aouf, L., Collard, F. (2010). Semiempirical Dissipation source functions for ocean waves. Part I: definition, calibration, and validation, J. Phys. Oceanogr., 40, 1917–1941. DOI: http://dx.doi.org/ 10.1175/2010JPO4324.1.

Babanin, A. V., Onorato, M., Qiao, F. (2012). Surface waves and wave-coupled effects in lower atmosphere and upper ocean. J. Geophys. Res. https://doi.org/10.1029/2012JC007932. C00J01.

Babanin, A.V. (2011). Breaking and Dissipation of Ocean Surface Waves. Cambridge University Press, 480, pp.

Barnard, P. L., Short, A., Harley, M. D., Splinter, K. D., Vitousek, S., Turner, I. L., Allan, J., Banno, M., Bryan, K. R., Doria, A., Hansen, J. E., Kato, S., Kuriyama, Y., Randall-Goodwin, E., Ruggiero, P., Walter, I. J., Heathfield, D. K. (2015). Coastal vulnerability across the Pacific dominated by El Niño/Southern Oscillation. Nature Geosciences, 8, 801–807.

Barnard, P. L., Hoover, D., Hubbard, D. M., Snyder, A., Ludka, B., Allan, J., Kaminsky, M., Ruggiero, P., Gallien, T. W., Gabel, L., McCandless, D., Weiner, H. M., Cohn, N., Anderson, D., Serafin, K. A. (2017). Extreme oceanographic forcing and coastal response due to the 2015–2016 El Niño. Nature Communications, 8, 14365.Battjes and Janssen (1978)

Bidlot, J.-R., Holmes, D. J., Wittmann, P. A., Lalbeharry, R., Chen, H. S. (2002). Intercomparison of the performance of operational ocean wave forecasting systems with buoy data. Weather Forecasting, 17, 287-310.

Bidlot, J.-R., Janssen, P., Abdalla, S. (2005). A revised formulation for ocean wave dissipation in CY29R1. ECMWF Technical Memorandum, R60.9/JB/0(1), 1-35.

Bidlot J.-R., Janssen, P., Abdalla, S. (2007). A revised formulation of ocean wave dissipation and its model impact. ECMWF Tech. Memo. 509. ECMWF, Reading, United Kingdom, 27pp. Available online at: http://www.ecmwf.int/publications/.

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

Bidlot, J.-R. (2017). Twenty-one years of wave forecast verification. ECMWF newsletter, No 150, 31-36. DOI: 10.18442/ECMWF-NL-201610.

Bidlot, J.-R., Lemos, G., Semedo, A., The C3S Reanalysis Team, ECMWF. (2019, November). ERA5 reanalysis & ERA5 based ocean wave hindcast. Presented at the 2nd International Workshop on Waves, Storm Surges and Coastal Hazards, Melbourne, Australia. Retrieved from http://waveworkshop.org/16thWaves/Presentations/R1%20Wave_Workshop_2019_Bidlot_et_al.pdf

Bitner-Gregersen, E. M., Eide, L. I., Hørte, T., Vanem, E. (2015). Impact of Climate Change and Extreme Waves on Tanker Design. SNAME Transactions 2014, USA.

Bitner-Gregersen, E. M., Gramstad, O. (2018). Potential changes in the joint probabilistic description of the North Atlantic wave climate. Proceedings of the ASME 2018 37th International Conference on Ocean, Offshore and Arctic Engineering - OMAE2018, June 17–22, 2018, Madrid, Spain.

Booij, N., Holthuijsen, L. H., Ris, R. C. (1996). THE "SWAN" WAVE MODEL FOR SHALLOW WATER. Coastal Engineering, 53, 668-676.

Booij, N., Ris, R. C., Holthuijsen, L. H. (1999). A third-generation wave model for coastal regions: 1. Model description and validation. Journal of Geophysical Research Oceans, 104, C4, 7649-7666. DOI: https://doi.org/10.1029/98JC02622.

Bricheno, L. M., Wolf, J. (2018). Future Wave Conditions of Europe, in Response to High-End Climate Change Scenarios. Journal of Geophysical Research: Oceans, 123, 8762–8791. DOI: https://doi.org/10.1029/2018JC013866.

Caires, S., Swail, V. R. (2004). Global wave climate trend and variability analysis. Proceedings of the. Eighth International Workshop on Wave Hindcasting and Forecasting, Oahu, HI, U.S. Army Engineer Research and Development Center's Coastal and Hydraulics Laboratory and Cosponsors, Al. [http://www.waveworkshop.org/8thWaves/Papers/A1.pdf.]

Camus, P., Losada, I, J., Izaguirre, C., Espejo, A., Menéndez, M., Pérez, J. (2017). Statistical wave climate projections for coastal impact assessments. Earth's Future 5, 918-933, doi:doi:10.1002/2017EF000609.

Casas-Prat, M., Wang, X. L., Swart, N. (2018). CMIP5-based global wave climate projections including the entire Arctic Ocean. Ocean Model, 123, 66–85.

Cavaleri, L., Fox-Kemper, B., Hemer, M. A. (2012). Wind waves in the coupled climate system. Bulletin of the American Meteorological Society, 93, 1651-1661.

Cazenave, A., Cozannet, G. L. (2014). Sea level rise and its coastal impacts. Earth's Future, 2, 15-34.

De Leo, F., Besio, G., Zolezzi, G., Bezzi, M. (2019). Coastal vulnerability assessment: through regional to local downscaling of wave characteristics along the Bay of Lalzit (Albania). Natural Hazards Earth Syst. Sci., 19, 287-298. DOI: https://doi.org/10.5194/nhess-19-287-2019.

Dietrich, J., Zijlema, M., Westerink, J., Holthuijsen, L., Dawson, C., Luettich Jr., R., Jensen, R., Smith, J., Stelling, G., Stone, G. (2011). Modeling hurricane waves and storm surge using integrally-coupled, scalable computations, Coast. Eng., 58, 45–65.

Dobrynin, M., Murawsky, J., Yang, S. (2012). Evolution of the global wind wave climate in CMIP5 expriments. Geophysical Research Letters, 39, DOI:10.1029/2012GL052843.

Dobrynin, M., Murawski, J., Baehr, J., Ilyina, T. (2015). Detection and Attribution of Climate Change Signal in Ocean Wind Waves. Journal of Climate, 28, 4, 1578-1591.

Döscher, R., Acosta, M., Alessandri, A., Anthoni, P., Arneth, A., Arsouze, T., Bergmann, T., Bernadello, R., Bousetta, S., Caron, L.-P., Carver, G., Castrillo, M., Catalano, F., Cvijanovic, I., Davini, P., Dekker, E., Doblas-Reyes, F. J., Docquier, D., Echevarria, P., Fladrich, U., Fuentes-Franco, R., Gröger, M., v. Hardenberg, J., Hieronymus, J., Karami, M. P., Keskinen, J.-P., Koenigk, T., Makkonen, R., Massonnet, F., Ménégoz, M., Miller, P. A., Moreno-Chamarro, E., Nieradzik, L., van Noije, T., Nolan, P., O'Donnell, D., Ollinaho, P., van den Oord, G., Ortega, P., Prims, O. T., Ramos, A., Reerink, T., Rousset, C., Ruprich-Robert, Y., Le Sager, P., Schmith, T., Schrödner, R., Serva, F., Sicardi, V., Sloth Madsen, M., Smith, B., Tian, T., Tourigny, E., Uotila, P., Vancoppenolle, M., Wang, S., Wårlind, D., Willén, U., Wyser, K., Yang, S., Yepes-Arbós, X., Zhang, Q. (2022). The EC-Earth3 Earth System Model for the Climate Model Intercomparison Project 6. Geosci. Model Dev., 15, 2973–3020. DOI: https://doi.org/10.5194/gmd-15-2973-2022.

ECMWF. (2016). IFS Documentation CY41R2. https://www.ecmwf.int/node/16651.

ECMWF. (2018). IFS Documentation CY45R1 - Part VII : ECMWF Wave Model. https://www.ecmwf.int/node/18717.

Erikson, L. H., Hegermiller, C. A., Barnard, P. L., Ruggiero, P., van Ormondt, M. (2015). Projected wave conditions in Eastern North Pacific under the influence of two CMIP5 climate scenarios. Ocean Modelling, 96, 1, 171-185.

Falloon, P., Challinor, A., Dessai, S., Hoang, L., Johnson, J., Koehler, A.-K. (2014). Ensembles and uncertainty in climate change impacts. Frontiers in Environmental Science, 2, 33. doi: 10.3389/fenvs.2014.00033.

Fan, Y., Held, I. M., Lin, S.-J., Wang, X. L. (2013). Ocean warming effect on surface gravity wave climate change for the end of the 21st century. Journal of Climate, 26, 6046-6066.

Foley, A. (2010). Uncertainty in Regional Climate Modelling: A Review. Progress in Physical Geography, 34(4). doi:10.1177/0309133310375654.

Harley, M. D., Turner, I. L., Kinsela, M. A., Middleton, J. H., Mumford, P. J., Splinter, K. D., Phillips, M. S., Simmons, J. A., Hanslow, D. J., Short, A. (2017). Extreme coastal erosion enhanced by anomalous extratropical storm wave direction. Scientific Reports, 7, 6033.

Hasselmann, K., Barnett, T. P., Bouws, E., Carlson, H., Cartwright, D. E., Enke, K., Ewing, J. A., Gienapp, H., Hasselmann, D. E., Kruseman, P., Meerburg, A., Müller, P., Olbers, D. J., Richter, K., Sell, W., Walden, H. (1973). Measurements of wind-wave growth and swell decay during the Joint North Sea Wave Project (JONSWAP), Deutsche Hydrographische Zeitschrift., Suppl. A, 8(12), 95p.

Hasselmann, S., Hasselmann, K., Allender, J.H., Barnett, T.P. (1985). Computations and parameterizations of the nonlinear energy transfer in a gravity-wave specturm. Part II: Parameterizations of the nonlinear energytransfer for application in wave models. J. Phys. Oceanogr. 15, 1378–1391.

Hawkins, E., Sutton, R. (2009). The potential to narrow uncertainty in regional climate predictions. Bulletin of the American Meteorological Society, 90, 1095–1107.

Hemer, M. A., Katzfey, J., Trenham, C. E. (2013a). Global dynamical projections of surface ocean wave climate for a future high greenhouse gas emission scenario. Ocean Modelling, 70, 221-245.

Hemer, M. A., Trenham, C. E. (2016). Evaluation of a CMIP5 derived dynamical global wind wave climate model ensemble. Ocean Modelling, 103, DOI: 10.1016/j.ocemod.2015.10.009.IPCC-AR5. Climate Change 2014: The Physical Science Basis; Cambridge University Press: Cambridge, UK, 2014.

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., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R., 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. (2020). The ERA5 global reanalysis. Quarterly Journal of the Royal Meteorological Society, 2020, 1-51.

Högström, U., Smedman, A., Sahleé, E., Drennan, W. M., Kahma, K. K., Pettersson, H., Zhang, F. (2009). The atmospheric boundary layer during swell: A field study and interpretation of the turbulent kinetic energy budget for high wave ages. Journal of Atmospheric Science, 66, 2764-2779.

Högström, U., Smedman, A.-S., Semedo, A., Rutgersson, A. (2011). Comments on “A global climatology of wind wave interaction”. Journal of Physical Oceanography, 41, 1811–1813.

Janssen, P. A. E. M. (2004). The Interaction of Ocean Waves and Wind. Cambridge: Cambridge University Press.

Jeffreys, H. (1924). On the formation of waves by wind. Proceedings of the Royal Society, A107, 189-206.

Jeffreys, H. (1925). On the formation of waves by wind II. Proceedings of the Royal Society, A110, 341–347.

Jones, R. N. et al. in Climate Change 2014: Impacts, Adaptation, and Vulnerability (eds Field, C. B. et al.) Ch. 2 (IPCC, Cambridge Univ. Press, 2014).

Kamranzad, B., Etemad-Shahidi, A., Chegini, V., Yeganeh-Bakhtiary, A. (2015). Climate change impact on wave energy in the Persian Gulf. Climate Dynamics, 65, 777-794. DOI: https://doi.org/10.1007/s10236-015-0833-y.

Kamranzad, B., Mori, N. (2018). Regional Wave Climate Projection Based on Super-High-Resolution MRI-AGCM3.2S, Indian Ocean. Journal of Japan Society of Civil Engineers, 74, 2, 1351-1355. DOI: https://doi.org/10.2208/kaigan.74.I_1351.

Kamranzad, B., Mori, N. (2019). Future wind and wave climate projections in the Indian Ocean based on a super-high-resolution MRI-AGCM3.2S model projection. Climate Dynamics, 53, 2391-2410. DOI: https://doi.org/10.1007/s00382-019-04861-7.

Kirezci, E., Young, I. R., Ranasinghe, R., Muis, S., Nicholls, R. J., Lincke, D., Hinkel, J. (2020). Projections of global-scale extreme sea levels and resulting episodic coastal flooding over the 21st Century. Scientific Reports, 10, 11629. DOI: https://doi.org/10.1038/s41598-020-67736-6.

Knutti, R., Sedláček, J. (2012). Robustness and uncertainties in the new CMIP5 climate model projections. Nature Climate Change, 3, 369–373.

Komen, G. J., Cavalieri, L., Doneland, M., Hasselmann, S., Janssen, P. A. E. M. (1994). Dynamics and modeling of ocean waves. Cambridge: Cambridge University Press.

Kumar, R., Lemos, G., Semedo, A., Alsaaq, F. (2022). Parameterization-Driven Uncertainties in Single-Forcing, Single-Model Wave Climate Projections from a CMIP6-Derived Dynamic Ensemble. Climate, 10, 51. DOI: https://doi.org/10.3390/cli10040051.

Lemos, G., Semedo, A., Dobrynin, M., Behrens, A., Staneva, J., Bidlot, J.-R., Miranda, P. (2019). Mid-twenty-first century global wave climate projections: Results from a dynamic CMIP5 based ensemble. Global and Planetary Change, 172, 69-87.

Lemos, G., Menendez, M., Semedo, A., Camus, P., Hemer, M., Dobrynin, M., Miranda, P. (2020a). On the need of bias correction methods for wave climate projections. Global and Planetary Change, 186, 103109.

Lemos, G., Semedo, A., Dobrynin, M., Menendez, M., Miranda, P. (2020b). Bias-Corrected CMIP5-Derived Single-Forcing Future Wind-Wave Climate Projections toward the End of the Twenty-First Century. Journal of Applied Meteorology and Climatology, 59, 9, 1393–1414. http://dx.doi.org/10.1175/jamc-d-19-0297.1

Lemos, G., Semedo, A., Menendez, M., Miranda, P. M. A., Hemer, M. (2021a). On the decreases in North Atlantic significant wave heights from climate projections. Climate Dynamics. DOI: https://doi.org/10.1007/s00382-021-05807-8.

Lemos, G., Semedo, A., Hemer, M., Menendez, M., Miranda, P. M. A.. (2021b). Remote climate change propagation across the oceans – the directional swell signature. Environmental Research Letters. DOI: https://doi.org/10.1088/1748-9326/ac046b.

Liang, B., Gao, H., Shao, Z. (2019). Characteristics of global waves based on the third-generation wave model SWAN. Marine Structures, 64, 35-53.

Lionello, P., Malguzzi, P., Buzzi, A. (1998). Coupling between the Atmospheric Circulation and the Ocean Wave Field: An Idealized Case. Journal of Physical Oceanography, 21, 161-177.

Lobeto, H., Menendez, M., Losada, I. (2021a). Future behavior of wind wave extremes due to climate change. Scientific Reports, 11, 7869. DOI: 10.1038/s41598-021-86524-4.

Lobeto, H., Menendez, M., Losada, I. (2021b). Projections of Directional Spectra Help to Unravel the Future Behavior of Wind Waves. Frontiers in Marine Science, 8, 558. DOI: DOI: 10.3389/fmars.2021.655490.

Lobeto, H., Menendez, M., Losada, I., Hemer, M. (2022). The effect of climate change on wind-wave directional spectra. Global and Planetary Change, 213(4), 103820.

Magnan, A. K., Schipper, E. L. F., Burkett., M., Bharwani, S., Burton, I., Eriksen, S., Gemenne, F., Schaar, J., Ziervogel, G. (2016). Addressing the risk of maladaptation to climate change. WIREs Climate Change, 7, 646–665.

Melet, A., Meyssignac, B., Almar, R., Le Cozannet, G. (2018). Under-estimated wave contribution to coastal sea-level rise. Nature Climate Change, 8, 234–239.

Menendez, M., Mendez, F., Losada, I., Graham, N. (2008). Variability of extreme wave heights in the northeast Pacific Ocean based on buoy measurements. Geophysical Research Letters, 35. DOI: 10.1029/2008GL035394.

Meucci, A., Young, I. R., Hemer, M., Trenham, C., Watterson, I. G. (2023). 140 years of global ocean wind-wave climate derived from CMIP6 ACCESS-CM2 and EC-Earth3 GCMs: Global trends, regional changes, and future projections. Journal of Climate, 36(6), 1605-1631.

Mori, N., Yasuda, T., Mase, H., Tom, T., Oku, Y. (2010). Projections of extreme wave climate change under global warming. Hydrological Research Letters, 4, 15-19.

Morim, J., Hemer, M. A., Cartwright, N., Strauss, D., Andutta, F. (2018). On the concordance of 21st century wind-wave climate projections. Global and Planetary Change, 167, 160-171.

Morim, J., Hemer, M. A., Wang, X., Cartwright, N., Trenham, C., Semedo, A., Young, I., Bricheno, L., Camus, P., Casas-Prat, M., Erikson, L., Mentaschi, L., Mori, N., Shimura, T., Timmermans, B., Aarnes, O., Breivik, Ø., Behrens, A., Dobrynin, M., Menendez, M., Staneva, J., Wehner, M., Wolf, J., Kamranzad, B., Webb, A., Stopa, J., Andutta, F. (2019). Robustness and uncertainties in global multivariate wind-wave climate projections. Nature Climate Change.

Morim, J., Wahl, T., Vitousek, S., Santamaria-Aguilar, S., Young, I., Hemer, M. (2023). Understanding uncertainties in contemporary and future extreme wave events for broad-scale impact and adaptation planning. Science Advances, 9(2). DOI: 10.1126/sciadv.ade3170.

NGDC. (2006). 2-minute Gridded Global Relief Data (ETOPO2) v2. NOAA National Centers for Environmental Information. DOI: https://doi.org/10.7289/V5J1012Q.

O'Neill, B. C. O., Tebaldi, V., van Vuuren, D. P., Eyring, V., Friedlingstein, P., Hurtt, G., Knutti, R., Kriegler, E., Lamarque, J.-F., Lowe, J., Meehl, G. A., Moss, R., Riahi, K., Sanderson, B. M. (2016). The Scenario Model Intercomparison Project (ScenarioMIP) for CMIP6. Geosci. Model Dev., 9, 3461-3482.

Payne, M. R., Barange, M., Cheung, W., MacKenzie, B., Batchelder, H., Cormon, X., Eddy, T., Fernandes, J., Hollowed, A., Jones, M., Link, J., Neubauer, P., Ortiz, I., Queirós, A., Paula, J. (2015). Uncertainties in projecting climate-change impacts in marine ecosystems. ICES Journal of Marine Science, 73(5), 1272–1282, https://doi.org/10.1093/icesjms/fsv231.

Ramon, J., Lledó, L., Torralba, V., Soret, A., Doblas-Reyes, F. J. (2019). What global reanalysis best represents near-surface winds? Quarterly Journal of the Royal Meteorological Society, 145, 3236-3251.

Rauser, F., Schemann, V., Sonntag, S. (2015). Sustainable early-career networks. Nat. Geosci., 8, 745–746.

Riahi, K., Rao, S., Krey, V., Cho, C., Chirkov, V., Fischer, G., Kindermann, G., Nakicenovic, N., Rafaj, P. (2011). RCP 8.5 – a scenario of comparatively high greenhouse gas emissions. Climatic Change, 109, 33-57, DOI: 10.1007/s10584-011-0149-y.

Rogers, W. E., Hwang, P. A., Wang, D. W. (2003). Investigation of wave growth and decaying the SWAN model: three regional-scale applications. J. Phys. Oceanogr., 33, 366–389.

Rogers, W. E., Babanin, A. V., Wang, D. W. (2012). Observation-Consistent Input and White capping Dissipation in a Model for Wind-Generated Surface Waves: Description and Simple Calculations. Journal of Atmospheric and Oceanic Technology, 29 (9), 1329–1346, doi: 10.1175/JTECH-D-11-00092.1.

Ruggiero, P., Komar, P. D., McDougal, W. G., Marra, J. J., Beach, R. A. (2001). Wave runup, extreme water levels and the erosion of properties backing beaches. J. Coast. Res. 2001, 17, 407–419.

Rutgersson, A., Saetra, Ø., Semedo, A., Carlson, B., Kumar, R. (2010). Impact of surface waves in a Regional Climate Model. Meteorol. Z., 19, 3, 247–257.

Semedo, A., Sušelj, K., Rutgersson, A. (2008). Variability of Wind Sea and Swell Waves in the North Atlantic Based on ERA-40 Reanalysis. Proceedings of the 8th European Wave and Tidal Energy Conference, September 7-10, 2008, Uppsala, Sweden.

Semedo, A., Sætra, Ø., Rutgersson, A., Kahma, K. K., Pettersson, H. (2009). Wave induced wind in the marine boundary layer. Journal of Atmospheric Science, 66, 2256-2271.

Semedo, A. (2010). Atmosphere-ocean Interactions in Swell Dominated Wave Fields. Ph.D. thesis. Uppsala University, Sweden.

Semedo, A., Sušelj, K., Rutgersson, A., Sterl, A. (2011). A Global View on the Wind Sea and Swell Climate and Variability from EERA-40. Journal of Climate, 24, 5, 1461-1479.

Semedo, A., Behrens, R., Sterl, A., Bengtsson, L., Günther, H. (2013). Projection of Global Wave Climate Change toward the End of the Twenty-First Century. Journal of Climate, 26, 8269-8288.

Semedo, A., Vettor, R., Breivik, Ø., Sterl, A., Reistad, M., Soares, C. G., Lima, D. (2014). The wind sea and swell waves climate in the Nordic seas. Ocean Dynamics, DOI: 10.1007/s10236-014-0788-4.

Semedo, A., Dobrynin, M., Lemos, G., Behrens, A., Staneva, J., de Vries, H., Sterl, A., Bidlot, J., Döscher, R., Murawski, J., Miranda, P. (2018b). CMIP5 Derived Single-Forcing, Single-Model and Single-Scenario Wind Wave Climate Simulations: Ensemble Configuration and Performance Evaluation. Journal of Marine Science and Engineering, 6, 90, 28pp.

Shih, S. M., Komar, P. D., Tillotson, K. J., McDougal, W.G., Ruggiero, P. (1995). Wave run-up and sea-cliff erosion. Coast. Eng., 2170–2184.

Stocker, T. F., et al., Eds., IPCC. (2013). Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report 1 of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge and New York, 1535 pp.

Stopa, J. E., 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, 65-83.

Stopa, J. E. (2018). Wind forcing calibration and wave hindcast comparison using multiple reanalysis and merged satellite wind datasets. Oceanogr. Meteorol., 127, 55–69.

Stopa, J. E., Semedo, A., Dobrynin, M., Behrens, A., Staneva, J., Lemos, G. (2019). A sampling technique to compare climate simulations with sparse satellite observations: Performance evaluation of a CMIP5 EC-Earth forced dynamical wave climate ensemble with altimeter observations. Ocean Modelling, 134, 18-29.

Sullivan, P. P., Edson, J. B., Hristov, T., McWilliams, J. C. (2008). Large-eddy Simulations and observations of atmospheric marine boundary layers above nonequilibrium surface waves. Journal of Atmospheric Science, 65, 1225-1245.

SWAN Team. (2022). SWAN user manual. SWAN Cycle III version 41.41. Delft University of Technology; 2022. p. 154. Technical documentation.

Tolman, H., Chalikov, D. (1996). Source Terms in a Third-Generation Wind Wave Model. Journal of Physical Oceanography, 26(11), 2497-2518.

Chawla, A., Tolman, H. L. (2007). Automated grid generation for WAVEWATCH III. Tech. Note 254, NOAA/NWS/NCEP/MMAB, 71 pp.

Chawla, A., Tolman, H. L. (2008). Obstruction grids for spectral wave models. Ocean Mod., 22, 12–25.

Tolman, H. (2009). User manual and system documentation of WAVEWATCH III version 3.14. NOAA / NWS / NCEP / MMAB Technical Note, 276, 2009.

Torralba, V., Doblas-Reyes, F. J. Gonzalez-Reviriego, N. (2017). Uncertainty in recent near-surface wind speed trends: a global reanalysis intercomparison. Environmental Research Letters 12, 114019, doi:10.1088/1748-9326/aa8a58.

Tuomi, L., Kahma, K., Pettersson, H. (2011). Wave hindcast statistics in the seasonally ice-covered Baltic Sea. Boreal Environment Research, 16, 451-472.

Vitousek, S., Barnard, P. L., Fletcher, C. H., Frazer, N., Erikson, L., Storlazzi, C. D. (2017). Doubling of coastal flooding frequency within decades due to sea-level rise. Sci. Rep., 7, 1399.

Vousdoukas, M. I., Mentaschi, L., Voukouvalas, E., Verlaan, M., Jevrejeva, S., Jackson, L. P., Feyen, L. (2018). Global probabilistic projections of extreme sea levels show intensification of coastal flood hazard. Nat. Commun., 9, 2360.

Wallace, J. M., Deser, C., Smoliak, B. V., Phillips, A. S. (2015). Attribution of Climate Change in the Presence of Internal Variability. In Climate Change: Multidecadal and Beyond (Eds: Chang, C.-P., Ghil, M., Latif, M., Wallace, J. M.), World Scientific.

WAMDI Group. (1988). The WAM Model – A Third Generation Ocean Wave Prediction Model. Journal of Physical Oceanography, 18, 1775-1810.

Wang, X. L., Feng, Y., Swail, V. R. (2015). Climate change signal and uncertainty in CMIP5-based projections of global ocean surface wave heights. Journal of Geophysical Research, 120, 5, 3859-3871.

Watterson, I. G. (1996). Non-dimensional measures of climate model performance. Int. J. Climatol., 16, 379-391.

Watterson, I. G., Bathols, J., Heady, C. (2014). What influences the skill of climate models over the continents? Bull. Am. Metorol. Soc., 95, 689-700.

The WAVEWATCH III Development Group [WW3DG]. (2019). User Manual and System Documentation of WAVEWATCH III-Version 6.07; Tech. Note 333; NOAA/NWS/NCEP/MMAB: College Park, MD, USA, 465p.

Young, I. R. (1999). Seasonal variability of the global ocean wind and wave climate. International Journal of Climatology, 19, 931-950.

Young, I. R., Zieger, S., Babanin, V. (2011). Global trends in wind speed and wave height. Science, 332, 451-455.

Zieger, S., Babanin, A. V., Rogers, E. W., Young, I. R. (2015). Observation-based source terms in the third-generation wave model WAVEWATCH. Ocean Modelling, 96, 2–25.

Zijlema, M. van Vledde, G. P., Holthuijsen, L. H. (2012). Bottom friction and wind drag forwave models. Coast Eng, 65, 19–26.

FIGURES AND TABLES

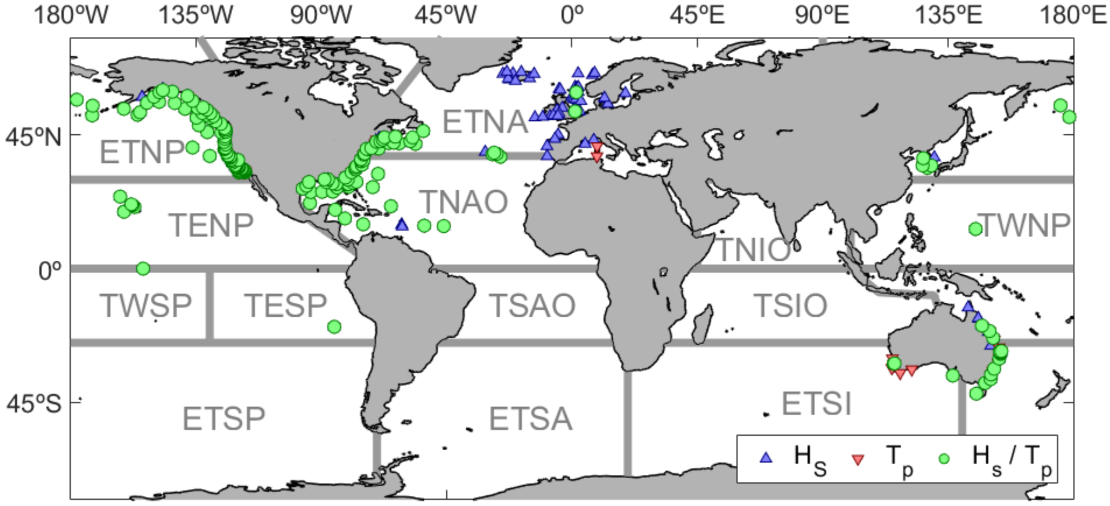


Figure 1 – Map with areas, following Alves (2006). Selected areas: extratropical North Atlantic (ETNA), extratropical South Atlantic (ETSA), extratropical North Pacific (ETNP), extratropical South Pacific (ETSP), extratropical South Indian (ETSI), tropical North Atlantic (TNAO), tropical South Atlantic (TSAO), tropical western North Pacific (TWNP), tropical eastern North Pacific (TENP), tropical western South Pacific (TWSP), tropical eastern South Pacific (TESP), tropical North Indian (TNIO), tropical South Indian (TSIO). Further details can be seen on Table SM1. Selected *in-situ* locations are marked according to the available wave parameters: (blue) H_S, (red) T_p and (green) both.

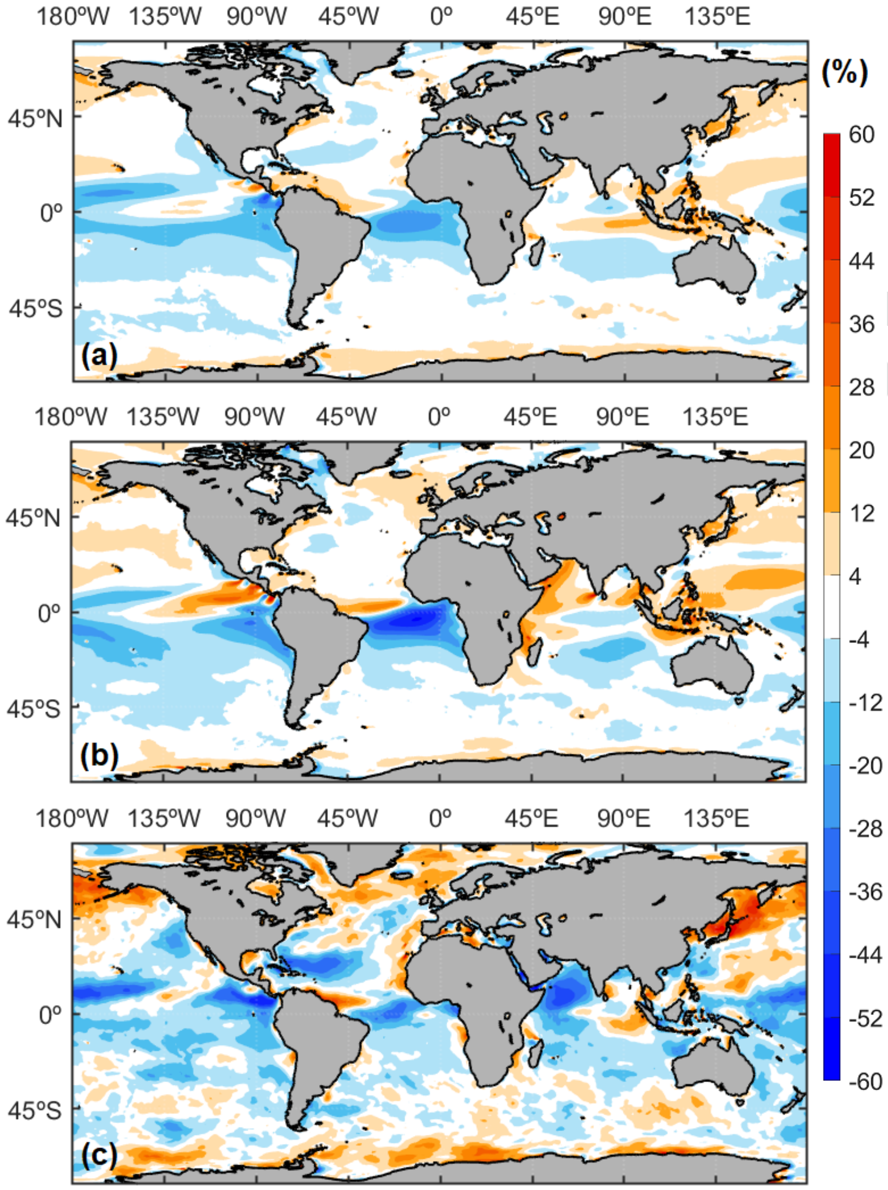


Figure 2 – Normalized differences (in %) between the (a) annual, (b) DJF and (c) JJA EC-Earth3 and ERA5 U_{10} means for the 1995-2014 historical period.

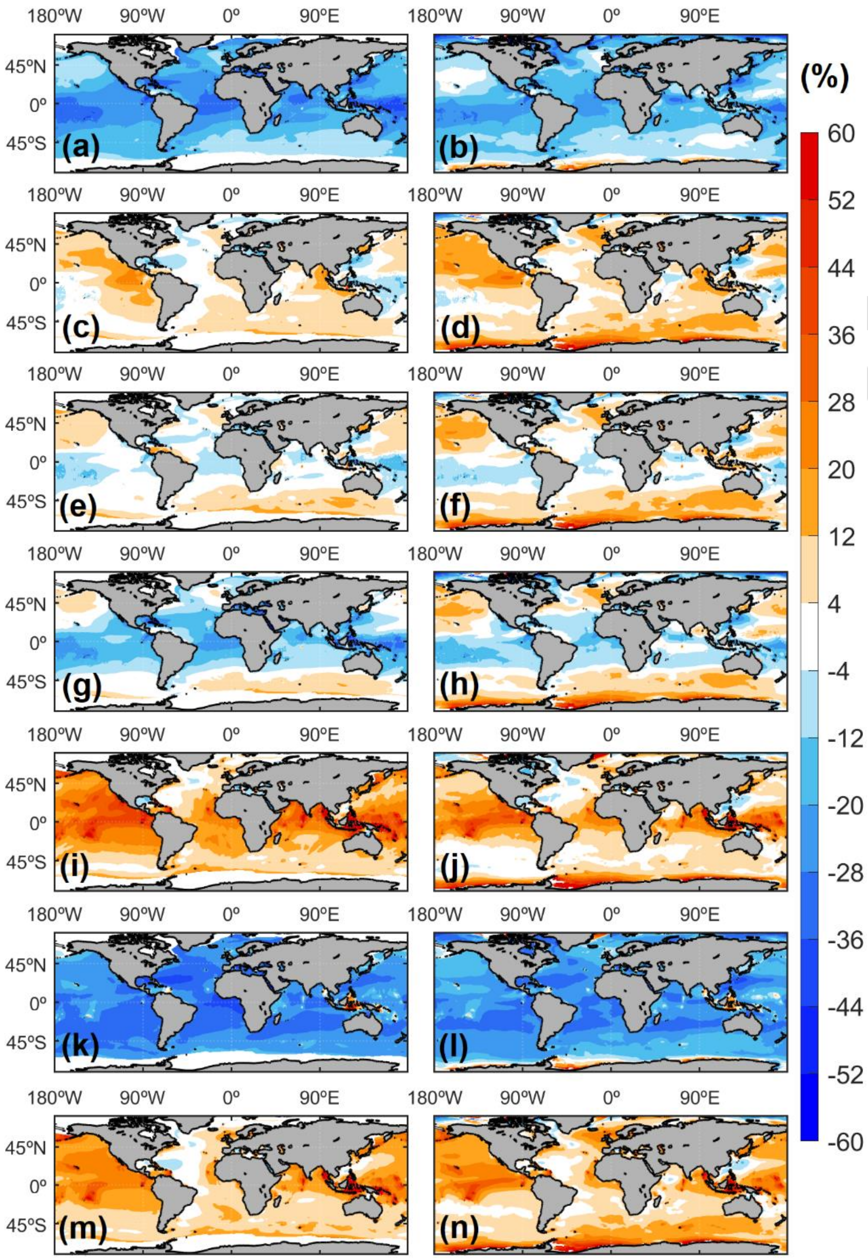


Figure 3 – Normalized differences (in %) between the (top) PC20-1 (WW3-ST2) to (bottom) -7 (WAM4.6) ensemble members' and ERA5 (left) annual mean H_S and (right) 95% percentile H_S (1995-2014).

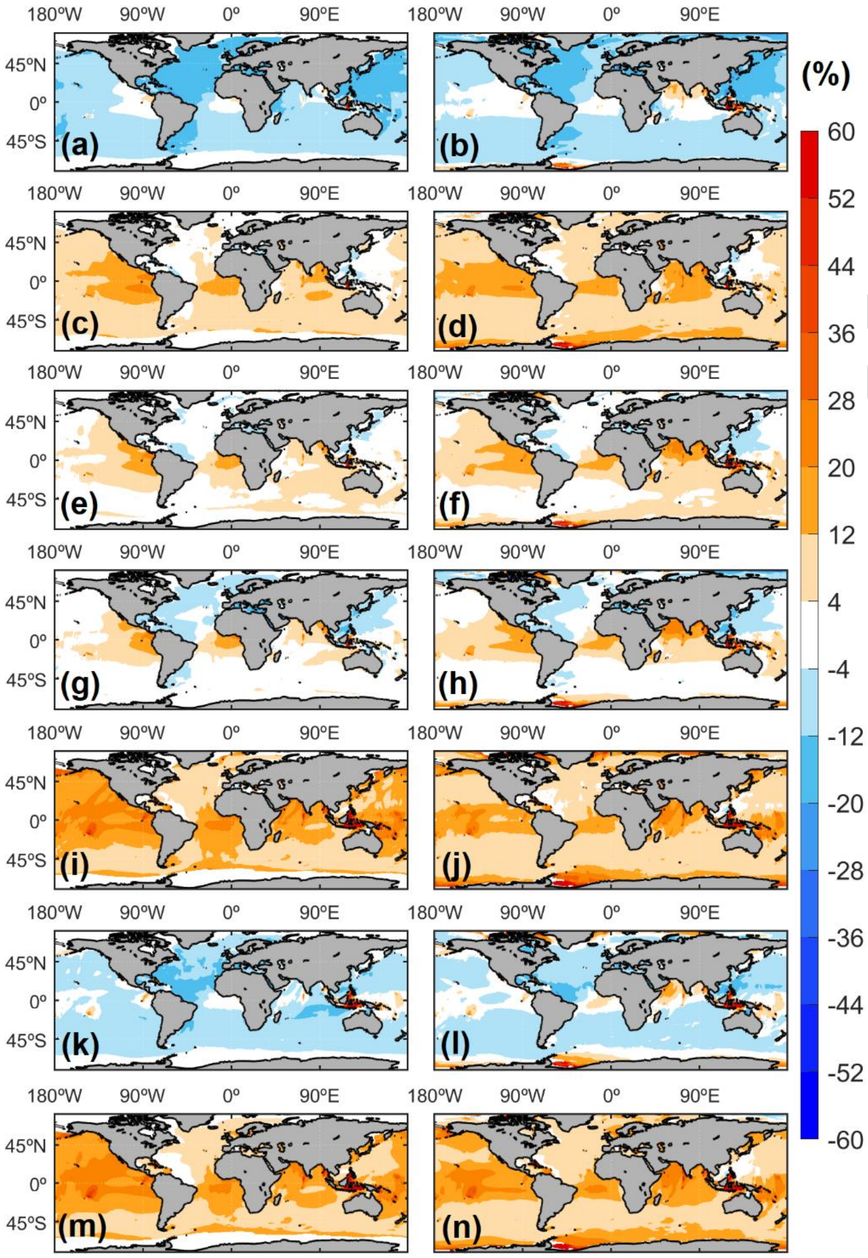


Figure 4 – Same as in Figure 3, but for T_m (s).

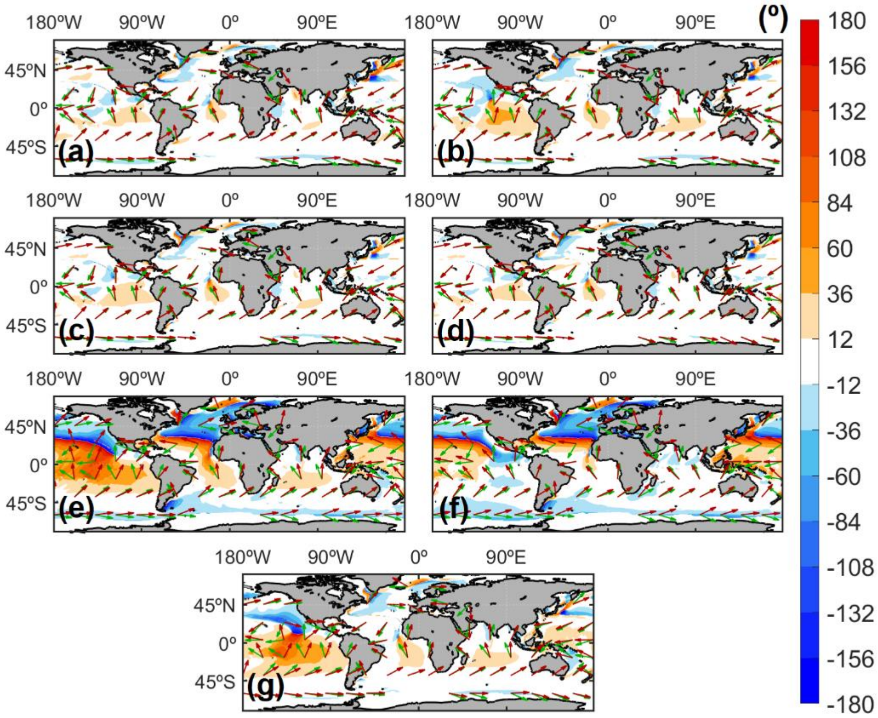


Figure 5 – Absolute differences (in º) between the (a) PC20-1 (WW3-ST2) to (g) -7 (WAM4.6) ensemble members' (red arrows) and ERA5 (green arrows) annual mean MWD (1995-2014).

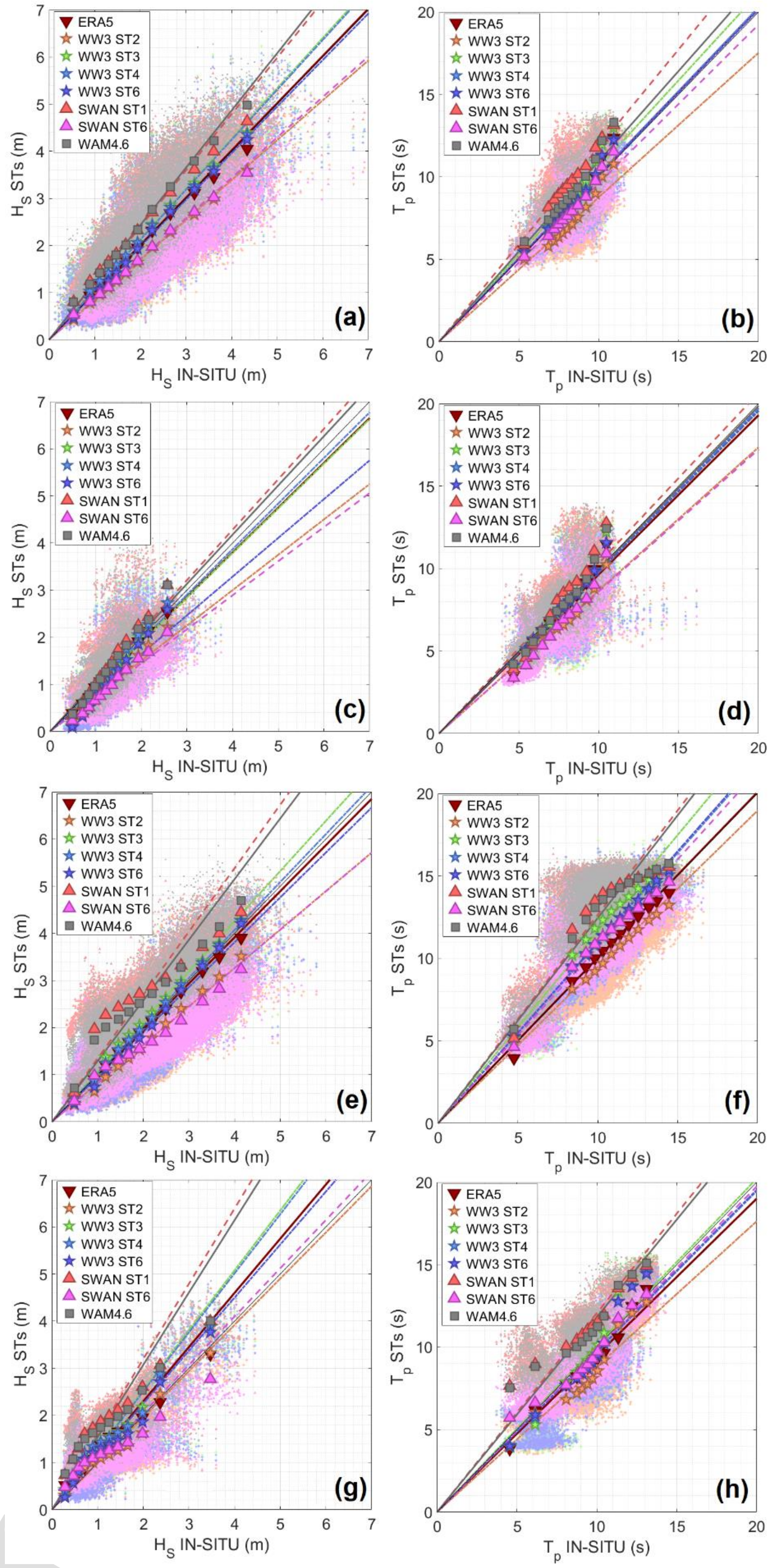


Figure 6 – Merged scatter-QQ-plots from the comparison between *in-situ* multi-year (1995-2014) daily (left) H_S and (right) T_p means, ERA5 and PC20-1 (WW3-ST2) to -7 (WAM4.6) ensemble members at the available *in-situ* locations across (a,b) ETNA, (c,d) TNAO, (e,f) ETNP and (g,h) TWSP/ETSP regional areas. Highlighted percentiles in the QQ-plots refer to the 1%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95% and 99% ones.

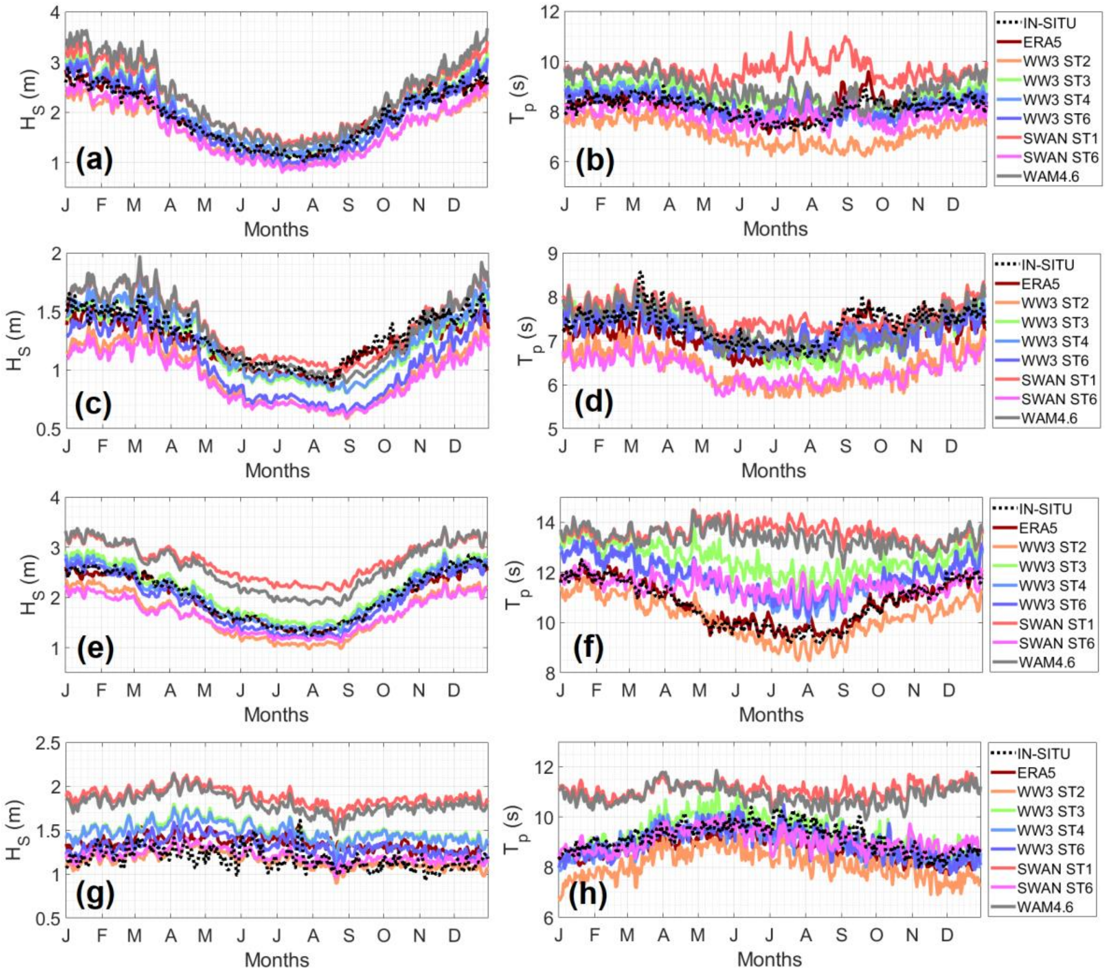


Figure 7 – Mean intra-annual (left) H_S and (right) T_p cycles (1995-2014) considering the *in-situ* observations, ERA5 and PC20-1 (WW3-ST2) to -7 (WAM4.6) ensemble members at the available *in-situ* locations across (a,b) ETNA, (c,d) TNAO, (e,f) ETNP and (g,h) TWSP/ETSP regional areas.

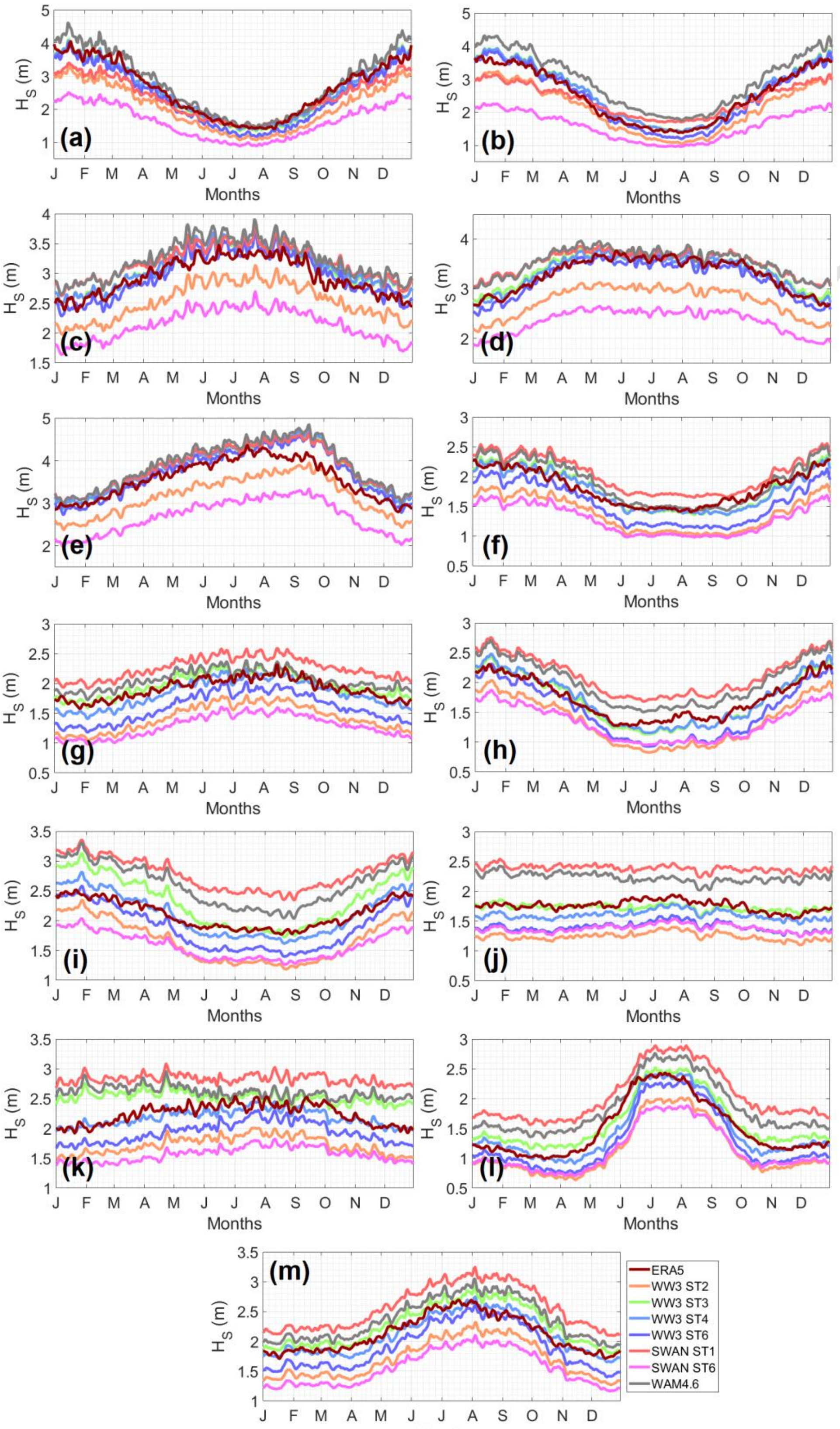


Figure 8 – Mean intra-annual H_S cycles (1995-2014) considering the ERA5 and PC20-1 (WW3-ST2) to -7 (WAM4.6) ensemble members across (a) ETNA, (b) ETNP, (c) ETSA, (d) ETSP, (e) ETSI, (f) TNAO, (g) TSAO, (h) TWNP, (i) TENP, (j) TWSP, (k) TESP, (l) TNIO and (m) TSIO regional areas.

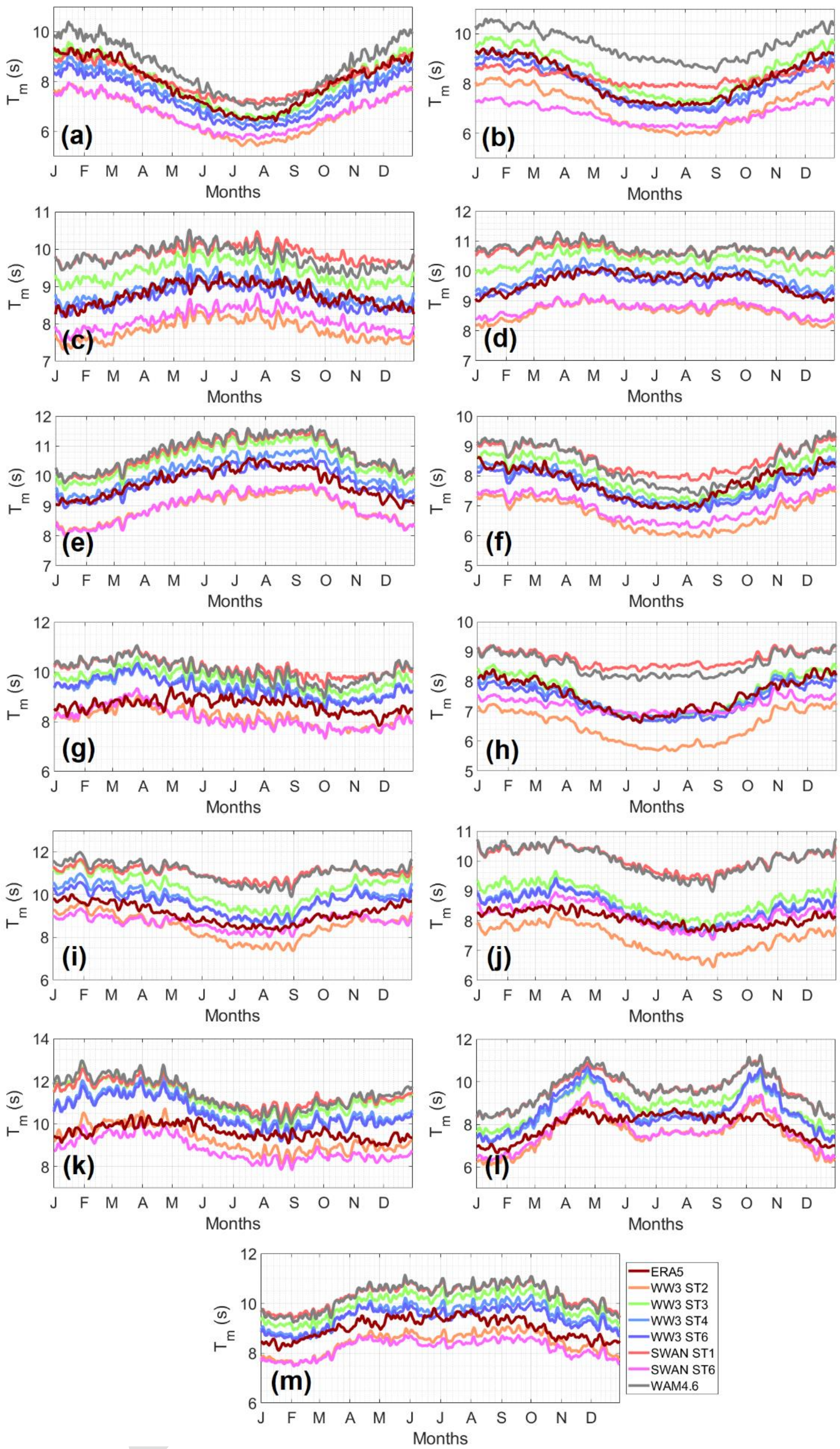


Figure 9 – Same as in Figure 8, but for T_m (s).

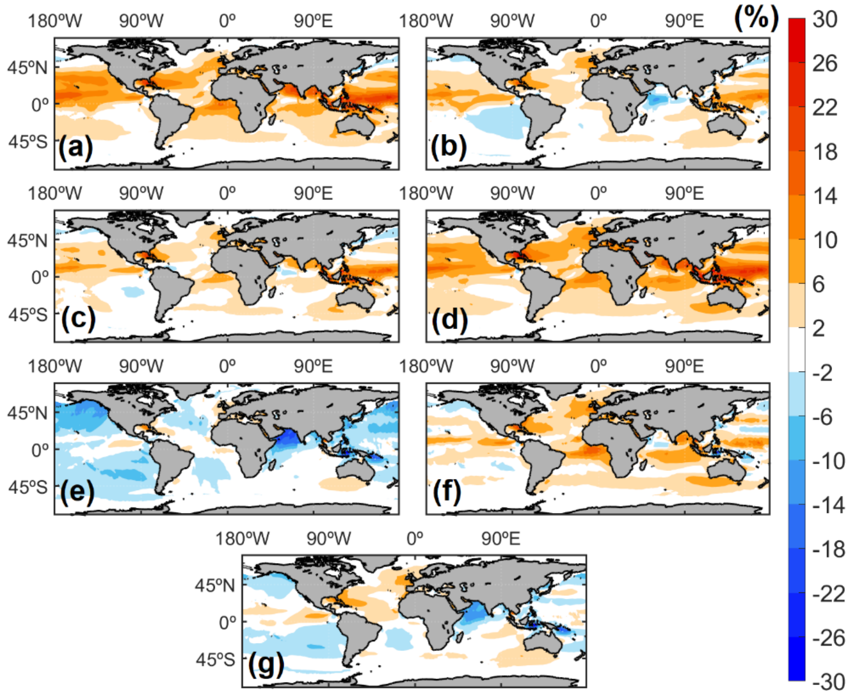


Figure 10 – Normalized differences (in %) between the (a) PC20-1 (WW3-ST2) to (g) -7 (WAM4.6) H_S MAVs and ERA5 ones (1995-2014).

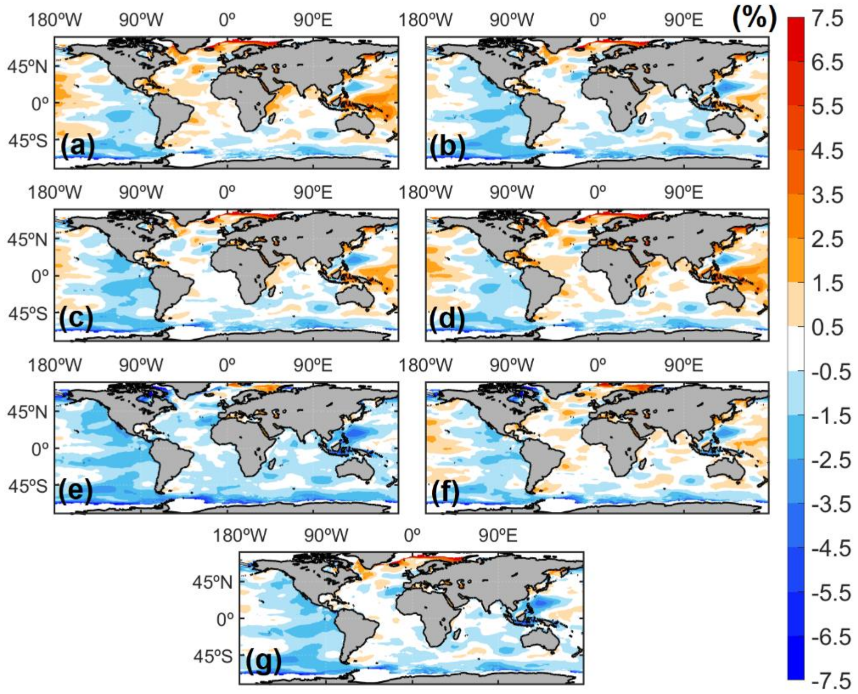


Figure 11 – Normalized differences (in %) between the (a) PC20-1 (WW3-ST2) to (g) -7 (WAM4.6) H_S IAVs and ERA5 ones (1995-2014).

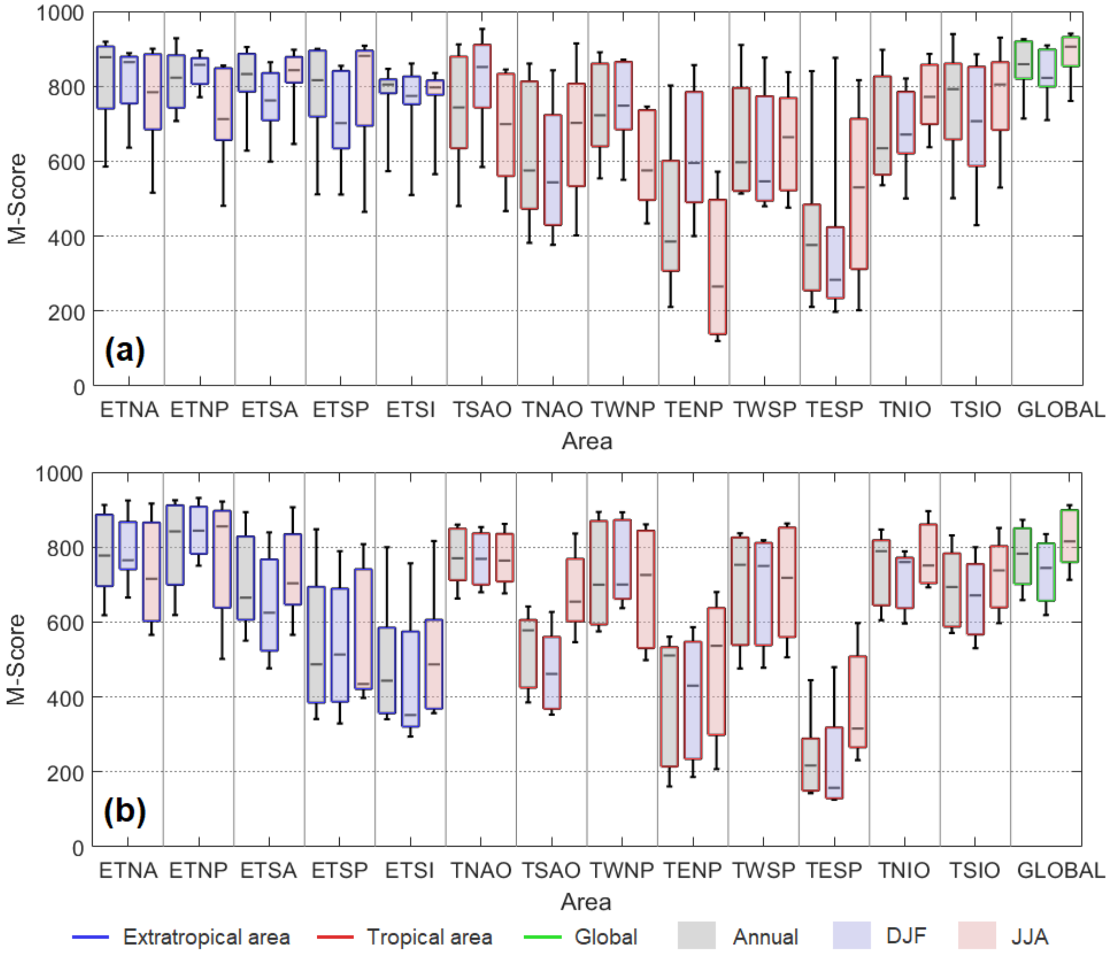


Figure 12 – Boxplots representing the range of (a) H_S and (b) T_m M-scores within the PC20 ensemble (1995-2014), from the comparison with ERA5, globally and across each of the regional areas. Extratropical areas are outlined in blue, tropical areas in red and the global in green. Grey shading represents the annual boxplots, while blue and red shadings represent DJF and JJA, respectively.

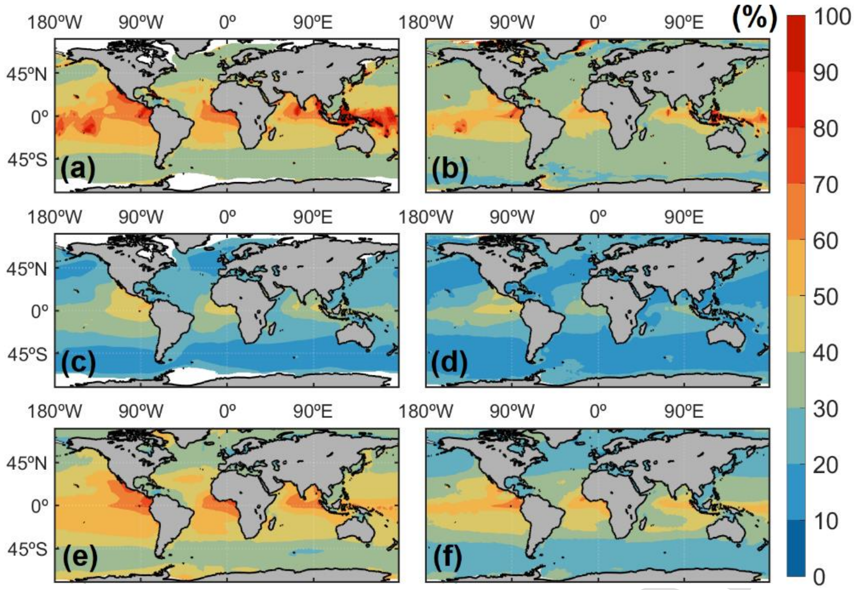


Figure 13 – Normalized uncertainty range (NUR; in %) for (a,b) the PC20 ensemble, (c,d) the WW3 subset of PC20 (*i.e.*, only PC20-1 to -4; ST2 to ST6) and (e,f) the SWAN subset of PC20 (*i.e.*, only PC20-5 and -6; ST1 and ST6), considering the annual (left) H_S means and (right) 95% percentiles (1995-2014).

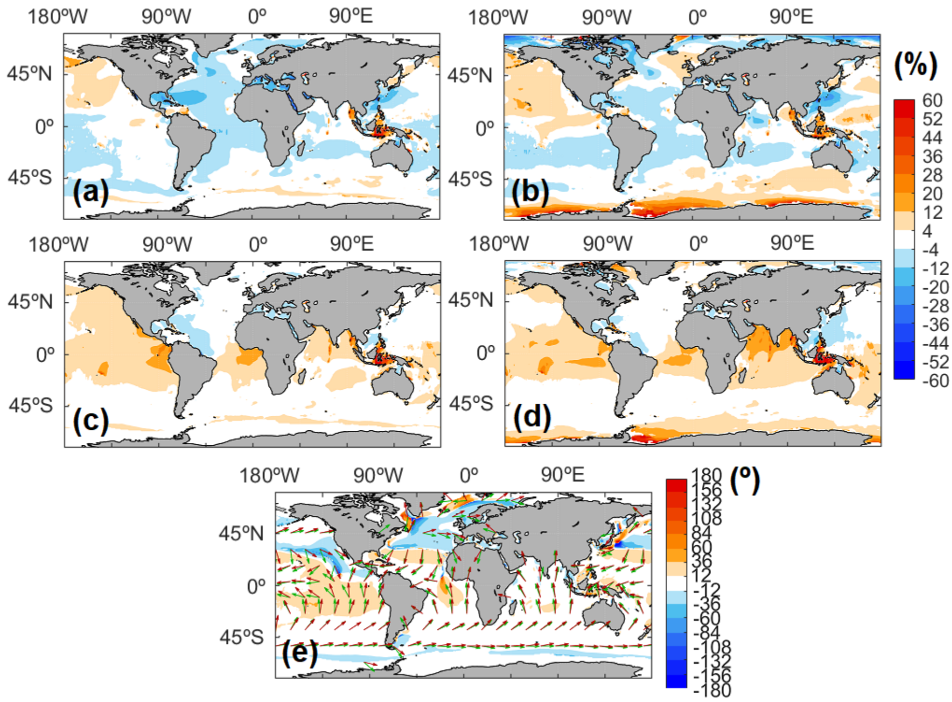


Figure 14 – 7-member PC20 full ensemble mean (a-d) normalized (in %) and (e) absolute (in °) differences in comparison with ERA5, considering the annual mean (a) H_S, (c) T_m and (e) MWD, as well as the annual 95% percentile (b) H_S and (d) T_m (1995-2014).

Table 1 – Statistic metrics representing the PC20-i (1 to 7) performance in representing the H_S climate at the selected *in-situ* locations across each area (89 at ETNA, 42 at TNAO, 84 at ETNP and 22 at TWSP/ETSP).

	ETNA (89)				
	Bias (m)	**RMSE (m)**	**R**	**SI**	**Slope**
WW3-ST2	-0.29	0.56	0.87	0.30	0.85
WW3-ST3	0.13	0.46	0.87	0.24	1.07
WW3-ST4	0.12	0.46	0.87	0.24	1.06
WW3-ST6	-0.02	0.46	0.87	0.24	0.99
SWAN-ST1	0.37	0.54	0.88	0.29	1.20
SWAN-ST6	-0.27	0.55	0.87	0.29	0.86
WAM4.6	0.41	0.60	0.88	0.32	1.22
	TNAO (42)				
	Bias (m)	**RMSE (m)**	**R**	**SI**	**Slope**
WW3-ST2	-0.32	0.46	0.83	0.36	0.75
WW3-ST3	-0.07	0.35	0.82	0.27	0.95
WW3-ST4	-0.04	0.34	0.83	0.27	0.97
WW3-ST6	-0.23	0.43	0.82	0.33	0.82
SWAN-ST1	0.09	0.40	0.78	0.31	1.08
SWAN-ST6	-0.36	0.50	0.79	0.39	0.72
WAM4.6	0.06	0.38	0.79	0.29	1.05
	ETNP (84)				
	Bias (m)	**RMSE (m)**	**R**	**SI**	**Slope**
WW3-ST2	-0.37	0.59	0.90	0.29	0.82
WW3-ST3	0.13	0.38	0.91	0.19	1.06
WW3-ST4	0.04	0.40	0.91	0.20	1.02
WW3-ST6	-0.10	0.44	0.91	0.22	0.95
SWAN-ST1	0.70	0.78	0.84	0.38	1.35
SWAN-ST6	-0.38	0.65	0.87	0.32	0.82
WAM4.6	0.58	0.66	0.88	0.32	1.29
	TWSP / ETSP (22)				
	Bias (m)	**RMSE (m)**	**R**	**SI**	**Slope**
WW3-ST2	-0.02	0.48	0.79	0.41	0.98
WW3-ST3	0.32	0.43	0.82	0.39	1.27
WW3-ST4	0.30	0.44	0.80	0.38	1.25
WW3-ST6	0.15	0.42	0.81	0.35	1.13
SWAN-ST1	0.70	0.67	0.80	0.57	1.60
SWAN-ST6	0.29	0.50	0.76	0.42	1.02
WAM4.6	0.63	0.62	0.80	0.52	1.54

Table 2 – Summary of PC20-i (1 to 7) H_S biases (in m) in comparison with *in-situ* observations at ETNA, TNAO, ETNP and TWSP/ETSP, at specific percentiles (10%, 50%, 90%, 95% and 99%).

	ETNA (89)				
	Bias P10%	**Bias P50%**	**Bias P90%**	**Bias P95%**	**Bias P99%**
WW3-ST2	-0.11	-0.27	-0.52	-0.63	-0.74
WW3-ST3	0.13	0.10	0.14	0.09	0.06
WW3-ST4	0.13	0.10	0.08	0.04	0.01
WW3-ST6	-0.03	-0.05	0.01	-0.04	-0.09
SWAN-ST1	0.34	0.33	0.43	0.39	0.29
SWAN-ST6	-0.09	-0.25	-0.47	-0.58	-0.80
WAM4.6	0.28	0.33	0.61	0.61	0.64
	TNAO (42)				
	Bias P10%	**Bias P50%**	**Bias P90%**	**Bias P95%**	**Bias P99%**
WW3-ST2	-0.36	-0.29	-0.32	-0.33	-0.35
WW3-ST3	-0.23	-0.04	0.08	0.09	0.16
WW3-ST4	-0.26	-0.0004	0.12	0.12	0.16
WW3-ST6	-0.39	-0.21	-0.08	-0.07	0.02
SWAN-ST1	-0.16	0.08	0.31	0.28	0.54
SWAN-ST6	-0.33	-0.36	-0.39	-0.47	-0.46
WAM4.6	-0.12	0.04	0.24	0.22	0.54
	ETNP (84)				
	Bias P10%	**Bias P50%**	**Bias P90%**	**Bias P95%**	**Bias P99%**
WW3-ST2	-0.29	-0.38	-0.52	-0.59	-0.63
WW3-ST3	0.05	0.13	0.12	0.12	0.14
WW3-ST4	-0.06	0.02	0.08	0.09	0.13
WW3-ST6	-0.19	-0.14	0.01	0.02	0.06
SWAN-ST1	1.04	0.80	0.41	0.33	0.30
SWAN-ST6	0.05	-0.37	-0.74	-0.84	-0.91
WAM4.6	0.80	0.60	0.47	0.47	0.55
	TWSP / ETSP (22)				
	Bias P10%	**Bias P50%**	**Bias P90%**	**Bias P95%**	**Bias P99%**
WW3-ST2	0.11	0.02	-0.36	0.08	-0.14
WW3-ST3	0.27	0.38	0.19	0.51	0.40
WW3-ST4	0.26	0.36	0.07	0.51	0.42
WW3-ST6	0.10	0.23	-0.12	0.34	0.29
SWAN-ST1	0.59	0.80	0.62	0.70	0.52
SWAN-ST6	0.26	0.10	-0.37	-0.40	-0.70
WAM4.6	0.62	0.68	0.55	0.65	0.54

Table 3 – Similar to Table 1, but for T_p (s).

	ETNA (24)				
	Bias	**RMSE**	**R**	**SI**	**Slope**
WW3-ST2	-1.01	1.51	0.72	0.19	0.88
WW3-ST3	0.47	1.17	0.71	0.14	1.06
WW3-ST4	0.11	0.95	0.73	0.12	1.01
WW3-ST6	0.06	1.02	0.69	0.13	1.01
SWAN-ST1	1.49	1.83	0.63	0.22	1.18
SWAN-ST6	-0.31	1.09	0.68	0.13	0.96
WAM4.6	0.77	1.31	0.69	0.16	1.09
	TNAO (39)				
	Bias	**RMSE**	**R**	**SI**	**Slope**
WW3-ST2	-0.97	1.55	0.75	0.21	0.87
WW3-ST3	-0.17	1.15	0.78	0.16	0.98
WW3-ST4	-0.10	1.02	0.78	0.14	0.99
WW3-ST6	-0.16	1.08	0.77	0.15	0.98
SWAN-ST1	0.24	1.25	0.84	0.17	1.03
SWAN-ST6	-1.02	1.30	0.81	0.18	0.86
WAM4.6	-0.05	1.03	0.82	0.14	0.99
	ETNP (79)				
	Bias	**RMSE**	**R**	**SI**	**Slope**
WW3-ST2	-0.56	1.18	0.85	0.11	0.95
WW3-ST3	1.78	2.12	0.81	0.20	1.17
WW3-ST4	1.00	1.41	0.86	0.13	1.09
WW3-ST6	1.08	1.49	0.82	0.14	1.10
SWAN-ST1	2.85	3.37	0.68	0.31	1.26
SWAN-ST6	0.72	1.58	0.78	0.15	1.07
WAM4.6	2.63	3.10	0.74	0.29	1.25
	TWSP / ETSP (29)				
	Bias	**RMSE**	**R**	**SI**	**Slope**
WW3-ST2	-1.09	1.84	0.83	0.20	0.88
WW3-ST3	0.10	1.33	0.88	0.14	1.01
WW3-ST4	-0.25	1.26	0.87	0.14	0.97
WW3-ST6	-0.22	1.28	0.86	0.14	0.98
SWAN-ST1	1.90	2.17	0.80	0.24	1.21
SWAN-ST6	-0.12	1.04	0.84	0.11	0.99
WAM4.6	1.67	2.07	0.75	0.23	1.18

Table 4 – Similar to Table 2, but for T_p (s).

	ETNA (24)				
	Bias P10%	**Bias P50%**	**Bias P90%**	**Bias P95%**	**Bias P99%**
WW3-ST2	-1.07	-1.20	-0.81	-0.19	-0.16
WW3-ST3	0.15	0.35	1.10	1.74	2.04
WW3-ST4	0.08	-0.06	0.33	1.10	1.37
WW3-ST6	-0.04	-0.09	0.34	1.10	1.29
SWAN-ST1	1.29	1.48	1.96	2.18	2.21
SWAN-ST6	-0.45	-0.40	-0.05	0.40	0.58
WAM4.6	0.53	0.58	1.29	1.99	2.37
	TNAO (39)				
	Bias P10%	**Bias P50%**	**Bias P90%**	**Bias P95%**	**Bias P99%**
WW3-ST2	-0.62	-1.05	-1.16	-0.94	-0.23
WW3-ST3	-0.56	-0.20	0.20	0.88	1.74
WW3-ST4	-0.30	0.03	-0.14	0.20	1.10
WW3-ST6	-0.42	-0.09	-0.12	0.14	1.11
SWAN-ST1	-0.85	0.68	0.56	1.33	2.28
SWAN-ST6	-1.31	-0.89	-1.07	-0.71	0.44
WAM4.6	-0.52	0.04	0.15	0.84	1.94
	ETNP (79)				
	Bias P10%	**Bias P50%**	**Bias P90%**	**Bias P95%**	**Bias P99%**
WW3-ST2	-0.27	-0.60	-0.89	-1.01	-1.23
WW3-ST3	1.80	2.00	1.53	1.38	1.11
WW3-ST4	1.05	1.05	0.95	0.86	0.61
WW3-ST6	1.14	1.08	1.02	1.04	0.87
SWAN-ST1	3.33	3.38	1.91	1.60	1.09
SWAN-ST6	1.08	0.89	0.37	0.30	0.18
WAM4.6	2.80	3.11	1.96	1.71	1.30
	TWSP / ETSP (29)				
	Bias P10%	**Bias P50%**	**Bias P90%**	**Bias P95%**	**Bias P99%**
WW3-ST2	-0.09	-1.66	0.02	-0.15	-0.38
WW3-ST3	-0.84	-0.22	2.03	2.02	1.95
WW3-ST4	-0.40	-0.69	1.45	1.52	1.45
WW3-ST6	-0.25	-0.70	1.42	1.47	1.33
SWAN-ST1	2.82	1.61	2.21	1.99	1.83
SWAN-ST6	0.49	-0.47	0.40	0.33	0.09
WAM4.6	2.68	1.21	2.42	2.20	2.00

Performance Evaluation of a Global CMIP6 Single Forcing, Multi Wave Model Ensemble of Wave Climate Simulations – Highlights

- A performance evaluation of a new CMIP6 ensemble of wave climate simulations
- Ensemble architecture focused on model-parameterization-induced uncertainties
- Evaluation through comparison with ERA5 and an extensive *in-situ* observational set
- Ensemble spreads found to exceed even high-end projected change rates until 2100

Performance Evaluation of a CMIP6 Single Forcing, Multi Wave Model Ensemble of Wave Climate Simulations

AUTHOR STATEMENT

Gil Lemos: Conceptualization, Methodology, Software, Formal Analysis, Writing – Original Draft

Alvaro Semedo: Conceptualization, Methodology, Validation, Investigation, Supervision

Rajesh Kumar: Conceptualization, Software, Validation, Investigation, Resources,

Mikhail Dobrynin: Software, Investigation, Resources

Adem Akpinar: Software, Investigation, Resources, Writing – Review & Editing

Bahareh Kamranzad: Software, Investigation, Resources

Jean Bidlot: Resources, Data Curation

Hector Lobeto: Validation, Data Curation, Writing – Review & Editing

Declaration of interests

☒ 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.

☐ The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: