



HAL
open science

Historical global ocean wave data simulated with CMIP6 anthropogenic and natural forcings

Anindita Patra, Guillaume Dodet, Mickael Accensi

► **To cite this version:**

Anindita Patra, Guillaume Dodet, Mickael Accensi. Historical global ocean wave data simulated with CMIP6 anthropogenic and natural forcings. *Scientific Data*, 2023, 10 (1), 325 (10p.). 10.1038/s41597-023-02228-6 . hal-04204065

HAL Id: hal-04204065

<https://hal.science/hal-04204065>

Submitted on 14 Sep 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



OPEN

DATA DESCRIPTOR

Historical global ocean wave data simulated with CMIP6 anthropogenic and natural forcings

Anindita Patra , Guillaume Dodet & Mickaël Accensi

This dataset presents historical ocean wave climate during 1960–2020, simulated using the numerical model WAVEWATCH III (WW3) forced by Coupled Model Intercomparison Project phase 6 (CMIP6) simulations corresponding to natural-only (NAT), greenhouse gas-only (GHG), aerosol-only (AER) forcings, combined forcing (natural and anthropogenic; ALL), and pre-industrial control conditions. Surface wind at 3-hourly temporal resolution, and sea-ice area fraction at monthly frequency, from a CMIP6 model - MRI-ESM2.0 are used to force WW3 over the global ocean. Model calibration and validation of the significant wave height are carried out using inter-calibrated multi-mission altimeter data produced by the European Space Agency Climate Change Initiative, with additional corroboration using ERA-5 reanalysis. The simulated dataset is assessed for its skill to represent mean state, extremes, trends, seasonal cycle, time consistency, and spatial distribution over time. Numerically simulated wave parameters for different individual external forcing scenario is not available yet. This study produces a novel database particularly useful for detection and attribution analysis to quantify the relative contributions of natural and anthropogenic forcings to historical changes.

Background & Summary

The relevance of surface ocean wind waves is manifold, from environmental to socioeconomic impacts on human society. A wide range of applications of wave information already stated by many previous literature include ship routing¹, coastal and offshore engineering activities², assessment of coastal hazards³, wave energy extraction planning⁴, estimation of ocean–atmosphere exchange⁵, among others. The importance of dynamic coupling of wave processes into coupled ocean–atmosphere global climate models has been emphasized since long back as surface wind-waves are capable of directly affecting climate system through continuous ocean–atmosphere exchange⁵.

Despite being a key component of the climate system⁶, ocean waves data are not available from the state-of-art global climate models (GCM). Using outputs of CMIP5/CMIP6 GCMs to force wave model, several ensembles of global wave climate projections are available for historical as well as for future periods, especially within the framework of the Coordinated Ocean Wave Climate Project (COWCLIP)⁷. In order to understand wave climate response to growing anthropogenic interference, we need ocean wave information for different forcing scenario. Anthropogenic greenhouse gases and aerosols are major external forcings to global climate change^{8,9}. CMIP6 multi-model simulations¹⁰ provide scenarios for various external forcings under the Detection and Attribution Model Intercomparison Project (DAMIP)¹¹. The goal of DAMIP is to attribute historically observed climate change to anthropogenic and natural causes, to derive observationally constrained projections of future climate change, and to improve understanding of the mechanisms by which particular forcings affect climate. Following the approach of Tier 1 experiments of DAMIP, we aim to create novel wave database corresponding to CMIP6 historical, hist-nat (natural-only), hist-GHG (greenhouse-only) and hist-aer (aerosol-only) experiments (described in Table 1) separately. Although few recent wave climate projections^{12–14} are available using CMIP6 forcings, these include all forcings together. The DECK (Diagnostic, Evaluation and Characterization of Klima) experiment¹⁰ within CMIP6 framework provides pre-industrial control simulation, which is representative of the period prior to the onset of large-scale industrialization, with 1850 being the reference year. The assumptions of the control simulation is that there are no changes in both naturally occurring forcings (e.g. volcanoes, orbital or solar changes) and human-induced forcings. Control simulation can thus be used to study the unforced

Univ Brest, CNRS, Ifremer, IRD, Laboratoire d’Océanographie Physique et Spatiale (LOPS), 29280, Plouzané, France. ✉e-mail: anindita.patra@ifremer.fr

Experiment	Description	Duration
CTL	Pre-industrial control simulation time invariant representative of the year around 1850	60-year chunks (3)
ALL	CMIP6 historical simulation and SSP2-4.5 forced by evolving external forcing such as solar variability, volcanic aerosols, and changes in atmospheric composition (GHGs and aerosols) caused by human activities.	1960–2020
NAT	Natural-only historical simulations (solar irradiance, stratospheric/ volcanic aerosol)	1960–2020
GHG	Historical greenhouse-gas-only simulations forced by well-mixed greenhouse gas changes only	1960–2020
AER	Anthropogenic-aerosol-only historical simulations	1960–2020

Table 1. CMIP6 experiments from MRI-ESM2.0 used in this study.

internal variability of the climate system¹¹. In addition to CMIP6 historical simulations, we aim to simulate wave conditions corresponding to pre-industrial control simulation using DECK results which will be useful to estimate the range of internal climate variability.

Here, we provide 6-hourly time-series of wave parameters (e.g. significant wave height, mean wave period, mean wave direction) for 1960–2020 for different forcing scenario. This extensive digital wave database can be used by various scientific community and policymakers. A direct application is detection and attribution (D&A) studies of historical wave climate change to anthropogenic and natural causes, specifically relative contribution of anthropogenic greenhouse and anthropogenic aerosols. This kind of D&A studies have never been attempted before for wave climate, to the best of our knowledge, owing to lack of dataset for different forcing scenario.

Methods

CMIP6 forcings. We derive surface winds (u and v components) at 3 hourly interval and sea ice area fraction at monthly interval from CMIP6 dataset (<https://esgf-node.llnl.gov/projects/cmip6/>) for historical (natural plus all anthropogenic, called ALL), greenhouse-only (GHG), natural-only (NAT), and aerosol-only forcing (AER) experiments. In CMIP6 dataset, there are many models which can provide surface winds at daily and monthly frequency for each of the above four experiments, but not at the 3 hourly frequency required to accurately reproduce synoptic-scale ocean wave generation. Therefore, we selected the MRI-ESM2.0, the Earth System Model (Version 2.0) by the Meteorological Research Institute¹⁵, as it provides 3 hourly surface wind ($1.125^\circ \times 1.1209^\circ \sim 1.12148^\circ$), and monthly sea-ice area fraction for each of these four CMIP6 experiments: ALL, GHG, AER, NAT.

MRI-ESM2.0 is a global climate model consisting of four major component models: an atmospheric general circulation model (AGCM) with land processes, an ocean–sea-ice general circulation model, an aerosol model, and an atmospheric chemistry model. This climate model has been shown to display realistic reproduction of both mean climate and interannual variability¹⁵. Following a preindustrial spin-up, a preindustrial control experiment was conducted with MRI-ESM2.0 as well as a set of historical simulations from the 1850 to the present, which were driven by forcing based on observations. These experiments are conformed to the protocol specified by CMIP6.

The ALL simulations generally end in 2014, but simulations are extended to 2020 with the Shared Economic Pathway (SSP) 2–4.5 scenario. SSP2-4.5¹⁶ is intermediate level of greenhouse gas emissions and its future aerosol and land use changes represent a broad range of SSP-based integrated assessment model projections¹¹. Following DAMIP project, CMIP6 historical + SSP2-4.5 experiment is called “historical” here and covers the period 1960–2020. Five variants (r1i1p1f1, r2i1p1f1, r3i1p1f1, r4i1p1f1, r5i1p1f1) are available for the whole 1960–2020 period for each of the four experiments (Table 1 and Fig. 1). The different variant label defines the realization number (r), the initialization index (i), the physics index (p), and the forcing index (f); all being positive integer. For a given CMIP6 experiment (for example: Hist-GHG), the “ripf” identifier is used to uniquely identify each simulation of an ensemble of runs contributed by a single model (for more details, visit https://github.com/WCRP-CMIP/CMIP6_CVs). The realization index distinguishes among members of an ensemble of simulations that differ only in their initial conditions (e.g. initialized from different points in a control run). The initialization index is used to distinguish simulations performed under the same conditions but with different initialization procedures. The physics index indicates the physics version used by the model, and the forcing index is used to distinguish different variants of forcing applied. Since the control simulation (CTL) is considered as time invariant, available 200 years of MRI-ESM2.0 CTL data is divided into 3 sub-periods of 60 years which is the length of historical simulations.

Wave modelling. WAVEWATCH III version 7.14 (The WAVEWATCH-III Development Group, 2019) was implemented over a near-global grid (78° S to 80° N) with a spatial resolution of 1.0° and using the GEBCO ocean bathymetry. The wave spectrum is discretized in 24 directions (15° directional resolution), and 36 exponentially spaced frequencies from 0.034 to 0.95 Hz, with a 1.1 increment factor from one frequency to the next. The maximum global time step, the maximum Courant–Friedrichs–Lewy (CFL) time steps for advection (x-y) and refraction (k-theta), and the minimum source term time step are set as 1800s, 300 s, 300 s, and 10 s, respectively. The wind input and wave dissipation source terms are from ST4 parameterization described in Ardhuin *et al.*¹⁷, with adjustments described in Alday *et al.*¹⁸. Non-linear wave interactions are modelled using the Discrete Interaction Approximation (DIA)¹⁹.

Here, we adjusted the β_{max} parameter, which is a non-dimensional wind-wave growth coefficient that has been used as a tuning parameter to calibrate spectral wave models at global scale for wind strength biases²⁰. Other specifications for model parameterizations are from test T475 from the WW3 hindcast implemented at Ifremer¹⁸. Model calibration and performance analysis are done by comparing the significant wave heights

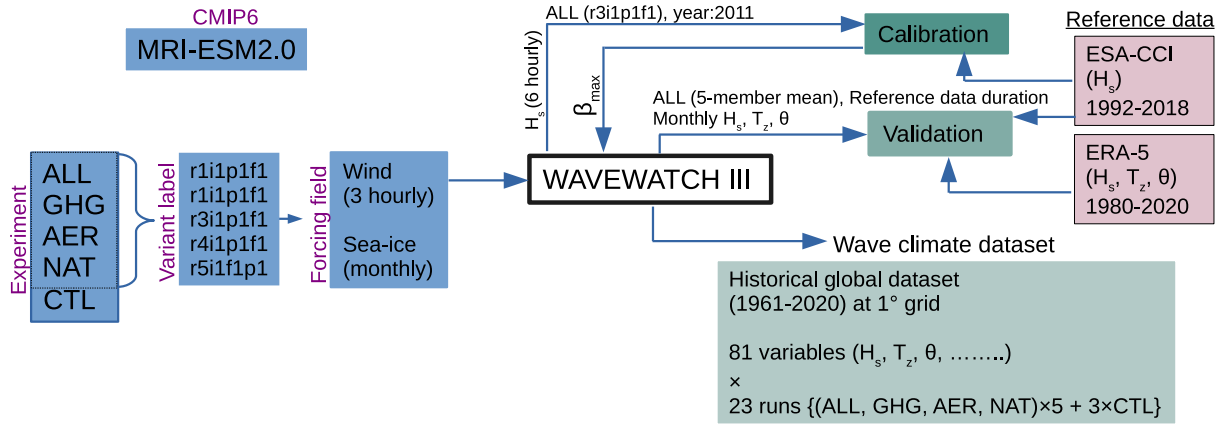


Fig. 1 Schematic diagram of the methods used in this study to generate the wave climate dataset.

β_{max}	0.9	1.0	1.1	1.75
Bias (m)	-0.16	-0.04	0.06	0.66
Nbias (%)	-6.81	-1.82	2.87	27.28

Table 2. Absolute and relative systematic error metrics for SWH (global averaged 60°S to 60°N) for the β_{max} sensitivity runs using ALL simulation (r3i1p1f1).

(SWH) obtained with the historical-ALL simulations and SWH from the inter-calibrated and denoised multi-mission satellite altimeter dataset (L4 product, version 1.1²¹) provided by the European Space Agency Climate Change Initiative (ESA-CCI)²² (Fig. 1). For model calibration, we looked at the systematic error statistics (bias and normalized bias: Nbias) computed over the full year of 2011 (same as Alday *et al.*¹⁸) using β_{max} values increasing from 0.9 to 1.75 (Table 2). The calibration analysis is based on a single year (enough to capture seasonal variability) in order to deal with computational demand. In this calibration analysis, the observation is 1 Hz along-track altimeter data and 6-hourly model outputs are interpolated at the position and time of altimeter measurements. First, the error metrics are computed for a single member from ALL simulations. We start with $\beta_{max} = 1.75$, as used by the reference study¹⁸, which displays high positive bias (not shown) and Nbias (Fig. 2d). After reducing β_{max} around 1, lower Nbias values are found. The area-averaged (60° S to 60° N) statistical values presented in Table 2 show the lowest bias (-0.04 m) and Nbias (-1.82%), when $\beta_{max} = 1$. These global average values are calculated after removing the outliers, from the spatial map, which are defined as values more than three scaled median absolute deviations from the median.

Following $\beta_{max} = 1.0$, all the members are shown to produce low bias (not shown here). Additionally, the standard deviations (of bias and Nbias) among the members are quite low. Here we acknowledge that the CMIP6 wind-forcing resolution is not adequate to reproduce the precise timing of storms²³. Hence the random error metrics (scatter index: SI) using 6-hourly outputs are usually large (around 40%), even though the model captures the long-term climate well. After the $\beta_{max} = 1.0$ chosen with this calibration, the model setting is retained for the entire simulation. Thereafter, the model simulations are validated through Nbias, SI, SC (spatial correlation), NSD (normalized standard deviation) and CRMSD (centered root mean squared difference) metrics using monthly mean values for the entire reference data duration. The statistical parameters, used in this study, are defined as follows:

$$bias = \frac{\sum(M - R)}{count} \quad (1)$$

$$Nbias = \frac{\sum(M - R)}{\sum R} \quad (2)$$

$$SI = \sqrt{\frac{\sum[(M - \bar{M}) - (R - \bar{R})]^2}{\sum R^2}} \quad (3)$$

$$SC = \frac{\sum_x(MR - \bar{M}\bar{R})}{\sum_x(M^2 - \bar{M}^2)(R^2 - \bar{R}^2)} \quad (4)$$

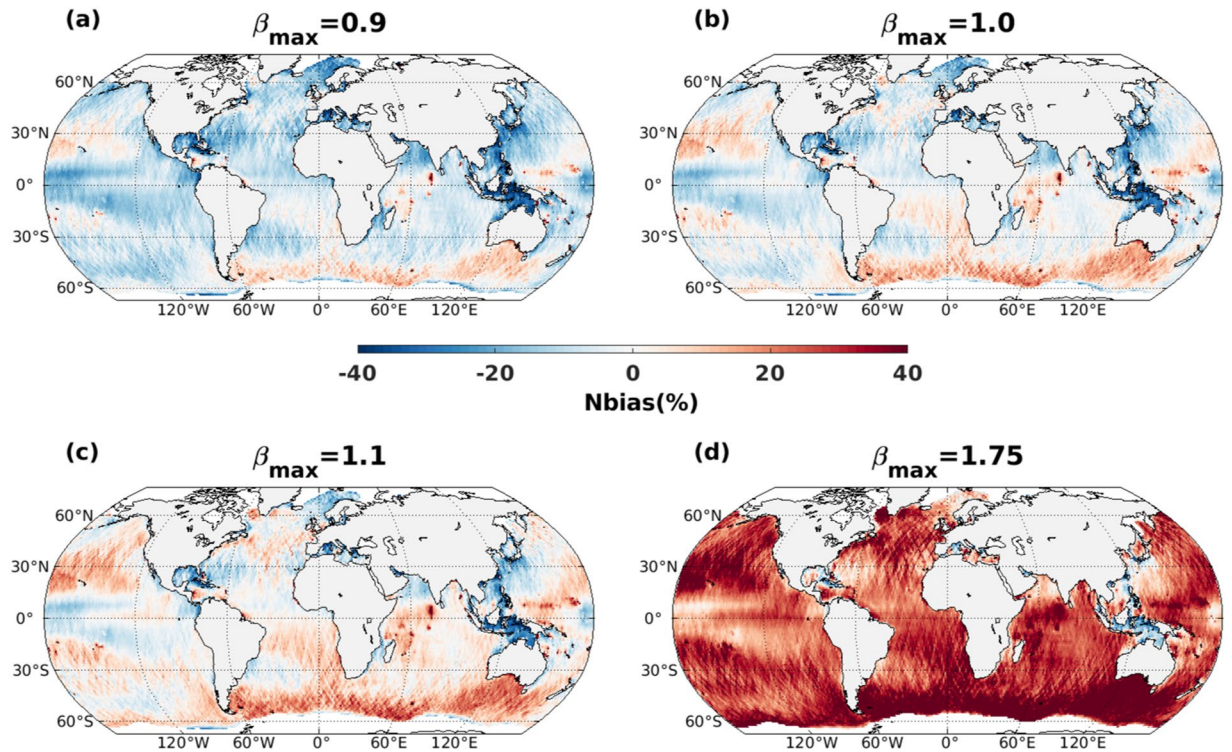


Fig. 2 Normalized bias (%) between CMIP6_WW3 SWH from a single member of ALL simulations (r3i1p1f1) and altimeter SWH, for four β_{max} values, averaged over the year 2011.

$$NSD = \frac{\sigma_M}{\sigma_R} \quad (5)$$

$$CRMSD = \sqrt{\frac{\sum [(M - \bar{M}) - (R - \bar{R})]^2}{count}} \quad (6)$$

where M and R are significant wave heights from model simulation and reference data, respectively. And σ is standard deviation.

Data Records

The full global wave climate dataset²⁴ can be accessed at <https://doi.org/10.12770/0983962b-4acc-4f8f-9484-e2195029b87b>.

The dataset consists of 6 hourly wave parameters (81 variables including significant wave height, mean wave periods, wave mean direction for total wave energy and partitions, among others), which are detailed in Supplementary Table 1, for 5 variants for each of the four CMIP6 historical experiments (ALL, NAT, GHG, AER) and three sets for CTL experiments (total = 23 runs) covering 60 years (1961–2020).

The 6-hourly outputs are saved into monthly files, so we have $12 \times 60 \times 23$ netcdf files.

The hierarchy for folders is set as: GLOBID_exp_id/0r/year/.

The gridded outputs are in the folder /FIELD_NC;

and the filenames are defined as: LOPS_WW3-glob-1d_YearMonth.nc, where Exp_id is the name of CMIP6 experiments, ALL, NAT, GHG, AER, CTL and 0r is the realization number of the variants (for example: 01 for r1i1p1f1, 02 for r2i1p1f1, and so on).

LOPS_WW3-glob-1d is WW3 global hindcast produced at LOPS for global ocean at 1° resolution.

The folders other than /FIELD_NC are:

“data/“ folder which has all the configuration files including bottom topography, namelist, etc.

“work/“ folder which has preprocessing file description used for each run.

“output/“ folder having output field description and others.

“SPEC_NC/“ folder having spectral energy information at the points listed in “/data/points_glob-1d.txt”.

Technical Validation

In order to use this dataset for future wave climate research, we evaluate the skill of our model simulations (ensemble mean of ALL) using various metrics with reference to ESA-CCI altimeter²² and ERA5²⁵ dataset. The skill of the wave database to capture the mean state is assessed through Nbias and SI maps from monthly mean data for annual and seasonal scales. At annual scale, the Nbias values are typically within 10% (Fig. 3), and

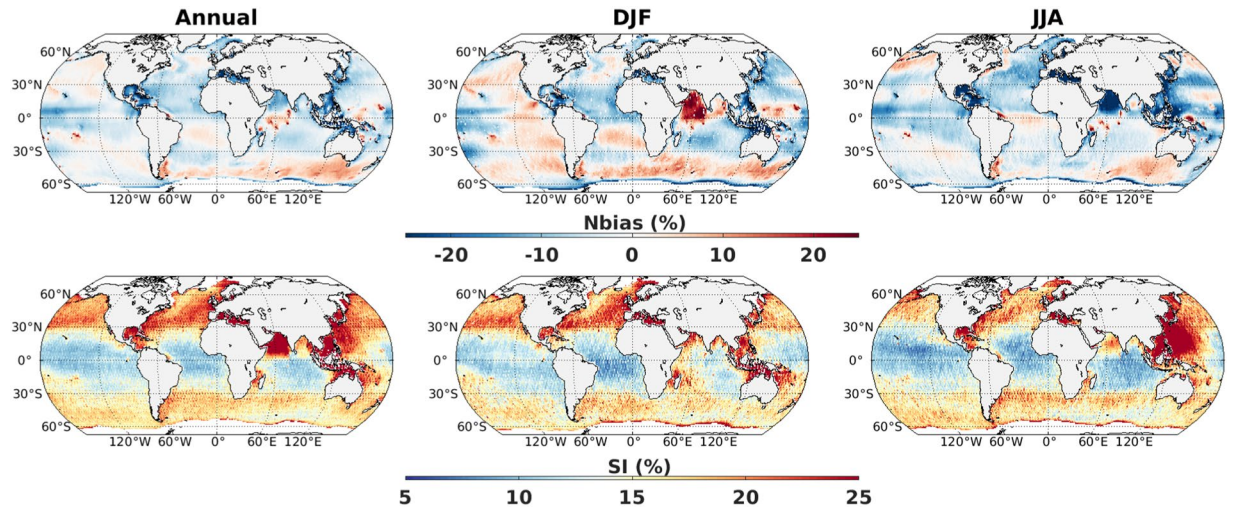


Fig. 3 Normalized bias (%) and scatter index (%) of SWH from CMIP6_WW3 ALL simulations (5 member mean) with reference to altimeter during 1992–2018 using monthly mean data. Left columns are considering all months, middle columns are for December-January-February and right columns are for June-July-August.

SI values are typically within 15%, with values higher than 15% mostly located in high latitude, coastal and enclosed seas. Overall, it represents good agreement with altimeter observations with global averaged values -3% and 17% for Nbias and SI, respectively. The Nbias values are significantly lower than the ones obtained for the calibration of the model (shown on Fig. 2), as we used monthly mean and ensemble mean of 5 members. In DJF and JJA seasons, the Nbias and SI values are of similar ranges, but with regional variations (Fig. 3). Exception is the Arabian Sea, which shows overestimation in DJF season and underestimation in JJA season by WW3 simulation compared to altimeter observation. It seems that MRI-ESM2.0 has difficulties in resolving the South Asian Monsoon accurately. In addition, few localized abnormally high bias values are scattered over tropical Indian Ocean which might occur due to unresolved sub-grid scale physics by our WW3 set-up with 1.0° horizontal resolution. Apparently, the current model discretization is coarse to capture island (Hawaii, Seychelles, Maldives, Chagos) blocking effect. It is noteworthy to mention that altimeter SWH measurements at low sea states (<0.75 m) is usually less accurate and noisier²⁶, which might also be the reason for high SI values in coastal areas. We see high SI values over the regions with strong current, specially western boundary currents. In this context, Alday *et al.*¹⁸ reported reduction in scatter index for the regions with stronger current (such as the Gulf stream, the Kuroshio, the Agulhas, etc.) when current is added as forcing to wave model. Unfortunately, MRI-ESM2.0 does not provide data for ocean currents. Note that, although MRI-ESM2.0 has improved over the previous versions and demonstrates realistic aspects of many climate components, still it has drawbacks in terms of double intertropical convergence zone (ITCZ) bias, cold biases in the Southern Hemisphere (SH) surface air temperature, among others¹⁵. The inconsistency between model and observation can arise due to combination of multiple factors, including forcing field, parameterization, model discretization, and altimeter measurement errors. Here, model discretizations are a balance between the computational cost and the level of accuracy. The 60 years of hindcast at 6-hourly interval for 5 CMIP6 experiments, each having 5 variants, already demands a large computation storage.

A slightly stronger Nbias and reduced SI values are seen when the model simulations are compared with ERA5 reanalysis (Fig. 4). Interestingly, spatial structure (positive and negative errors) and seasonal variation of Nbias pattern are similar for both reference data set (ESA-CCI altimeter and ERA5), which is partly due to the assimilation of altimeter data in ERA5. More specifically, estimates of Nbias computed with ERA5 show higher overestimation over the mid-latitudes and lesser underestimation over the intertropical regions compared to Nbias computed with altimeter, in agreement with the known differences between altimeter and ERA5²². The lower scattered index estimates can be explained by smoother ERA5 values compared to altimeter observations. The global averaged values are 2% and 12% for Nbias and SI, respectively.

It is important to acknowledge that differences in altimeter data processing methods may induce significant differences between multi-mission altimeter products^{27,28}. Moreover, different global wave hindcasts and/or wave reanalysis often show disparate results regarding climatology, variability; as evidenced by Morim *et al.*²⁹ and Erikson *et al.*³⁰. As a consequence, the calibration and validation results may potentially vary depending on the selected reference dataset (observations or model). Since it is not the objective of this paper to compare different altimeter and reanalysis products, the reader could refer to Morim *et al.*²⁹ and Erikson *et al.*³⁰ for comparative assessments of reference altimeter and model datasets.

As expected, higher error is seen for monthly maxima of SWH than for monthly mean values when compared with altimeter (Figure not shown). Model overestimates almost everywhere with larger values over the mid-latitudes of both hemisphere, and global averaged values are around 20% . The spatial pattern of SI map bears resemblance to that of monthly mean values (Fig. 3), but with amplified magnitude. As acknowledged in Young and Ribal³¹ and Timmermans *et al.*²⁷, differences in extremes between products are more likely to be

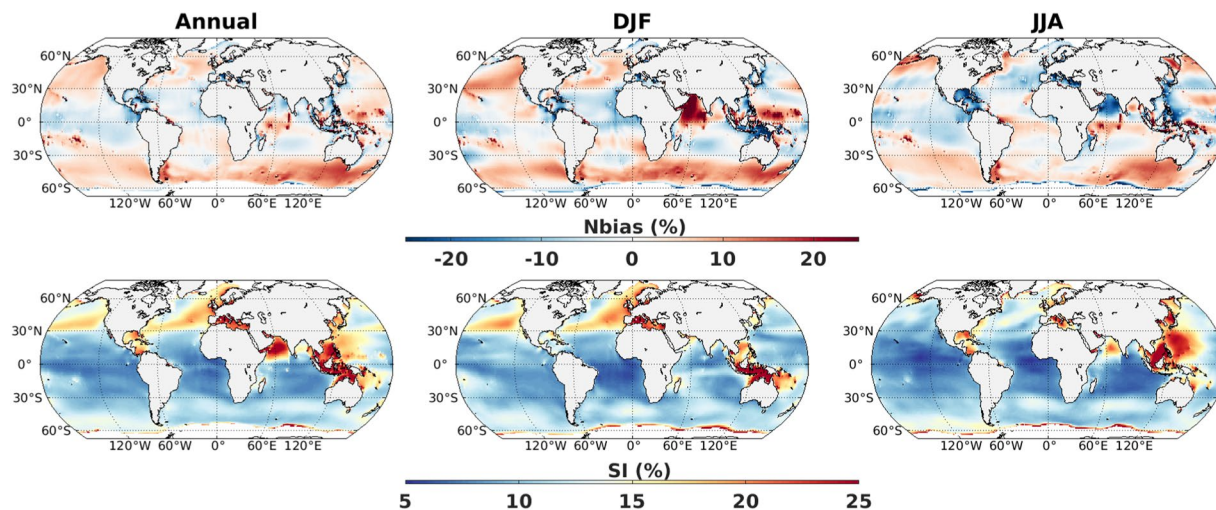


Fig. 4 Normalized bias (%) and scatter index (%) of SWH from CMIP6_WW3 ALL simulations (5 member mean) with reference to ERA5 during 1980–2020 using monthly mean data. Left columns are considering all months, middle columns are for December–January–February and right columns are for June–July–August.

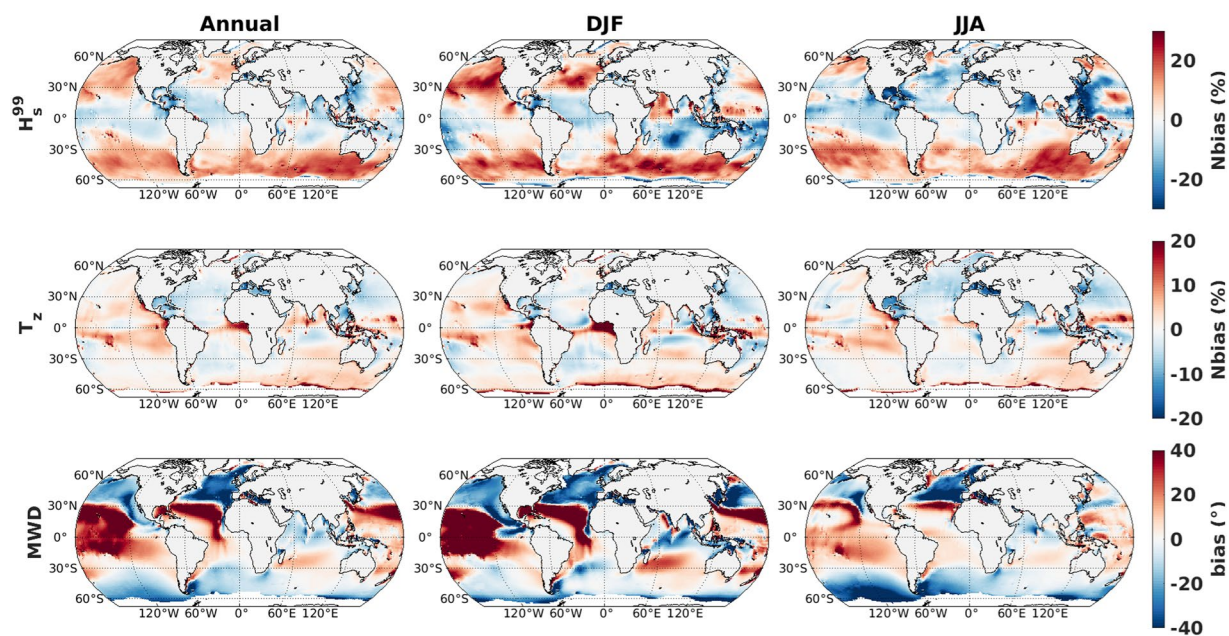


Fig. 5 Normalized bias (%) of 99th percentile SWH (H_s^{99}) and mean zero-crossing wave period (T_z); and bias of mean wave direction (MWD) from CMIP6_WW3 ALL simulations (5 member mean) with reference to ERA5 during 1980–2020 using monthly data. Left columns are considering all months, middle columns are for December–January–February and right columns are for June–July–August.

higher than in mean values. These large errors can be partially explained by the low revisit time period of altimeter missions, causing significant undersampling errors in multi-mission altimeter products³², which poorly resolve the extreme storm events.

Given the noisiness of maximum wave heights, monthly 99th percentile SWHs (H_s^{99}) are also validated with reference to ERA-5 wave height (Fig. 5). The spatial pattern of Nbias for H_s^{99} shows resemblance to that of monthly mean values with higher overestimation in the storm track regions (high latitudes). Notable negative biases prevail over the Arabian Sea and east China sea during summer monsoon season. Contrasting relative difference for wave period (T_z), the Nbias values are within 10%, except for east tropical Atlantic. However, limited overestimation is evident along the tropical latitudes of both hemispheres, in the so-called swell pools³³. The mean wave direction from model shows positive (clockwise) biases along the tropical band, and slightly lower negative

Reference	SC	NSD	CRMSD	Nbias%	SI%
ESA-CCL_ H_s (m)	0.99	0.61	0.45	-3.1	16.8
ESA-CCI H_s^{max} (m)	0.99	0.69	1.10	24.8	26.2
ERA5_ H_s (m)	0.99	0.66	0.33	2.4	12.6
ERA5 H_s^{99} (m)	0.99	0.89	1.04	3.1	20.5
ERA5_ T_z (s)	0.95	0.62	0.48	2.2	6.8
ERA5_ MWD ($^{\circ}$)	0.96	0.52	35.01	0.9	N/A

Table 3. Global averaged error metrics for CMIP6_WW3 ALL simulations (5-member mean) for reference data period (ESA-CCI :1992–2018 and ERA5: 1980–2020). Note that for MWD, the absolute bias value is provided (in $^{\circ}$), instead of the normalized bias (%).

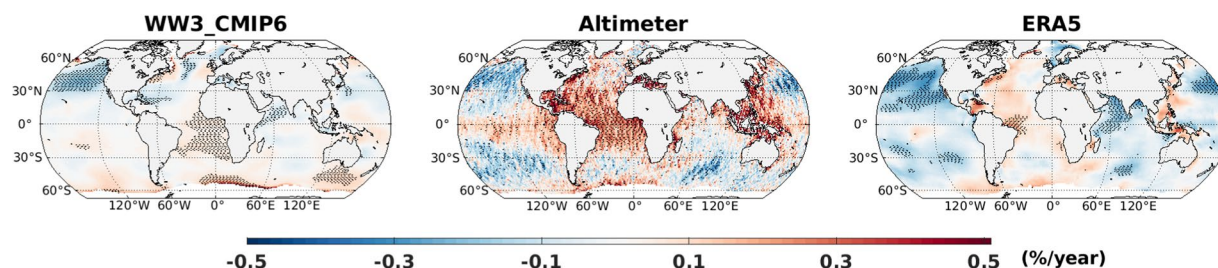


Fig. 6 Linear trend (%) expressed as percentage change per year during 1992–2018 for SWH from CMIP6_WW3 ALL simulations, ESA-CCI altimeter, and ERA5 dataset.

(anticlockwise) biases over the high-latitudes. The clockwise rotation can be explained by strong westerly component along the storm belt and the anticlockwise rotation by strong southerly component associated with strong swells from Southern Ocean; and similar pattern is reported by Lemos *et al.*³⁴ when comparing with ERA5 dataset. Overall, our model simulations can reproduce the major spatial patterns and seasonal variations as in ERA5. The global average values for normalized bias for H_s^{99} , T_z and bias for MWD (using ERA5 as reference) are 3.07(%), 2.2(%), and 0.9 ($^{\circ}$), respectively (Table 3). In addition to Nbias and SI, metrics such as SC, NSD, and CRMSD are also provided; as in COWCLIP framework²³. The SC values are close to unity, while NSD values are 0.5–0.9. The CRMSD values are 1.04 m, 0.48 s, 35.01 $^{\circ}$ for H_s^{99} , T_z , and MWD, respectively. Moreover, these metrics are comparable with Song *et al.*¹².

To assess multi-annual variability, comparison of linear trend in annual mean wave height is presented in Fig. 6. Generally, the percentage change per year among three datasets agree on spatial distribution and sign of trends. Quantitatively, the range of the trends are also in agreements, although the altimeter trends are slightly stronger in certain regions (e.g. tropical Atlantic). Strong negative trends in the North Pacific is clearly consistent among the datasets. Other regions of consistent trends include Atlantic and eastern Pacific sector of Southern Ocean and tropical Atlantic. Nevertheless, certain differences in trends among three dataset are evident over Arabian Sea and mid-latitudes of south Pacific.

Furthermore, we validate the ability of CMIP6-forced wave model to represent seasonal cycle. The modeled seasonality (difference between January and July wave height) is strongest in Northern Hemisphere (NH) extra-tropical storm belt with lower values over tropics (Fig. 7a). In SH, it is strongest over Indian Ocean sector of Southern Ocean, which is in agreement with Hemer and Trenham³⁵. Looking at the bias in seasonality (Fig. 7b), it is encouraging that difference of seasonal amplitude between CMIP6 forced simulation and altimeter is generally lower than 0.5 m, except for the Arabian Sea and certain sectors of the Southern Ocean, which coincides with regions with strong bias (Fig. 3). Overall, the reduced agreement in Southern Ocean can be because of combination of factors, such as presence of circumpolar current, presence of both sea-ice and icebergs, the parameterizations of wave-ice interactions. In a quantitative view, difference in seasonal amplitude is reasonable with a low global mean value (-0.12 m). The temporal correlation map (using monthly mean values) reveal statistically significant relation almost everywhere in global ocean, with higher values (>0.7) in NH (Fig. 7c). Surprisingly, low to no correlation is seen over narrow tropical band and over a narrow latitudinal band close to Southern Ocean. These regions having low correlation coincides with regions having low seasonal or intra-annual variability (Fig. 7a). The spatial patterns of simulated dataset correlated well with the spatial pattern of altimeter observation over the years (blue curve on Fig. 7d), always higher than 0.75 except for the year 1992, probably because of altimeter undersampling issue³². To further evaluate spatial variability, difference in spatial standard deviation between WW3 simulation and altimeter is presented (red curve on Fig. 7d). The spatial variance of simulated dataset are generally within the same range as the observation, to be more specific, model variance is slightly stronger than observation since the year 2000. To summarise, this WW3 simulations produce considerably realistic representation of the historical global wave climate.

To present a brief overview of our simulated wave data, 60-year mean climatology for all the CMIP6 experiments: ALL, GHG, AER, and NAT are displayed in Fig. 8. Note that for GHG, AER and NAT scenarios, anomaly

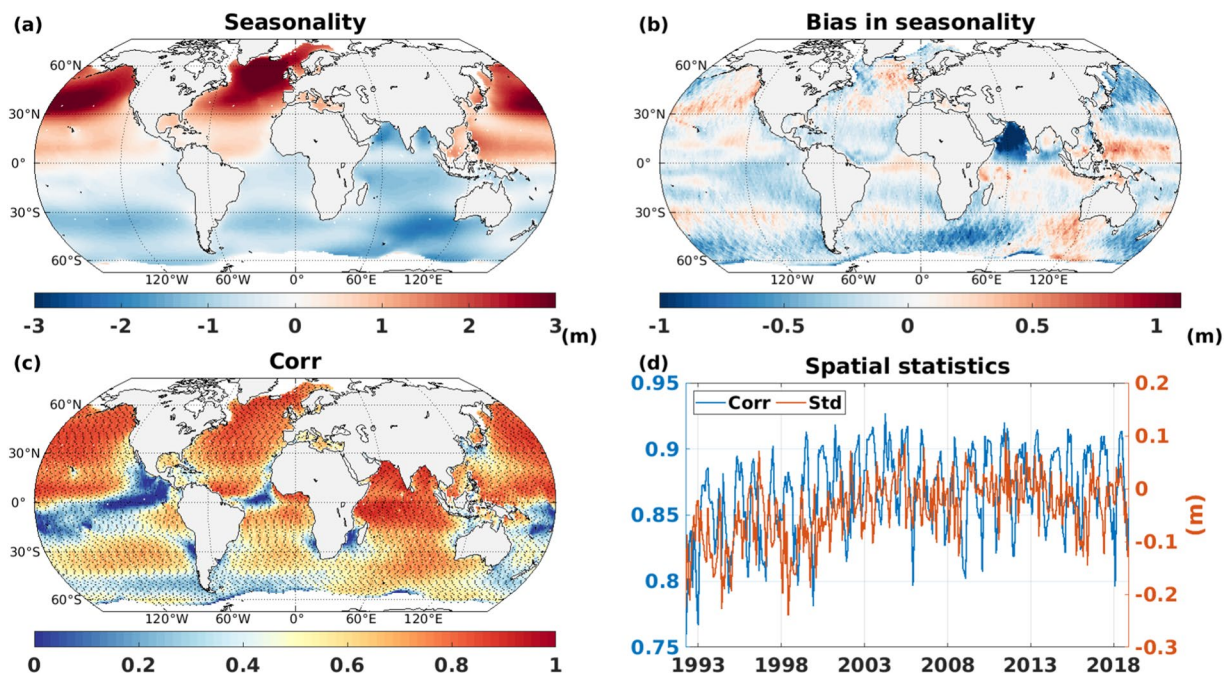


Fig. 7 (a) Seasonality (January minus July) of modeled SWH (m) from ALL simulations, (b) Difference between ALL simulations and altimeter seasonality (m), (c) Time correlation (Pearson correlation coefficient) between ALL simulation and altimeter SWH with black dots showing statistical significance at 5% level, (d) spatial correlation coefficient (blue) and difference in spatial standard deviation (red) between ALL simulation and altimeter SWH (m) during 1992–2018.

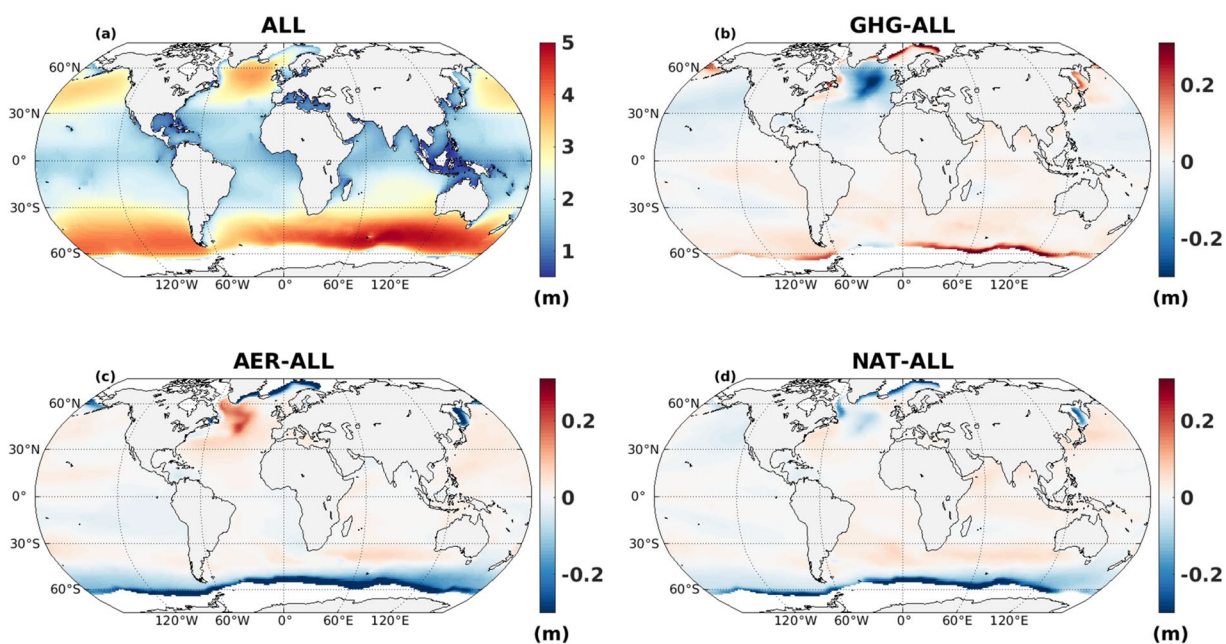


Fig. 8 (a) Climatology of SWH (m) from ALL; and anomaly of climatology (with respect to ALL) for (b) GHG, (c) AER, and (d) NAT simulation during 1961–2020 (in m).

of climatology (60-year mean of GHG/AER/NAT - 60-year mean of ALL) with respect to ALL simulation is displayed. The ALL simulation clearly shows a global zonal distribution with larger values over extra-tropics and lower values in the tropics. Anomaly maps corresponding to different forcings differ substantially from each other and obviously from ALL forcing. In NAT, the anomaly amplitude is weaker than GHG and AER close to the Arctic circle. In the Southern Ocean, negative anomaly exists for AER and NAT in contrast to GHG positive anomaly.

It is noteworthy that stronger anomalies are seen in the proximity of sea-ice. The contrasting behaviors among the difference forcing scenario is quite intriguing. However, the potential reason behind such pattern is not in the scope of current study but clearly warrants further research. Detailed attribution of the differences between datasets would be desirable, through understanding skill of the forcing obtained from climate model, although many uncertainty remains in the reference observations²⁷. We acknowledge that, our wave product is using single-method modelling, so uncertainty arising from different global wave products should be considered when assessing coastal hazards/risk for extremes³⁶. Our dataset will be of service to future research to clarify how external anthropogenic and natural forcing have influenced historical wave climate changes.

Code availability

The WAVEWATCH III setup files are available in the data link (https://data-dataref.ifremer.fr/ww3/GLOBID_CMIP6/)²⁴ for the purpose of replicating the data described in this paper. The “wavesetup.env” file can be found in the following repository of the above link: GLOBID_exp_id/0r/year/. Please refer to Data Records section for more description of the repository.

Received: 9 November 2022; Accepted: 11 May 2023;

Published online: 26 May 2023

References

- Grifoll, M., Borén, C. & Castells-Sanabra, M. A comprehensive ship weather routing system using CMEMS products and A* algorithm. *Ocean Engineering* **255**, 111427, <https://doi.org/10.1016/j.oceaneng.2022.111427> (2022).
- Dean, R. G. & Dalrymple, R. A. *Coastal Processes with Engineering Applications* (Cambridge University Press, Cambridge, 2001).
- Sierra, J. P. & Casas-Prat, M. Analysis of potential impacts on coastal areas due to changes in wave conditions. *Climatic Change* **124**, 861–876, <https://doi.org/10.1007/s10584-014-1120-5> (2014).
- Kamranzad, B., Lin, P. & Iglesias, G. Combining methodologies on the impact of inter and intra-annual variation of wave energy on selection of suitable location and technology. *Renewable Energy* **172**, 697–713, <https://doi.org/10.1016/j.renene.2021.03.062> (2021).
- Cavaleri, L., Fox-Kemper, B. & Hemer, M. Wind Waves in the Coupled Climate System. *Bulletin of the American Meteorological Society* **93**, 1651–1661, <https://doi.org/10.1175/BAMS-D-11-00170.1> (2012).
- Hemer, M. A., Wang, X. L., Weisse, R. & Swail, V. R. Advancing Wind-Waves Climate Science: The COWCLIP Project. *Bulletin of the American Meteorological Society* **93**, 791–796, <https://doi.org/10.1175/BAMS-D-11-00184.1> (2012).
- Morim, J. *et al.* A global ensemble of ocean wave climate projections from CMIP5-driven models. *Scientific Data* **7**, 105, <https://doi.org/10.1038/s41597-020-0446-2> (2020).
- Najafi, M. R., Zwiers, F. W. & Gillett, N. P. Attribution of Arctic temperature change to greenhouse-gas and aerosol influences. *Nature Climate Change* **5**, 246–249, <https://doi.org/10.1038/nclimate2524> (2015).
- Bilbao, R. A. F., Gregory, J. M., Bouttes, N., Palmer, M. D. & Stott, P. Attribution of ocean temperature change to anthropogenic and natural forcings using the temporal, vertical and geographical structure. *Climate Dynamics* **53**, 5389–5413, <https://doi.org/10.1007/s00382-019-04910-1> (2019).
- Eyring, V. *et al.* Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development* **9**, 1937–1958, <https://doi.org/10.5194/gmd-9-1937-2016> (2016).
- Gillett, N. P. *et al.* The Detection and Attribution Model Intercomparison Project (DAMIP v1.0) contribution to CMIP6. *Geoscientific Model Development* **9**, 3685–3697, <https://doi.org/10.5194/gmd-9-3685-2016> (2016).
- Song, Z. *et al.* Centuries of monthly and 3-hourly global ocean wave data for past, present, and future climate research. *Scientific Data* **7**, 226, <https://doi.org/10.1038/s41597-020-0566-8>. Number: 1 Publisher: Nature Publishing Group (2020).
- Badriana, M. R. & Lee, H. S. Multimodel Ensemble Projections of Wave Climate in the Western North Pacific Using CMIP6 Marine Surface Winds. *Journal of Marine Science and Engineering* **9**, 835, <https://doi.org/10.3390/jmse9080835>. Number: 8 Publisher: Multidisciplinary Digital Publishing Institute (2021).
- Kumar, R., Lemos, G., Semedo, A. & Alsaqa, F. Parameterization-Driven Uncertainties in Single-Forcing, Single-Model Wave Climate Projections from a CMIP6-Derived Dynamic Ensemble. *Climate* **10**, 51, <https://doi.org/10.3390/cli10040051>. Number: 4 Publisher: Multidisciplinary Digital Publishing Institute (2022).
- Yukimoto, S. *et al.* The Meteorological Research Institute Earth System Model Version 2.0, MRI-ESM2.0: Description and Basic Evaluation of the Physical Component. *Journal of the Meteorological Society of Japan. Ser. II* **97**, 931–965, <https://doi.org/10.2151/jmsj.2019-051> (2019).
- O'Neill, B. C. *et al.* The Scenario Model Intercomparison Project (ScenarioMIP) for CMIP6. *Geoscientific Model Development* **9**, 3461–3482, <https://doi.org/10.5194/gmd-9-3461-2016> (2016).
- Ardhuin, F. *et al.* Semiempirical Dissipation Source Functions for Ocean Waves. Part I: Definition, Calibration, and Validation. *Journal of Physical Oceanography* **40**, 1917–1941, <https://doi.org/10.1175/2010JPO4324.1> (2010).
- Alday, M., Accensi, M., Ardhuin, F. & Dodet, G. A global wave parameter database for geophysical applications. Part 3: Improved forcing and spectral resolution. *Ocean Modelling* **166**, 101848, <https://doi.org/10.1016/j.ocemod.2021.101848> (2021).
- Hasselmann, S. & Hasselmann, K. Computations and Parameterizations of the Nonlinear Energy Transfer in a Gravity-Wave Spectrum. Part I: A New Method for Efficient Computations of the Exact Nonlinear Transfer Integral. *Journal of Physical Oceanography* **15**, 1369–1377, [10.1175/1520-0485\(1985\)015<1369:CAPOTN>2.0.CO;2](https://doi.org/10.1175/1520-0485(1985)015<1369:CAPOTN>2.0.CO;2) (1985).
- Stopa, J. E., Ardhuin, F., Stutzmann, E. & Lecocq, T. Sea State Trends and Variability: Consistency Between Models, Altimeters, Buoys, and Seismic Data (1979–2016). *Journal of Geophysical Research: Oceans* **124**, 3923–3940, <https://doi.org/10.1029/2018JC014607>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1029/2018JC014607> (2019).
- Piollé, J.-F., Dodet, G. & Quilfen, Y. ESA Sea State Climate Change Initiative (Sea_state_cci): Global remote sensing merged multi-mission monthly gridded significant wave height, L4 product, version 1.1, <https://doi.org/10.5285/47140D618DCC40309E1EDBCA7E773478>. Medium: application/xml Type: dataset (2020).
- Dodet, G. *et al.* The Sea State CCI dataset v1: towards a sea state climate data record based on satellite observations. *Earth System Science Data* **12**, 1929–1951, <https://doi.org/10.5194/essd-12-1929-2020> (2020).
- Morim, J. *et al.* Robustness and uncertainties in global multivariate wind-wave climate projections. *Nature Climate Change* **9**, 711–718, <https://doi.org/10.1038/s41558-019-0542-5> (2019).
- Patra, A., Accensi, M. & Dodet, G. *GLOBID_cmip6*: Historical global ocean wave data simulated with CMIP6 anthropogenic and natural forcings. *IFREMER* <https://doi.org/10.12770/0983962b-4acc-4f8f-9484-e2195029b87b> (2022).
- Hersbach, H. *et al.* The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society* **146**, 1999–2049, <https://doi.org/10.1002/qj.3803>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/qj.3803> (2020).
- Smith, W. H. F. & Scharroo, R. Waveform Aliasing in Satellite Radar Altimetry. *IEEE Transactions on Geoscience and Remote Sensing* **53**, 1671–1682, <https://doi.org/10.1109/TGRS.2014.2331193>. Conference Name: IEEE Transactions on Geoscience and Remote Sensing (2015).

27. Timmermans, B. W., Gommenginger, C. P., Dodet, G. & Bidlot, J. Global Wave Height Trends and Variability from New Multimission Satellite Altimeter Products, Reanalyses, and Wave Buoys. *Geophysical Research Letters* **47**, <https://doi.org/10.1029/2019GL086880> (2020).
28. Ribal, A. & Young, I. R. 33 years of globally calibrated wave height and wind speed data based on altimeter observations. *Scientific Data* **6**, 77, <https://doi.org/10.1038/s41597-019-0083-9>. Number: 1 Publisher: Nature Publishing Group (2019).
29. Morim, J. *et al.* A global ensemble of ocean wave climate statistics from contemporary wave reanalysis and hindcasts. *Scientific Data* **9**, 358, <https://doi.org/10.1038/s41597-022-01459-3> (2022).
30. Erikson, L. *et al.* Global ocean wave fields show consistent regional trends between 1980 and 2014 in a multi-product ensemble. *Communications Earth & Environment* **3**, 320, <https://doi.org/10.1038/s43247-022-00654-9> (2022).
31. Young, I. R. & Ribal, A. Multiplatform evaluation of global trends in wind speed and wave height. *Science* **364**, 548–552, <https://doi.org/10.1126/science.aav9527> (2019).
32. Jiang, H. Evaluation of altimeter undersampling in estimating global wind and wave climate using virtual observation. *Remote Sensing of Environment* **245**, 111840, <https://doi.org/10.1016/j.rse.2020.111840> (2020).
33. Semedo, A., Sušelj, K., Rutgersson, A. & Sterl, A. A Global View on the Wind Sea and Swell Climate and Variability from ERA-40. *Journal of Climate* **24**, 1461–1479, <https://doi.org/10.1175/2010JCLI3718.1>. Place: Boston MA, USA Publisher: American Meteorological Society (2011).
34. Lemos, G., Semedo, A., Dobrynin, M., Menendez, M. & Miranda, P. M. A. 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**, 1393–1414, <https://doi.org/10.1175/JAMC-D-19-0297.1>. Place: Boston MA, USA Publisher: American Meteorological Society (2020).
35. Hemer, M. A. & Trenham, C. E. Evaluation of a CMIP5 derived dynamical global wind wave climate model ensemble. *Ocean Modelling* **103**, 190–203, <https://doi.org/10.1016/j.ocemod.2015.10.009> (2016).
36. Morim, J. *et al.* Understanding uncertainties in contemporary and future extreme wave events for broad-scale impact and adaptation planning. *Science Advances* **9**, eade3170, <https://doi.org/10.1126/sciadv.ade3170> (2023).

Acknowledgements

AP received a post-doctoral grant from IFREMER. This research is part of the Sea State CCI project of the Climate Change Initiative (CCI) (ESA ESRIN, contract no. 4000123651/18/I-NB). The authors acknowledge the World Climate Research Programme which coordinated CMIP6. We thank the climate modelling group of MRI-ESM2.0. ESA Sea State CCI L2P product (v1.1) is available on <https://climate.esa.int/en/projects/sea-state/data/>.

Author contributions

A.P. conceived the experiment, executed the model, analyzed the simulation, and wrote the manuscript. G.D. contributed to methodology and writing manuscript. M.A. contributed to numerical simulations. All the authors discussed during the experiment.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41597-023-02228-6>.

Correspondence and requests for materials should be addressed to A.P.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023