



OPEN

DATA DESCRIPTOR

# Global 3-hourly wind-wave and swell data for wave climate and wave energy resource research from 1950 to 2100

Xingjie Jiang<sup>1,2,3</sup>, Botao Xie<sup>4</sup>, Ying Bao<sup>1,2,3</sup> & Zhenya Song<sup>1,2,3</sup>

Ocean wave climate, including wind waves and swells, is essential to human marine activities and global or regional climate systems, and is highly related to harnessing wave energy resources. In this study, a global 3-hourly instantaneous wave dataset was established with the third-generation wave model MASNUM-WAM and wind forcings derived from the products of the First Institute of Oceanography-Earth System Model version 2.0, the climate model coupled with wave model, under the unified framework of the Coupled Model Intercomparison Project phase 6. This dataset contains 17 wave parameters, including the information associated with wave energy and spectral shape geometries, from one historical (1950–2014) simulation and three future (2015–2100) scenario experiments (ssp125, ssp245, and ssp585). Moreover, all the parameters can be accessed separately in the form of wind waves and swells. The historical results show that the simulated wave characteristics agree well with satellite observations and the ERA5 reanalysis products. This dataset can provide the community with a unique and informative data source for wave climate and wave energy resource research.

## Background & Summary

Ocean waves, arriving from specific wind events that are occurring locally (wind waves) or have occurred somewhere else on the sea surface (swells), can reach tens of meters in height or travel thousands of miles and can bring serious threats to various marine activities, such as sea voyages<sup>1–3</sup>, ocean fishing<sup>4,5</sup>, and oil exploitation<sup>6–8</sup>. Ocean waves can also be intimately involved in the energy and material exchange between the atmosphere and ocean, playing a crucial role in global and regional climate systems<sup>9,10</sup>. Therefore, understanding the wave climate and change is valuable for offshore engineering structure safety, shoreline protection, global warming prevention, etc. Moreover, harnessing wave energy (the most concentrated and high-available source of marine renewable energy with great potential for exploitation<sup>11,12</sup>) needs to consider its annual or seasonal spatial distributions and temporal variability, which is highly correlated with the wave climatology of the target area.

In studies on wave climate, especially in those interested in future climatic scenarios, simulation with numerical wave models is the main method used. With specific wind forcings, the simulation can generate long-term wave parameters with continuous coverage in space and time, which can be applied in further analyses. A series of wave datasets are proposed to support such research, e.g., the earlier proposed ERA-Interim<sup>13</sup> provides the basic bulk wave parameters, such as significant wave height, mean wave period, and mean wave direction. The later presented ERA5 reanalysis<sup>14–16</sup> and EMC/NCEP wave hindcast<sup>17</sup> datasets can further exhibit those wave parameters above in the forms of wave spectral partitions, i.e., wind waves and swells, and the ERA5 dataset can also provide the parameters associated with spectral shape geometries, which can be adopted in extreme wave event analyses<sup>18–21</sup>. In addition to the reanalysis and hindcast datasets, wave data products from global and regional climate models can be helpful in understanding the response of the global wave climate to both historical and future climate change. For example, the First Institute of Oceanography-Earth System Model version 2.0 (FIO-ESM v2.0)<sup>22</sup> was used to carry out Coupled Model Intercomparison Project

<sup>1</sup>First Institute of Oceanography, and Key Laboratory of Marine Science and Numerical Modeling, Ministry of Natural Resources, Qingdao, 266061, China. <sup>2</sup>Laoshan Laboratory, Qingdao, 266237, China. <sup>3</sup>Shandong Key Laboratory of Marine Science and Numerical Modeling, Qingdao, 266061, China. <sup>4</sup>China National Offshore Oil Corporation Research Institute, Beijing, 100028, China. ✉e-mail: [songroy@fio.org.cn](mailto:songroy@fio.org.cn)

phase 6 (CMIP6)<sup>23,24</sup> experiments, and a global monthly and 3-hourly ocean wave dataset spanning centuries was produced<sup>25</sup>. Through the Coordinated Ocean Wave Climate Project (COWCLIP)<sup>10</sup> phase 2, a global ensemble of ocean wave climate projections from CMIP5-driven<sup>26</sup> models could be accessed<sup>27</sup>. Similarly, through COWCLIP2, a global ensemble of ocean wave climate statistics from contemporary reanalysis and hindcasts was also released recently<sup>28</sup>.

However, in those wave datasets derived from climate models, wave parameters associated with the wave energy resources, as well as the information related to shape geometries of wave spectra may be less considered, and detailed wave characters from spectral partitions are also unavailable. In this study, the wind products simulated by FIO-ESM v2.0 under the unified framework of CMIP6 were adopted again to force a standalone wave model, and then, a global 3-hourly wave dataset was generated. The dataset described here contains up to 17 wave parameters, including the characteristics associated with wave energy and spectral geometries, from one historical (1950–2014) simulation and three future (2015–2100) scenario experiments. Moreover, all the wave parameters can be accessed separately in the forms of wind waves and swells. Therefore, the newly proposed dataset differs from the other wave datasets mentioned above and can provide the community with a unique data source for wave climate and wave energy resource research.

## Methods

This dataset was established with the third-generation wave model MASNUM-WAM and wind forcings derived from the products of FIO-ESM v2.0 under the framework of CMIP6. In this section, we introduce the wave model and the wind forcings adopted in this study. Brief introductions to calculating wave parameters and identifying wind waves and swells are also presented.

**MASNUM-WAM and modeling configuration.** The MASNUM-WAM (formerly LAGFD-WAM)<sup>29–32</sup> is a third-generation wave model developed by the Key Laboratory of Marine Science and Numerical Modeling (MASNUM), FIO of MNR (Ministry of Nature Resources) of China. MASNUM-WAM solves the energy spectrum balance equation in wavenumber space and uses a complicated characteristic inlaid scheme<sup>29</sup> in spherical coordinates<sup>31</sup> to perform shoaling and refraction effects in shallow waters, the modulation of background current to wave evolution, and the refraction of waves propagating along great circles.

In this work, the ST6 source function package<sup>33–35</sup> is adopted to simulate the effects of wind input, white-capping dissipation, and swell dissipation on the evolution of waves, and the DIA<sup>36,37</sup> scheme is adopted to calculate the nonlinear energy transfer between waves. A global computational grid is used in the simulation, covering the region from 80°S to 80°N and 0°(360°)E to 359°E with a 1° × 1° horizontal resolution. The modeling spectral space is set as 24 directions with intervals of 15° and 35 wavenumbers spaced logarithmically from the minimum of 0.0071 up to 4.6341 with intervals of  $k_{i+1}/k_i = 1.21$ , which are equivalent to frequencies from 0.042 Hz to 1.073 Hz with a ratio of 1.1 at infinite depth. Finally, bathymetric data are obtained from ETOPO1<sup>38</sup> of the National Oceanic and Atmospheric Administration (NOAA) National Geophysical Data Centre (NGDC).

**FIO-ESM v2.0 and the winds from CMIP6 experiments.** FIO-ESM v2.0 is a global climate model that consists of five coupled components: an atmosphere general circulation model (Community Atmosphere Model version 5<sup>39</sup>), a land surface model (Community Land Model version 4.0<sup>40</sup>), an ocean general circulation model (Parallel Ocean Program version 2<sup>41</sup>), an ocean surface wave model (MASNUM-WAM<sup>29–32</sup>), and a sea ice model (Los Alamos sea ice model version 4<sup>42</sup>). FIO-ESM v2.0 has considered four distinctive physical processes, including nonbreaking surface wave-induced vertical mixing<sup>43,44</sup>, the effects of Stokes drift on momentum and heat fluxes, the effects of sea spray on heat flux, and the SST diurnal cycle.

Recently, FIO-ESM v2.0 was used to carry out CMIP6 experiments<sup>23</sup>. The wind data products are derived from the CMIP6 historical data, and then, the three future scenario experiments are adopted as driving forcings in this work. The historical simulation represents climate change over the 1850–2014 period, and the future scenario experiments, which belong to the CMIP6-Endorsed Scenario Model Intercomparison Project (ScenarioMIP)<sup>24</sup>, are the projections of future (during 2015–2100) climate change. In this work, the three future scenarios are forced by the latest proposed shared socioeconomic pathways (SSPs), denoted ssp126, ssp245, and ssp585, representing the low, medium, and high ends of the range of future forcing pathways to produce radiative forcings of 2.6 W m<sup>−2</sup>, 4.5 W m<sup>−2</sup>, and 8.5 W m<sup>−2</sup> in 2100, respectively. All four experiments mentioned above were forced by the forcing datasets provided by CMIP6 (<https://esgf-node.llnl.gov/search/input4mips/>).

The derived wind forcing data are the zonal and meridional wind velocities at 10 m above the sea surface. The horizontal resolution is a finite volume grid (approximately 0.9° × 1.25°), and the time resolution is three hours.

**Wave parameters.** As a spectral model adopted in the simulation, wave parameters can be obtained by integrating the simulated 2-D wave spectra. The 2-D wavenumber-direction spectra simulated in MASNUM-WAM can be expressed as  $E(k, \theta)$ , which can be easily transformed to the frequency-direction energy spectra  $S(f, \theta)$  as follows:

$$S(f, \theta) = \frac{1}{2\pi} c_g E(k, \theta). \quad (1)$$

In Eq. (1),

$$c_g = \frac{2\pi f}{k} \left( 1 + \frac{2kd}{\sinh(2kd)} \right) \quad (2)$$

is the group velocity of waves, and wavenumber  $k$  associated with given frequency  $f$  and water depth  $d$  (unit: m) is defined implicitly through the following dispersion relationship:

$$(2\pi f)^2 = gk \tanh(kd). \quad (3)$$

Then, the one-dimensional spectrum  $S(f)$  can be obtained as follows:

$$S(f) = \int_0^{2\pi} S(f, \theta) d\theta, \quad (4)$$

and the  $n$ -th spectral moment can be expressed as follows:

$$m_n = \int_0^\infty f^n S(f) df. \quad (5)$$

The bulk wave parameters, such as the significant wave height  $H_s$ , the wave energy period  $T_e$ , the mean zero-crossing wave period  $T_{m02}$ , and the mean wave period  $T_{m01}$ , can be obtained directly via the spectral moments as follows:

$$H_s = 4\sqrt{m_0} \quad (6)$$

$$T_e = \frac{m_{-1}}{m_0} \quad (7)$$

$$T_{m02} = \sqrt{\frac{m_0}{m_2}} \quad (8)$$

$$T_{m01} = \frac{m_0}{m_1} \quad (9)$$

The mean wave direction  $\theta_m$  can be expressed as follows:

$$\theta_m = \tan^{-1}(b/a) \quad (10)$$

in which

$$\begin{cases} a = \int_0^{2\pi} \int_0^\infty \cos\theta S(f, \theta) df d\theta \\ b = \int_0^{2\pi} \int_0^\infty \sin\theta S(f, \theta) df d\theta \end{cases}. \quad (11)$$

Parameters associated with spectral shape geometries can also be obtained directly from the simulated spectra, e.g., the peak frequency  $f_p$  is defined as the frequency representing the maximum value of  $S(f)$ ; the peak wave direction  $d_p$  is expressed as follows:

$$d_p = \tan^{-1}(b_p/a_p) \quad (12)$$

with

$$\begin{cases} a_p = \int_0^{2\pi} \cos\theta S(f_p, \theta) d\theta \\ b_p = \int_0^{2\pi} \sin\theta S(f_p, \theta) d\theta \end{cases}; \quad (13)$$

The wavelength at the peak wavenumber  $k_p$  is calculated as follows:

$$L_p = \frac{2\pi}{k_p} \quad (14)$$

and  $k_p$  is associated with  $f_p$  according to Eq. (3), and the peak wave period  $T_p$  is defined as the reciprocal of  $f_p$  and calculated using a parabolic fit around  $f_p$ . In this work, additional geometries are also provided, such as the Goda peakedness<sup>45</sup>

$$Q_p = \frac{2}{m_0^2} \int_0^\infty f \left[ \int_0^{2\pi} S(f, \theta) d\theta \right]^2 df, \quad (15)$$

the spectral bandwidth<sup>46</sup>

$$\nu = \sqrt{\frac{m_0 m_2}{m_1^2} - 1}, \quad (16)$$

and the mean directional spreading<sup>47</sup>

$$\sigma_\theta = \left\{ 2 \left[ 1 - \left( \frac{a^2 + b^2}{m_0^2} \right)^{1/2} \right] \right\}^{1/2} \quad (17)$$

(where  $a$  and  $b$  are obtained according to Eq. (11)).

Moreover, wave parameters, which can be applied to wave energy assessment and characterization, are also presented. According to the Technical Specification proposed by the International Electrotechnical Commission<sup>48</sup> (IEC TS 62600-101:2015; hereafter IEC2015), the (omni-directional) wave power density (WPD) can be estimated as follows:

$$WPD = \rho g \int_0^{2\pi} \int_0^\infty C_g S(f, \theta) df d\theta \quad (18)$$

where  $\rho$  (taken as  $1023 \text{ kg/m}^3$  in this work) denotes the density of seawater,  $g = 9.81 \text{ m/s}^2$  is the acceleration of gravity, and  $C_g$  is the group velocity of waves in Eq. (2). IEC2015 also recommends directionally resolved WPD, i.e., resolving omni-directional WPD in a specific direction  $\theta_j$ :

$$WPD_{\theta_j} = \rho g \int_0^{2\pi} \int_0^\infty C_g S(f, \theta) \cos(\theta - \theta_j) \delta df d\theta, \quad \begin{cases} \delta = 1, \cos(\theta - \theta_j) \geq 0 \\ \delta = 0, \cos(\theta - \theta_j) < 0 \end{cases} \quad (19)$$

In this work,  $\theta_j$  are assigned to the 24 discrete directions in the spectral space, and the maximum value of  $WPD_{\theta_j}$  (denoted as  $WPD_{\theta_j, \max}$ ) and the corresponding direction  $\theta_{j, \max}$  are retained for each simulated spectrum. Furthermore, to measure the relative spread of wave energy in the  $f$  and  $\theta$  directions, IEC2015 recommends the following coefficients:

$$\epsilon_0 = \sqrt{\frac{m_0 m_{-2}}{m_{-1}^2} - 1} \quad (20)$$

and

$$d_\theta = \frac{WPD_{\theta_j, \max}}{WPD}, \quad (21)$$

respectively.

Finally, the wind-sea fraction (WSF) parameter is introduced to characterize the proportion of wind wave energy contained in each spectrum and is presented as follows:<sup>49,50</sup>

$$WSF = \frac{E_{U_p > c}}{m_0} \quad (22)$$

where  $E_{U_p > c}$  is the energy in the spectral space for which the projected wind speed  $U_p$  is larger than the local wave phase velocity  $c$ . The parameter  $U_p$  can be calculated as follows:

$$U_p = C_{mult} U_{10} \cos(\delta), \quad (23)$$

where  $U_{10}$  denotes the wind speed at the height of 10 m above the sea surface,  $\delta$  denotes the angle between the direction in the spectral space and the direction in which the wind is blowing, and  $C_{mult}$  is a coefficient set as 1.7 in this work. In the spectral space, wave phase velocity  $c$  is associated with frequency  $f$  and wavenumber  $k$ ; thus, according to Eq. (3),

$$c = \frac{2\pi f}{k} = \sqrt{\frac{g}{k}} \cdot \sqrt{\tanh(kd)}. \quad (24)$$

**Identification of wind waves and swells.** The spectral partitioning technique was introduced to demonstrate the historical and future wave characteristics in the forms of wind waves and swells. The spectral partitioning technique can be traced back to a digital image processing watershed algorithm<sup>51</sup>, which can be adopted to identify watershed lines, mountain peaks, and valleys in topographic maps. Because the 2D spectrum resembles a topological surface, it is logical to apply such an algorithm in this circumstance<sup>52</sup>. As described by Hanson and Phillips<sup>49</sup>, the basic approach to the spectral partitioning method is that by searching through the spectral matrix  $S(f, \theta)$ , the paths of steepest ascent leading to each peak or local energy maximum can be identified; then, all paths



No.	<para_id>	Description	Dimensions	Units
1	Hs	Significant wave height	[lon, lat, npt, time]	m
2	Tp	Peak wave period	[lon, lat, npt, time]	s
3	Te	Wave energy period	[lon, lat, npt, time]	s
4	Tm01	Mean wave period	[lon, lat, npt, time]	s
5	Tm02	Mean zero-crossing period	[lon, lat, npt, time]	s
6	fp	Peak frequency	[lon, lat, npt, time]	Hz
7	Lp	Peak wavelength	[lon, lat, npt, time]	m
8	Dirm	Mean wave direction (Cartesian To)	[lon, lat, npt, time]	degr.
9	Dirp	Peak wave direction (Cartesian To)	[lon, lat, npt, time]	degr.
10	Spr	Mean directional spreading <sup>47</sup>	[lon, lat, npt, time]	degr.
11	nu	Spectral bandwidth <sup>46</sup>	[lon, lat, npt, time]	—
12	Qp	Goda peakedness <sup>45</sup>	[lon, lat, npt, time]	—
13	WSF	Wind sea fraction	[lon, lat, npt, time]	%
14	WPD	Wave power density <sup>48</sup>	[lon, lat, npt, time]	kW/m
15	WPeps0	Relative spread of wave energy in frequency dimension <sup>48</sup>	[lon, lat, npt, time]	—
16	WPthmx	Maximum value of directionally resolved WPD <sup>48</sup>	[lon, lat, npt, time]	kW/m
17	thWPmx	Direction of WPthmx (Cartesian To) <sup>48</sup>	[lon, lat, npt, time]	degr.
19	windx	Wind speed toward east	[lon, lat, time]	m/s
20	windy	Wind speed toward north	[lon, lat, time]	m/s

**Table 1.** List of all variables in the dataset.

leading to the same peak can be grouped, and the members that lie on the collection of the paths are considered to belong to a distinct partition. Partitioning of wave spectra is widely adopted in research concerning data assimilation<sup>52–54</sup>, spatial and temporal tracking of wave systems<sup>49,55</sup>, and so on.

Notably, the wave parameters mentioned above can be calculated not only from the entire spectrum but also from a partition of it. In each partitioned spectrum, partitions whose WSF are greater than 33.33%, together with spectral elements ( $f$ ,  $\theta$ ) whose phase velocities (Eq. (24)) are less than the local projected wind speeds (Eq. (23)), are combined as a new partition, and the newly formed partition is considered to be under the direct influence of the wind and thus is identified as the wind wave system; the remaining partitions ( $WSF < 33.33\%$ ), including those incomplete partitions that have contributed some elements to the wind wave system, are then identified as individual swells. The swells in the same spectrum can be combined as a total swell partition.

The partitioning and identification program implemented in this work was developed based on the W3PARTMD module of WaveWatch III ver. 6.07<sup>56</sup>, in which an efficient FORTRAN routine was transformed from the MATLAB code<sup>57,58</sup> that was used to apply the watershed algorithm<sup>51</sup>. The wave parameters provided in this dataset are calculated from both the entire spectra and its wind wave and swell partitions; see the Data Records section for more details.

## Data Records

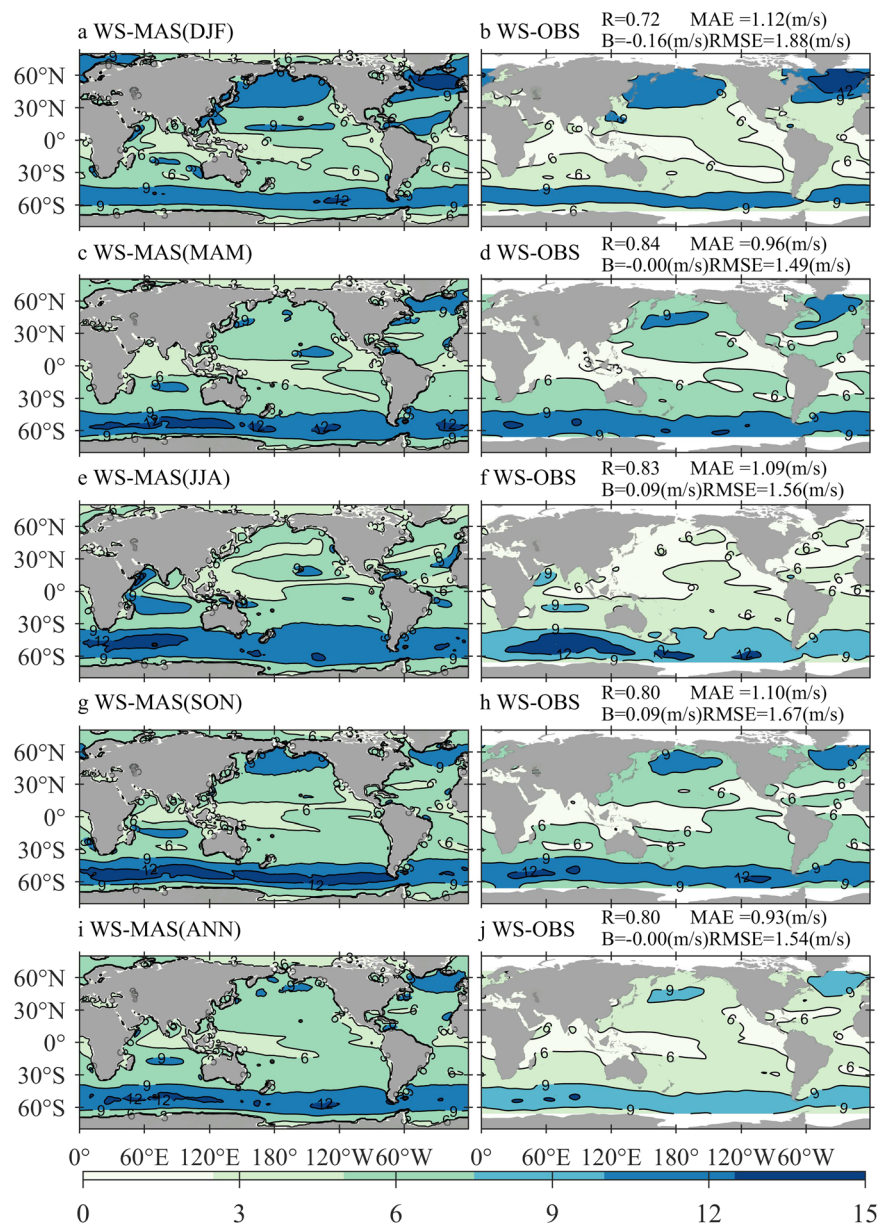
This dataset consists of up to 17 kinds of wave parameters, and wind speeds at 10 meters above the sea surface (toward the east and north) are also provided. The dataset covers the area of  $0^{\circ}E$ – $359^{\circ}E$ ,  $80^{\circ}S$ – $80^{\circ}N$  with spatial intervals of  $1^{\circ}$ , and the temporal intervals are 3 hours. The 17 wave parameters presented in this dataset are integrated from both the entire spectra (i.e., combined wind waves and swells, hereafter COMB) and their partitions, and the partitions are presented in the forms of wind waves (WSEA), total swells (TSWL), and the first three swells with the largest  $H_s$  (SWL1, SWL2, and SWL3). Moreover, the wave dataset spans a 65-year historical period (1950–2014) and three 86-year future scenarios (2015–2100).

The data mentioned above are stored monthly for each wave and wind parameter, and the filenames of the data are in the following format:

<para\_id>\_<exp\_id>\_<yyyymm>.nc,

where <para\_id> denotes the name of parameters, see Table 1; <exp\_id> represents the name of the CMIP6 driving conditions, which are ‘histor’, ‘ssp126’, ‘ssp245’, and ‘ssp585’; and <yyyymm> are expressed as 195001–201412 for the ‘histor’ data and as 201501–210012 for the three future scenarios. Since there are 780 and 1032 months during 1950–2014 and 2015–2100, respectively, and there are 17 wave parameters and 2 wind parameters to be exhibited, the number of files in the historical catalog is  $780 \times 19 = 14820$ , and the number of files for each future scenario is  $1032 \times 19 = 19608$ .

All data files are provided in NetCDF format and are archived in the ScienceDB<sup>59</sup>. The variables in each file, together with their descriptions, dimensions, and units, are outlined in Table 1. The dimension of time is presented as days since 1950-01-01 00:00:00; the dimensions of latitude and longitude are expressed as degrees north and east, respectively; and variables associated with waves contain the ‘npt’ dimension, and npt from 1 to 6 indicate variables derived from the COMB, WSEA, TSWL, and SWL1-3, respectively. In this dataset, ‘\_FillValue’ denotes land points; in particular, wave-associated variables may present a negative value, indicating that the spectrum or the spectral partition from which the parameter is obtained contains much less energy, such that the corresponding  $H_s$  is less than 0.05 m.

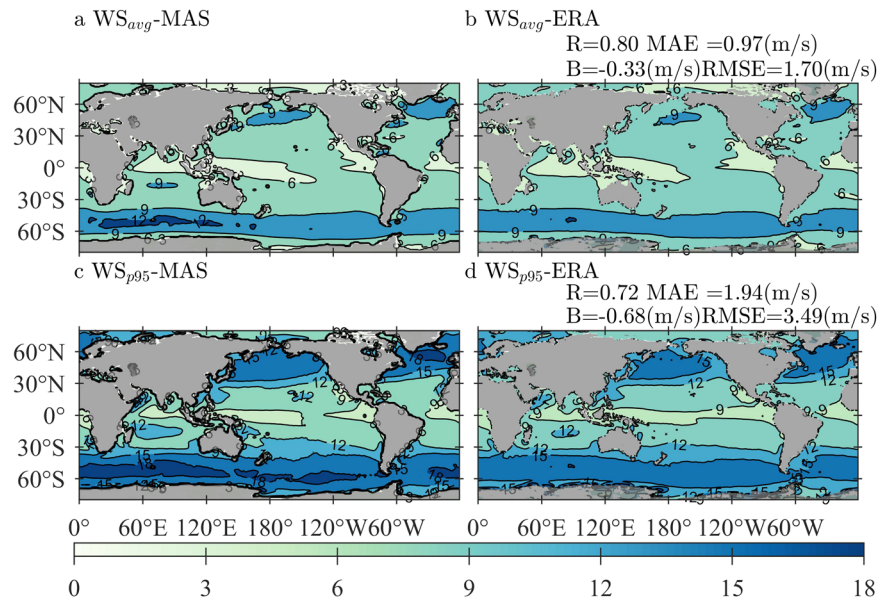


**Fig. 1** Climatological distributions of the averaged WS-MAS (wind speed adopted to force MASNUM-WAM, left column) and WS-OBS (wind speed from observation, right column) in boreal winter (**a,b**), boreal spring (**c,d**), boreal summer (**e,f**), boreal autumn (**g,h**), and 2013–2014 (**i,j**), with the corresponding quantitative errors exhibited in panels b, d, f, h, and j, respectively.

### Technical Validation

The MASNUM-WAM has been calibrated and adopted many times in previous scientific and engineering studies (e.g.<sup>60–65</sup>); moreover, MASNUM-WAM is now the ocean wave component of several operational ocean forecasting systems (OFS), such as the OFS for the seas of China and adjacent areas<sup>66</sup>, OFS for Southeast Asian Seas and OFS for the 21st-Century Maritime Silk Road<sup>67</sup>. Therefore, validation of the MASNUM-WAM is not shown in this study.

The validation of FIO-ESM v2.0 can be referred to in the work of Bao *et al.*<sup>22</sup>, in which FIO-ESM v2.0 was applied to conduct the CMIP6 DECK (Diagnostic, Evaluation and Characterization of Klima) and historical (1850–2014) experiments<sup>23</sup>. The results show that the time evolutions of surface air temperature, sea surface temperature, and Atlantic meridional overturning circulation in the past centuries are well reproduced; in particular, the common, large, warm sea surface temperature bias for all climate models is dramatically reduced, and the simulated El Niño–Southern Oscillation period is much closer to the observation within 2–7 years. Therefore, it is suggested that the performance of FIO-ESM v2.0 under the CMIP6 experimental framework is stable and reliable, including in both the historical and future scenarios. Moreover, Song *et al.*<sup>25</sup> performed the



**Fig. 2** Climatological distributions of (a)  $WS_{avg}$ -MAS (averaged wind speed from the forcings of MASNUM-WAM), (b)  $WS_{avg}$ -ERA (averaged wind speed from ERA5), (c)  $WS_{p95}$ -MAS (95th-percentile wind speed from the forcings of MASNUM-WAM) and (d)  $WS_{p95}$ -ERA (95th-percentile wind speed from ERA5), with the corresponding quantitative errors exhibited in panels b and d, respectively. The sampling period is from 2010 to 2014.

validation of the FIO-ESM v2.0 wave product in the CMIP6 historical experiment. In the comparison against the ERA5 reanalysis data from 1979–2014, the monthly mean  $H_s$ ,  $\theta_m$ ,  $T_p$ , and  $T_{m02}$  show good agreement in terms of the basic characteristics of spatial pattern and seasonal variation, as well as the 99th-percentile values of  $H_s$  derived from the 3-hourly data.

As the aim of this dataset is to aid in wave climate and wave energy resource research, and as most of the research characteristics of interest are closely related to the parameters of  $H_s$  and  $T_p$ , the validation focuses on the climatology of the two key parameters, denoted as  $H_s$ -MAS and  $T_p$ -MAS, in this dataset. In addition, the quality of wind forcing adopted to force MASNUM-WAM is also to be validated, and we focus on the wind speed (WS), denoted as WS-MAS, in this section.

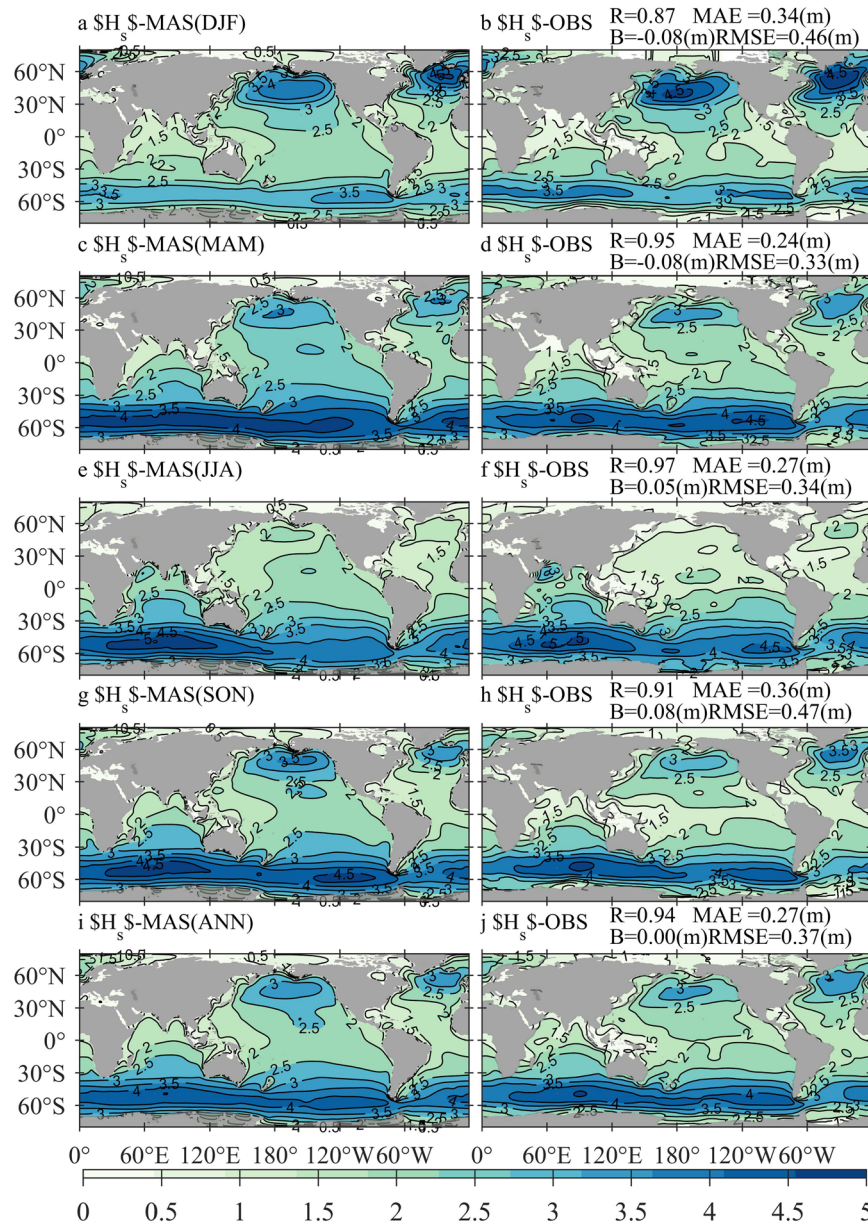
The AVISO gridded wind and wave products<sup>68</sup> are selected as the observation baseline to be compared with WS-MAS and  $H_s$ -MAS. The AVISO products comprise the daily WS and  $H_s$  observations, denoted as WS-OBS and  $H_s$ -OBS, respectively, merged from a set of missions, such as Envisat, Jason-1–3, AltiKa, and Sentinel-3A, and it covers  $0^\circ E$  to  $359^\circ E$ ,  $90^\circ S$  to  $89^\circ N$  with a horizontal resolution of  $1^\circ \times 1^\circ$ . Notably, to be comparable with the MASNUM-WAM simulated results in the historical scenario, the sampling periods are 2013–2014 and 2010–2014 for WS-OBS and  $H_s$ -OBS, respectively. To validate WS-MAS with a longer sampling period and to validate  $H_s$ -MAS and  $T_p$ -MAS in the forms of wind waves and swells separately, the ‘ERA5 hourly data on single levels from 1959 to present’ dataset<sup>69</sup> is adopted. The ERA5 hourly data can provide 10m u-v winds and separate COMB, WSEA, and TSWL wave characteristics, covering  $0^\circ E$ – $360^\circ E$ ,  $90^\circ S$ – $90^\circ N$  with spatial intervals of  $0.25^\circ$  and  $0.5^\circ$  for the wind and wave parameters, respectively; to be comparable with  $H_s$ -MAS and  $T_p$ -MAS, the original spatial and temporal resolutions are reduced and the sampling period is selected as 2010–2014. Then, the employed ERA5 WS,  $H_s$  and  $T_p$  products are denoted as WS-ERA,  $H_s$ -ERA and  $T_p$ -ERA, respectively. Finally, to perform a more general validation, a global ensemble of ocean wave climate statistics<sup>28</sup> is introduced. As a product of the COWCLIP2<sup>10,28</sup>, the statistics mentioned above comprise 14 contemporary wave reanalysis and hindcasts computed across 1980–2014, including general and extreme statistics of  $H_s$ , mean wave period (such as  $T_{m01}$  and  $T_{m02}$ ), and  $\theta_m$  at different frequency resolutions (monthly, seasonally, and annually). We employed the mean values of annually averaged  $H_s$ , denoted as  $H_s$ -COW, in the statistics as the baseline to validate the historical  $H_s$ -MAS. Similarly, the spatial resolution of  $H_s$ -COW has been adjusted to that of  $H_s$ -MAS. Furthermore, in the comparisons below, four quantitative errors are also exhibited: the Pearson’s correlation coefficient  $R$

$$R = \frac{\sum_i ((S_i - \bar{S})(O_i - \bar{O}))}{\sqrt{\sum_i (S_i - \bar{S})^2 \cdot \sum_i (O_i - \bar{O})^2}}, \quad (25)$$

the mean absolute error MAE

$$MAE = \frac{1}{N} \sum_{i=1}^N |S_i - O_i|, \quad (26)$$





**Fig. 3** Climatological distributions of the averaged Hs-MAS (significant wave height simulated by MASNUM-WAM, left column) and Hs-OBS (significant wave height from observation, right column) in boreal winter (a,b), boreal spring (c,d), boreal summer (e,f), boreal autumn (g,h), and 2010–2014 (i,j), with the corresponding quantitative errors exhibited in panels b, d, f, h, and j, respectively.

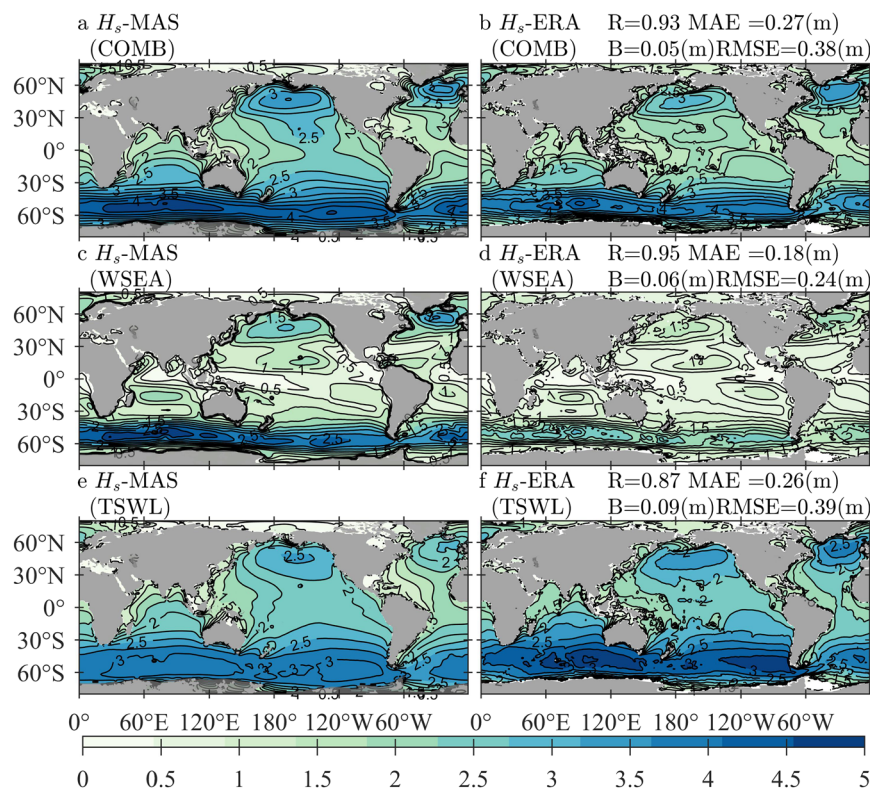
the mean bias  $B$

$$B = \frac{1}{N} \sum_{i=1}^N (S_i - O_i), \quad (27)$$

and the root-mean-square error RMSE

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (S_i - O_i)^2}. \quad (28)$$

In Eqs. (25–28),  $S$  denotes the key parameters from the described dataset, i.e., WS-MAS, Hs-MAS or Te-MAS, and  $O$  indicates the corresponding parameters from AVISO observations, ERA5 reanalysis, or COWCLIP2 products. Notably, the abovementioned  $S$  and  $O$  represent the annual-averaged statistics or the 95-th percentiles over the time in each dataset to be compared. Thus, the quantitative comparisons with Eqs. (25–28)



**Fig. 4** Climatological distributions of the averaged  $H_s$ -MAS (significant wave height simulated by MASNUM-WAM, left column) and  $H_s$ -ERA (significant wave height from ERA5, right column) for COMB (combined wind waves and swell, **a,b**), WSEA (wind waves, **c,d**), and TSWL (total swells, **e,f**) wave patterns, with the corresponding quantitative errors exhibited in panels b, d, and f, respectively. The average period is from 2010 to 2014.

are performed on space, i.e., the subscript  $i$  indicates the index of each water grid point in  $S$  and  $O$ , considering the total number of sampling grid points  $N$  and  $\bar{S} = \frac{1}{N} \sum_{i=1}^N S_i$  and  $\bar{O} = \frac{1}{N} \sum_{i=1}^N O_i$ . The spatial resolutions of  $S$  and  $O$  are unified as mentioned previously.

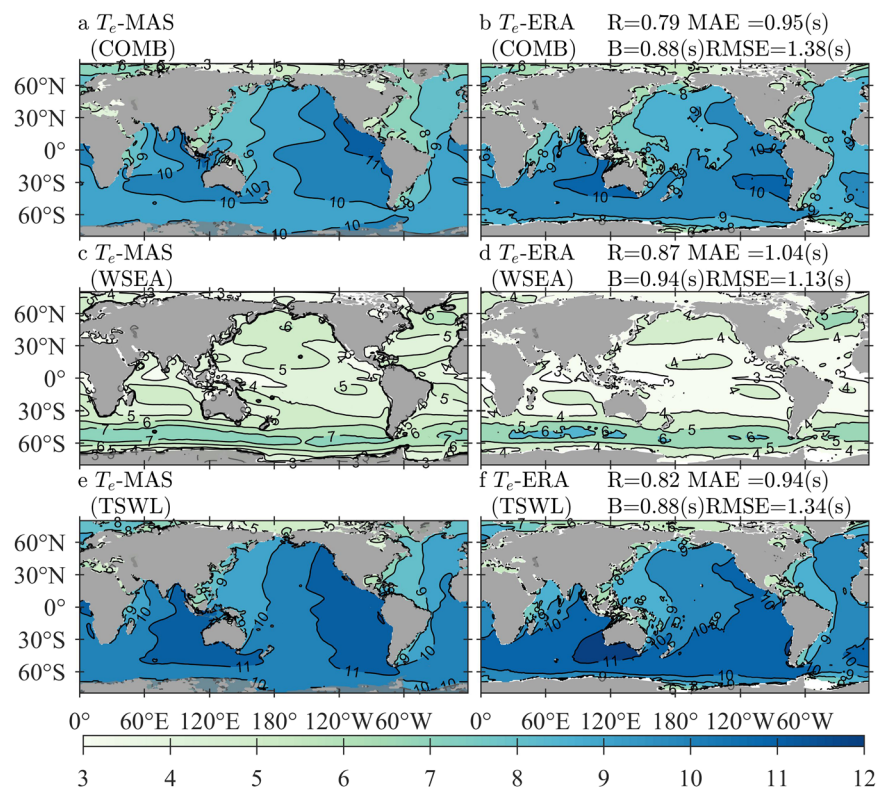
**Comparisons of WS-MAS against WS-OBS and WS-ERA.** Comparisons of WS-MAS against WS-OBS and WS-ERA can validate the quality of the forcing wind. Figure 1 illustrates the climatological distributions of WS-MAS and WS-OBS in boreal winter (December-January-February, DJF, panels a,b), spring (March-April-May, MAM, panels c,d), summer (June-July-August, JJA, panels e-f), and autumn (September-October-November, SON, panels g-h) during 2013–2014, as well as the annual (ANN, panels i-j) mean result. The quantitative errors between the seasonal- or annual-averaged WS-MAS and WS-OBS are also exhibited in the same rows of the corresponding panels. Figure 1 shows that WS-WAM and WS-OBS can achieve a strong level of agreement around the world. The distribution patterns of the two characteristics are quite similar for both seasonal and annual mean statistics. The mean value of quantitative error  $B$  can even be 0.00 m/s when all the samples are involved.

The climatological distributions of averaged WS-MAS and WS-ERA (panels a–b), together with the 95-th percentile of the two characters (panels c–d), are illustrated in Fig. 2, with the corresponding quantitative errors exhibited in panels b and d, respectively. The sampling period is from 2010 to 2014. As shown in Fig. 2, the distribution patterns of WS-MAS and WS-ERA can match strongly in both mean and extreme conditions, although understandably, the quantitative errors in extreme conditions are slightly higher than those in mean conditions.

Therefore, we can conclude that the quality of the forcing wind derived from the FIO-ESM v2.0 product is robust and reliable.

**Comparison between  $H_s$ -MAS and  $H_s$ -OBS.** A comparison between  $H_s$ -MAS and  $H_s$ -OBS can be used to assess the seasonal and annual mean state of the simulated  $H_s$  in spatial distribution patterns. The climatological distributions of averaged  $H_s$ -MAS and  $H_s$ -OBS in boreal winter (December-January-February, DJF, panels a,b), spring (panels c,d), summer (panels e,f), and autumn (panels g,h) during 2010–2014, together with the annual (panels i,j) mean result, are illustrated in Fig. 3. The quantitative errors between the seasonal- or annual-averaged  $H_s$ -MAS and  $H_s$ -OBS are also exhibited in the same rows of the corresponding panels.

Figure 3 shows that the comparison against the satellite-observed  $H_s$  has a good agreement. The seasonal- and annual-averaged  $H_s$ -MAS and  $H_s$ -OBS distribution patterns are very similar. The differences are generally



**Fig. 5** Climatological distributions of the 95-th percentile Hs-MAS (significant wave height simulated by MASNUM-WAM, left column) and Hs-ERA (significant wave height from ERA5, right column) for COMB (combined wind waves and swell, a,b), WSEA (wind waves, c,d), and TSWL (total swells, e,f) wave patterns, with the corresponding quantitative errors exhibited in panels b, d, and f, respectively. The sampling period is from 2010 to 2014.

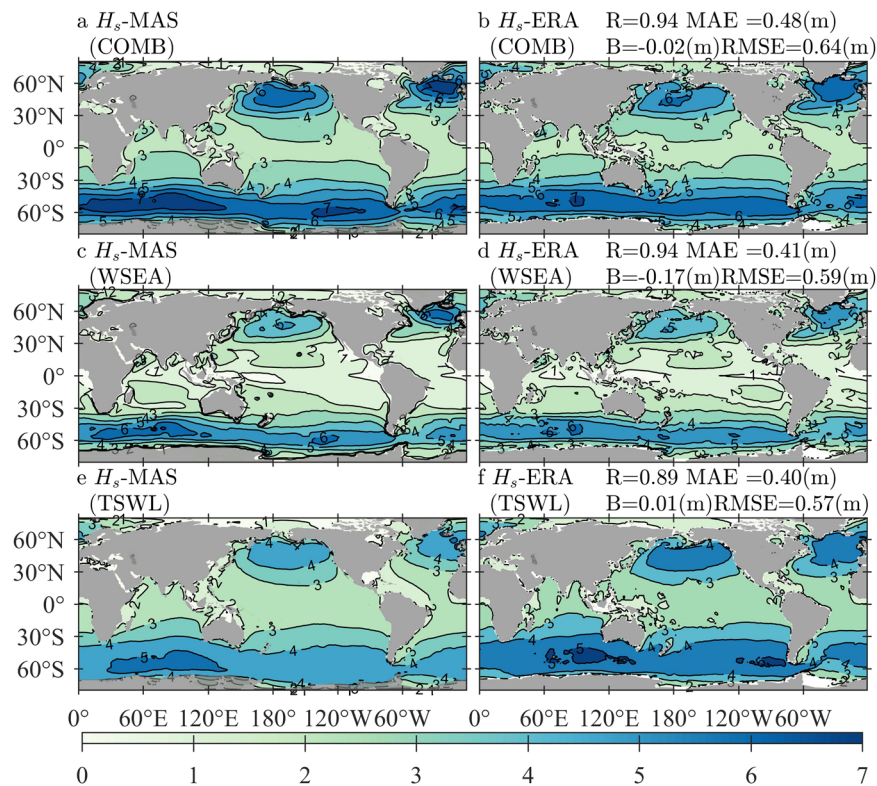
restricted within  $\pm 0.4$  m in most oceans around the world for the three-monthly averages and are even reduced to  $\pm 0.2$  m when the yearly mean values are considered. From the quantitative errors, the MAE and RMSE are generally less than 0.3 m and 0.4 m, respectively, the values of the R coefficients are generally greater than 0.9, and all the B values are very close to 0 m.

Although the MASNUM-WAM simulations with the FIO-ESM v2.0 wind forcings can capture the basic characteristics of ocean waves, large differences from observations in some sea areas are still inevitable. For example, Hs-MAS is larger than Hs-OBS by approximately 0.4–0.8 m in the seas south of 60°S throughout the years averaged, and the extreme value can be over 1.2 m in DJF and SON. In contrast, Hs-MAS is smaller than Hs-OBS in the Arctic Ocean, and extreme differences of  $-0.8$  m can be found in DJF. Moreover, Hs-MAS is smaller than Hs-OBS in all four seasons over the North Atlantic, where the absolute differences are smaller in JJA and larger in DJF, with values of approximately 0.2–0.4 m and 0.8–1.0 m, respectively.

**Comparisons of Hs-MAS and Te-MAS against Hs-ERA and Te-ERA.** Comparisons of Hs-MAS and Te-MAS against Hs-ERA and Te-ERA demonstrate the performance of the simulated  $H_s$  and  $T_e$  in the forms of wind waves and swells separately. Figure 4 illustrates the climatological distributions of annually averaged Hs-MAS and Hs-ERA in the forms of COMB (panels a,b), WSEA (panels c,d), and TSWL (panels e,f), with the quantitative errors exhibited in the same rows of the corresponding panels. In addition to the mean state, extreme conditions, i.e., the 95th-percentile values of Hs-MAS and Hs-OBS derived from 2010–2014, are presented in Fig. 5, where the panels and quantitative errors are arranged similarly to those in Fig. 4.

Figure 4 shows that the mean states of COMB Hs-MAS coincide well with those of COMB Hs-ERA, where the four errors are close to those exhibited in the comparison between the annual mean Hs-MAS and Hs-OBS. The consistency of the two datasets in WSEA is even better, but a larger difference can be found when considering TSWL conditions; nevertheless, quantitative errors between TSWL Hs-MAS and TSWL Hs-ERA still suggest acceptable goodness of fit. For extreme waves shown in Fig. 5, the comparisons also exhibit good agreements; the R coefficients continue at the high levels that have been found in Fig. 4, and the values of B, MAE and RMSE in Fig. 5 become larger due to higher wave heights involved in the statistical procedure. Differences in the spatial distribution of both Figs. 4, 5 are similar to Fig. 3. Hs-MAS is larger than Hs-ERA in the seas south of 60°S for both WSEA and TSWL conditions, and TSWL Hs-MAS is even larger than TSWL Hs-ERA by





**Fig. 6** Climatological distributions of the averaged Te-MAS (wave energy period simulated by MASNUM-WAM, left column) and Te-ERA (wave energy period from ERA5, right column) for COMB (combined wind waves and swell, **a,b**), WSEA (wind waves, **c,d**), and TSWL (total swells, **e,f**) wave patterns, with the corresponding quantitative errors exhibited in panels b, d, and f, respectively. The average period is from 2010 to 2014.

approximately 0.6–0.8 m in both mean and extreme conditions. In the Arctic Ocean,  $H_s$ -MAS is smaller than  $H_s$ -ERA, mainly due to the smaller estimated TSWL of  $H_s$ -MAS. In the North Atlantic,  $H_s$ -MAS is generally smaller than  $H_s$ -ERA for the mean state, but for extreme conditions, the former is larger than the latter in the eastern part of the ocean.

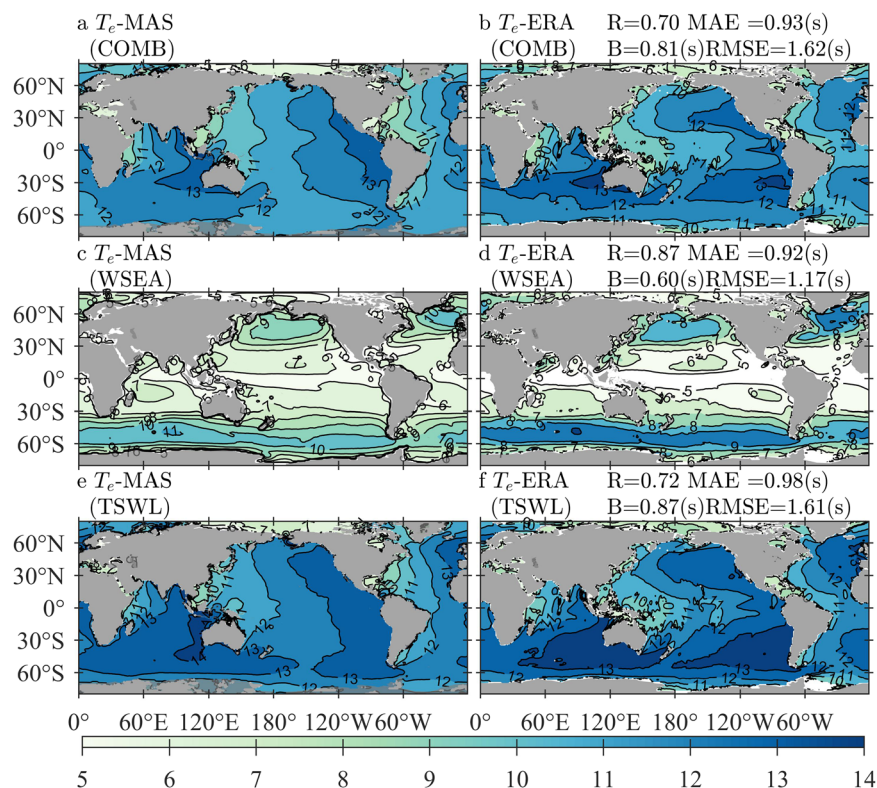
The comparisons between Te-MAS and Te-ERA for the mean and extreme conditions are shown in Figs. 6, 7, respectively. Panels for the wave patterns of COMB, WSEA, and TSWL are illustrated in the same way as Figs. 4, 5, as well as the quantitative errors. Parameters associated with wave periods may be influenced markedly by spectral shapes; thus, the simulated  $T_e$  in the two datasets might not easily achieve consistency, especially when spectral partitioning is conducted. Figure 6 shows that Te-MAS is larger than Te-ERA in almost all oceans around the world, including both WSEA and TSWL patterns, but the extreme bias is no more than 1.8 s. TSWL Te-MAS is smaller than TSWL Te-ERA by approximately 0.2–0.4 s in the North Atlantic, resulting in a smaller COMB Te-MAS in the same location. For the extreme conditions shown in Fig. 7, the differences between Te-MAS and Te-ERA are reduced for both WSEA and TSWL conditions, although the Te-MAS is still larger than the Te-ERA by approximately 1 s; it is noted that Te-MAS is estimated to be smaller than Te-ERA in the North Atlantic for both WSEA and TSWL wave patterns.

**Comparisons between  $H_s$ -MAS and  $H_s$ -COW.** Figure 8 illustrates the climatological distributions of the averaged  $H_s$ -MAS (left-column) and  $H_s$ -COW (right-column) for the annual mean (panels a–b) and 95-th percentile (panels c–d) statistics. The corresponding quantitative errors are exhibited in panels b and d, respectively.

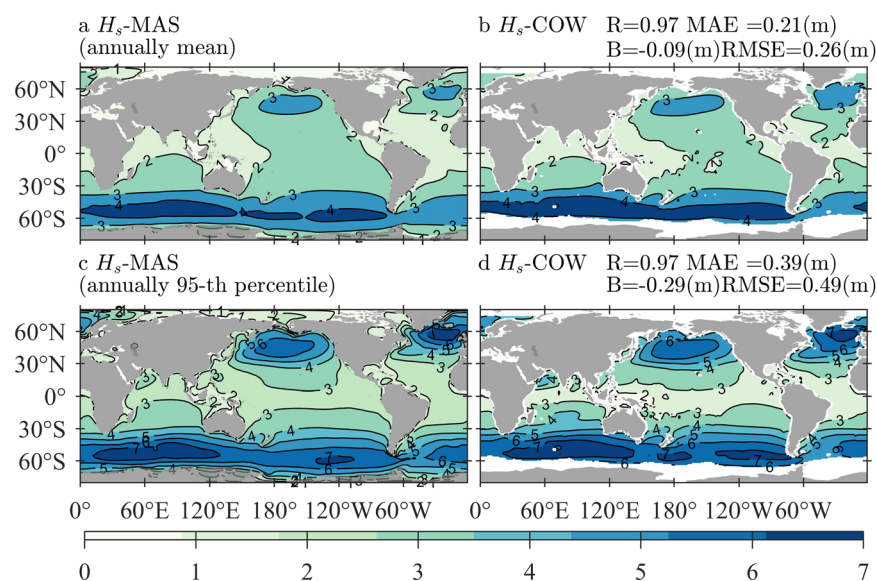
Good agreement between the statistics of  $H_s$ -MAS and  $H_s$ -COW can still be found in Fig. 8. The deviations in spatial distribution for both mean and extreme conditions are very similar to the corresponding panels in Figs. 4, 5, as well as the quantitative errors. In addition,  $H_s$ -MAS is also smaller than  $H_s$ -COW in the North Atlantic by approximately 0.4 m for the annual mean values and by approximately 1.0 m when the 95-th percentiles are considered.

Overall, the above analyses indicate that the mean state of WS,  $H_s$ , and  $T_e$  proposed in the newly established dataset can capture the basic characteristics of the satellite observations in seasonal and annual spatial distributions and can also be broadly consistent with the ERA5 products in both forms of WSEA and TSWL and both mean and extreme wave conditions. The comparisons against the more general wave statistics produced through COWCLIP2 confirm the conclusions mentioned above. However, the simulated  $H_s$  and  $T_e$  may still suffer biases, especially in the southern 60°S, Arctic, and North Atlantic Oceans.





**Fig. 7** Climatological distributions of the 95th-percentile  $T_e$ -MAS (wave energy period simulated by MANUM-WAM, left column) and  $T_e$ -ERA (wave energy period from ERA5, right column) for COMB (combined wind waves and swell, **a,b**), WSEA (wind waves, **c,d**), and TSWL (total swells, **e,f**) wave patterns, with the corresponding quantitative errors exhibited in panels b, d, and f, respectively. The sampling period is from 2010 to 2014.



**Fig. 8** Climatological distributions of the averaged  $H_s$ -MAS (significant wave height simulated by MASNUM-WAM, left-column) and  $H_s$ -COW (significant wave height derived from COWCLIP2 dataset, right-column) for the annual mean (panels **a,b**) and 95-th percentile (panels **c,d**) statistics, with the corresponding quantitative errors exhibited in panels b and d, respectively. The sampling period is from 1980 to 2014.

### Code availability

The source code of MASNUM-WAM is available to the public and can be downloaded from <https://doi.org/10.57760/sciencedb.0289370>. The dataset can be regenerated by using the wind parameter files, i.e., files with <para\_id> of 'windx' and 'windy', archived in ScienceDB<sup>59</sup> as wind forcings.

Received: 29 September 2022; Accepted: 13 April 2023;

Published online: 20 April 2023

## References

- In, K. *et al.* Analysis of a Marine Accident and Freak Wave Prediction with an Operational Wave Model. *The 19th International Offshore and Polar Engineering conference*. <http://legacy.iso.org/publications/proceedings/ISOPE/ISOPE%202009/data/papers/2009-TPC-426.pdf> (2009).
- Waseda, T. *et al.* Predicting freakish sea state with an operational third-generation wave model. *Nat. Hazards Earth Syst. Sci.* **14**, 945–957 (2014).
- Trulsen, K., Nieto Borge, J. C., Gramstad, O., Aouf, L. & Lefèvre, J.-M. Crossing sea state and rogue wave probability during the Prestige accident. *J. Geophys. Res. Oceans* **120**, 7113–7136 (2015).
- Waseda, T., Tamura, H. & Kinoshita, T. Freakish sea index and sea states during ship accidents. *J. Mar. Sci. Technol.* **17**, 305–314 (2012).
- Tamura, H., Waseda, T. & Miyazawa, Y. Freakish sea state and swell-windsea coupling: Numerical study of the Suwa-Marui incident. *Geophys. Res. Lett.* **36**, 2–6 (2009).
- Cavaleri, L. *et al.* The Draupner wave: A fresh look and the emerging view. *J. Geophys. Res. Oceans* **121**, 6061–6075 (2016).
- Guedes Soares, C., Cherneva, Z. & Antao, E. M. Characteristics of abnormal waves in North Sea storm sea states. *Appl. Ocean Res.* **25**, 337–344 (2003).
- Karin Magnusson, A. & Donelan, M. A. The Andrea Wave Characteristics of a Measured North Sea Rogue Wave. *J. Offshore Mech. Arct. Eng.* **135**, 031108 (2013).
- Young, I. R. & Ribal, A. Multiplatform evaluation of global trends in wind speed and wave height. *Science* **364**, 548–552 (2019).
- Hemer, M. A., Wang, X. L., Weisse, R. & Swail, V. R. Advancing Wind-Waves Climate Science: The COWCLIP Project. *Bull. Am. Meteorol. Soc.* **93**, 791–796 (2012).
- Pelc, R. & Fujita, R. M. Renewable energy from the ocean. *Marine Policy* **26**, 471–479 (2002).
- Vosough, A. Wave Energy. *Int. J. Multidiscip. Sci. Eng.* **2**, 60–63 (2011).
- Dee, D. P. *et al.* The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Q. J. R. Meteorol. Soc.* **137**, 553–597 (2011).
- Hersbach, H. *et al.* Global reanalysis: goodbye ERA-Interim, hello ERA5. *ECMWF Newsletter* **159**, 17–24, <https://www.ecmwf.int/node/19027> (2019).
- Bell, B. *et al.* The ERA5 global reanalysis: Preliminary extension to 1950. *Q. J. R. Meteorol. Soc.* **147**, 4186–4227 (2021).
- Hersbach, H. *et al.* The ERA5 global reanalysis. *Q. J. R. Meteorol. Soc.* **146**, 1999–2049 (2020).
- Chawla, A., Spindler, D. M. & Tolman, H. L. Validation of a thirty year wave hindcast using the Climate Forecast System Reanalysis winds. *Ocean Model.* **70**, 189–206 (2013).
- Janssen, P. & Bidlot, J. J.-R. On the extension of the freak wave warning system and its verification. *ECMWF Technical Memorandum* 588 <https://www.ecmwf.int/node/10243> (2009).
- Janssen, P. A. E. M. On some consequences of the canonical transformation in the Hamiltonian theory of water waves. *J. Fluid Mech.* **637**, 1–44 (2009).
- Janssen, P. A. E. M. On a random time series analysis valid for arbitrary spectral shape. *J. Fluid Mech.* **759**, 236–256 (2014).
- Janssen, P. A. E. M. Shallow-water version of the Freak Wave Warning System. *ECMWF Technical Memorandum* 813 <https://www.ecmwf.int/node/18063> (2017).
- Bao, Y., Song, Z. & Qiao, F. FIO-ESM Version 2.0: Model Description and Evaluation. *J. Geophys. Res. Oceans* **125** (2020).
- Eyring, V. *et al.* Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geosci. Model Dev.* **9**, 1937–1958 (2016).
- O'Neill, B. C. *et al.* The Scenario Model Intercomparison Project (ScenarioMIP) for CMIP6. *Geosci. Model Dev.* **9**, 3461–3482 (2016).
- Song, Z. *et al.* Centuries of monthly and 3-hourly global ocean wave data for past, present, and future climate research. *Sci. Data* **7**, 226 (2020).
- Taylor, K. E., Stouffer, R. J. & Meehl, G. A. An Overview of CMIP5 and the Experiment Design. *Bull. Am. Meteorol. Soc.* **93**, 485–498 (2012).
- Morim, J. *et al.* A global ensemble of ocean wave climate projections from CMIP5-driven models. *Sci. Data* **7**, 105 (2020).
- Morim, J. *et al.* A global ensemble of ocean wave climate statistics from contemporary wave reanalysis and hindcasts. *Sci. Data* **9**, 358 (2022).
- Yuan, Y., Hua, F., Pan, Z. & Sun, L. LAGDF-WAM numerical wave model - I. basic physical model. *Acta Oceanol. Sin.* **10**, 483–488 (1991).
- Yuan, Y., Hua, F., Pan, Z. & Sun, L. LAGDF-WAM numerical wave model - II: characteristics inlaid scheme and its application. *Acta Oceanol. Sin.* **11**, 13–23 (1992).
- Yang, Y., Qiao, F., Zhao, W., Teng, Y. & Yuan, Y. MASNUM ocean wave numerical model in spherical coordinates and its application. *Acta Oceanol. Sin.* **27**, 1–7 (2005).
- Qiao, F. *et al.* A Highly Effective Global Surface Wave Numerical Simulation with Ultra-High Resolution. *The International Conference for High Performance Computing, Networking, Storage and Analysis (SC'16)* **5**, 1–11, <https://doi.org/10.1109/SC.2016.4> (2016).
- Erick Rogers, W., Babanin, A. V. & Wang, D. W. Observation-consistent input and whitecapping dissipation in a model for wind-generated surface waves: Description and simple calculations. *J. Atmospheric Ocean. Technol.* **29**, 1329–1346 (2012).
- Zieger, S., Babanin, A. V., Erick Rogers, W. & Young, I. R. Observation-based source terms in the third-generation wave model WAVEWATCH. *Ocean Model.* **96**, 2–25 (2015).
- Liu, Q. *et al.* Numerical simulations of ocean surface waves under hurricane conditions: Assessment of existing model performance. *Ocean Model.* **118**, 73–93 (2017).
- 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. *J. Phys. Oceanogr.* **15**, 1369–1377 (1985).
- Hasselmann, S. & Hasselmann, K. Computations and Parameterizations of the Nonlinear Energy Transfer in a Gravity-Wave Spectrum. Part II: Parameterizations of the Nonlinear Energy Transfer for Application in Wave Models. *J. Phys. Oceanogr.* **15**, 1378–1391 (1985).
- NGDC NOAA, *ETOPO1 1 Arc-Minute Global Relief Model*. <https://www.ncei.noaa.gov/access/metadata/landing-page/bin/iso?id=gov.noaa.ngdc.mgg.dem:316> (2020).
- Neale, R. B. *et al.* Description of the NCAR Community Atmosphere Model (CAM 5.0) <https://opensky.ucar.edu/islandora/object/technotes%3A594/> (2012).
- Lawrence, D. M. *et al.* Parameterization improvements and functional and structural advances in Version 4 of the Community Land Model. *J. Adv. Model. Earth Syst.* **3**, 27 (2011).
- Smith, R. & Jones, P. *The parallel ocean program (POP) reference Manual*. <https://www.cesm.ucar.edu/models/cesm1.0/pop2/doc/sci/POPRefManual.pdf> (2010).

42. Hunke, E. & Lipscomb, W. *CICE: The Los Alamos Sea Ice Model. Documentation and Software User's Manual Version 4.1* [https://cdms.colorado.edu/w/images/CICE\\_documentation\\_and\\_software\\_user's\\_manual.pdf](https://cdms.colorado.edu/w/images/CICE_documentation_and_software_user's_manual.pdf) (2008).
43. Qiao, F. *et al.* Wave-induced mixing in the upper ocean: Distribution and application to a global ocean circulation model. *Geophys. Res. Lett.* **31** (2004).
44. Qiao, F. *et al.* A three-dimensional surface wave–ocean circulation coupled model and its initial testing. *Ocean Dyn.* **60**, 1339–1355 (2010).
45. Goda, Y. Numerical experiments on wave statistics with spectral simulation. *Rep. Port Harb. Res. Inst.* **9**, 3–57 (1970).
46. Longuet-Higgins, M. S. On the Joint Distribution of Wave Periods and Amplitudes in a Random Wave Field. *Proc. R. Soc. Math. Phys. Eng. Sci.* **389**, 241–258 (1983).
47. Kuik, A. J., van Vledder, G. P. & Holthuijsen, L. H. A Method for the Routine Analysis of Pitch-and-Roll Buoy Wave Data. *J. Phys. Oceanogr.* **18**, 1020–1034 (1988).
48. IEC. Part 101: Wave energy resource assessment and characterization. in *Marine energy - Wave, tidal and other water current converters* (BSI Standards Limited 2015, 2015).
49. Hanson, J. L. & Phillips, O. M. Automated Analysis of Ocean Surface Directional Wave Spectra. *J. Atmospheric Ocean. Technol.* **18**, 277–293 (2001).
50. Tracy, B., Devaliere, E., Hanson, J., Nicolini, T. & Tolman, H. *Wind Sea and Swell Delineation for Numerical Wave Modeling* [https://library.wmo.int/pmb\\_ged/wmo-td\\_1442\\_en/WWW/Papers/10th\\_wave\\_paper\\_tracy\\_dhnt.pdf](https://library.wmo.int/pmb_ged/wmo-td_1442_en/WWW/Papers/10th_wave_paper_tracy_dhnt.pdf) (2007).
51. Vincent, L. & Soille, P. Watersheds in Digital Spaces: An Efficient Algorithm Based on Immersion Simulations. *IEEE Trans. Pattern Anal. Mach. Intell.* **13**, 583–598 (1991).
52. Hasselmann, S., Brüning, C., Hasselmann, K. & Heimbach, P. An improved algorithm for the retrieval of ocean wave spectra from synthetic aperture radar image spectra. *J. Geophys. Res. Oceans* **101**, 16615–16629 (1996).
53. Young, I. R. & Glowacki, T. J. Assimilation of altimeter wave height data into a spectral wave model using statistical interpolation. *Ocean Eng.* **23**, 667–689 (1996).
54. Voorrips, A. C., Makin, V. K. & Hasselmann, S. Assimilation of wave spectra from pitch-and-roll buoys in a North Sea wave model. *J. Geophys. Res. Oceans* **102**, 5829–5849 (1997).
55. Devaliere, E.-M., Hanson, J. L. & Luetlich, R. Spatial Tracking of Numerical Wave Model Output Using a Spiral Search Algorithm. *The 2009 WRI World Congress on Computer Science and Information Engineering* 404–408 (IEEE, 2009).
56. Jose-Henrique, A. Source Code for: WAVEWATCH-III.v6.07 NOAA-EMC. *Github* <https://github.com/NOAA-EMC/WW3/releases/tag/6.07> (2020).
57. Hanson, J. L. & Jensen, R. E. *Wave system diagnostics for numerical wave models*. <http://www.waveworkshop.org/8thWaves/Papers/E3.pdf> (2004).
58. Hanson, J. L., Tracy, B. A., Tolman, H. L. & Scott, R. D. Pacific Hindcast Performance of Three Numerical Wave Models. *J. Atmospheric Ocean. Technol.* **26**, 1614–1633 (2009).
59. Jiang, X., Song, Z. & Bao, Y. Global 3-hourly wind-wave and swell data for wave climate and wave energy resource research from 1950 to 2100. *ScienceDB* <https://doi.org/10.57760/sciencedb.02242> (2022).
60. Yu, W., Qiao, F., Yuan, Y. & Pan, Z. Numerical modelling of wind and waves for Typhoon Betty (8710). *Acta Oceanol. Sin.* **16**, 459–473 (1997).
61. Qiao, F., Chen, S., Li, C., Zhao, W. & Pan, Z. The study of wind, wave, current extreme parameters and climatic characters of the South China Sea. *Mar. Technol. Soc. J.* **33**, 61–68 (1999).
62. Qiao, F.-L., Ma, J., Yang, Y. & Yuan, Y. Simulation of the Temperature and Salinity Along 36°N in the Yellow Sea with a Wave-Current Coupled Model. *J. Korean Soc. Oceanogr.* **39**, 35–45 (2004).
63. Zhao, B. *et al.* Sensitivity of typhoon modeling to surface waves and rainfall. *J. Geophys. Res. Oceans* **122**, 1702–1723 (2017).
64. Jiang, B. *et al.* Assessment of wave energy resource of the Bohai Sea, Yellow Sea and East China Sea based on 10-year numerical hindcast data. *OCEANS 2016* <http://ieeexplore.ieee.org/document/7485434> (2016).
65. Sun, M., Yin, X., Yang, Y. & Wu, K. An effective method based on dynamic sampling for data assimilation in a global wave model. *Ocean Dyn.* **67**, 433–449 (2017).
66. Wang, G., Zhao, C., Xu, J., Qiao, F. & Xia, C. Verification of an operational ocean circulation-surface wave coupled forecasting system for the China's seas. *Acta Oceanol. Sin.* **35**, 19–28 (2016).
67. Qiao, F. *et al.* China published ocean forecasting system for the 21st-Century Maritime Silk Road on December 10, 2018. *Acta Oceanol. Sin.* **38**, 1–3 (2019).
68. *Aviso+ Grided Wind/Wave Products* <https://www.aviso.altimetry.fr/en/data/products/wind/wave-products/mswh/mwind.html>
69. Hersbach, H. *et al.* ERA5 hourly data on single levels from 1959 to present. *Copernicus Climate Change Service (C3S) Climate Data Store (CDS)* <https://doi.org/10.24381/cds.adbb2d47> (2018).
70. Jiang, X., Song, Z. & Bao, Y. Custom Code of 'Global 3-hourly wind-wave and swell data for wave climate and wave energy resource research from 1950 to 2100' *ScienceDB* <https://doi.org/10.57760/sciencedb.02893> (2022).

## Acknowledgements

This paper is supported by the Marine S&T Fund of Shandong Province for Laoshan Laboratory (LSKJ202202100), National Key Research and Development Program of China (2022YFC3104801), National Natural Science Foundation of China (42022042 and 41821004), China-Korea Cooperation Project on Northwest Pacific Marine Ecosystem Simulation under Climate Change, and CAS Interdisciplinary Innovation Team (JCTD-2020-12).

## Author contributions

Xingjie Jiang and Zhenya Song conceived the research and drafted the manuscript. Xingjie Jiang carried out the experiments and produced the dataset. Botao Xie and Zhenya Song provided scientific guidance for this work. Ying Bao contributed to preparing the forcing data from FIO-ESM v2.0 CMIP6 experiments. All authors reviewed and edited the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Correspondence** and requests for materials should be addressed to Z.S.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://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