

Evaluating the physical and biogeochemical state of the global ocean component of UKESM1 in CMIP6 Historical simulations

Andrew Yool¹, Julien Palmiéri¹, Colin G. Jones², Lee de Mora³, Till Kuhlbrodt⁴, Ekatarina E. Popova¹, A.J. George Nurser¹, Joel Hirschi¹, Adam T. Blaker¹, Andrew C. Coward¹, Edward W. Blockley⁵, and Alistair A. Sellar⁵

¹National Oceanography Centre, European Way, Southampton SO14 3ZH, UK

²National Centre for Atmospheric Science, University of Leeds, Leeds LS2 9JT, UK

³Plymouth Marine Laboratory, Prospect Place, Plymouth PL1 3DH, UK

⁴National Centre for Atmospheric Science, University of Reading, Earley Gate, Reading RG6 6BB, UK

⁵Met Office, FitzRoy Road, Exeter, Devon EX1 3PB, UK

Correspondence: Andrew Yool (axy@noc.ac.uk)

Abstract. The ocean plays a key role in modulating the climate of the Earth system (ES). At the present time it is also a major sink both for the carbon dioxide (CO₂) released by human activities as well as for the excess heat driven by the resulting atmospheric greenhouse effect. Understanding the ocean's role in these processes is critical for model projections of future change and its potential impacts on human societies. A necessary first step in assessing the credibility of such future projections is an evaluation of their performance against the present state of the ocean. Here we use a range of observational **properties fields** to validate the physical and biogeochemical performance of the ocean component of UKESM1, a new Earth system **model** (ESM) for CMIP6 built upon the **HadGEM3-HadGEM3-GC3.1** physical climate model. Analysis focuses on the realism of the ocean's physical state and circulation, its key elemental cycles, and its marine productivity. UKESM1 generally performs well across a broad spectrum of properties, but it exhibits a number of notable biases. Physically, these include a global warm bias inherited from model spin-up, excess northern sea-ice but insufficient southern sea-ice, and sluggish interior circulation. Biogeochemical biases found include shallow remineralisation of sinking organic matter, excessive iron stress in regions such as the Equatorial Pacific, and generally lower surface alkalinity that results in decreased surface and interior dissolved inorganic carbon (DIC) concentrations. The mechanisms driving these biases are explored to identify consequences for the behaviour of UKESM1 under future climate scenarios, and avenues for model improvement. Finally, across key biogeochemical properties, UKESM1 improves in performance relative to its CMIP5 precursor, and **compares favourably to performs well alongside its** fellow members of the CMIP6 ensemble.

Key points

- Physical and biogeochemical evaluation of the ocean component of the UKESM1 model
- Identification and investigation of present-day biases in major properties
- UKESM1 improves on CMIP5 predecessor and compares well with CMIP6 ensemble

Copyright statement. The works published in this journal are distributed under the Creative Commons Attribution 4.0 License. This licence does not affect the Crown copyright work, which is re-usable under the Open Government Licence (OGL). The Creative Commons Attribution 4.0 License and the OGL are interoperable and do not conflict with, reduce or limit each other. © Crown copyright 2020

1 Introduction

25 The climate dynamics of the Earth system are a product in large part of the two interacting geophysical fluids at the planet's surface: the atmosphere and the ocean. Both are reservoirs for heat and the greenhouse gas carbon dioxide (CO₂), one of several climatically-relevant chemical constituents. Because of the high specific heat capacity of water, as well as the chemical buffering capacity of seawater, the ocean stores the majority of the Earth system's active reserves of both. Over the past few centuries, the atmospheric concentration of CO₂ has risen exponentially from its quasi-stable interglacial background of around
30 278 ppm to more than 400 ppm. This growth is largely ~~in response to~~ driven by the release of CO₂ through anthropogenic processes such as fossil fuel combustion, land clearance and cement production. This change in the ~~chemical composition~~ CO₂ airborne fraction of the atmosphere has also altered its radiative transfer properties toward retaining a greater fraction of outgoing long-wave radiation, resulting in atmospheric warming and change to the climate of the Earth system. Further, for the reasons identified above, the ocean is the destination for the majority of these anthropogenic perturbations in both heat and
35 carbon dioxide (e.g. Archer, 2005; Kuhlbrodt et al., *subm.*).

Leaving aside the relatively static inventory within the geosphere, the Earth's carbon cycle partitions this element dynamically between atmosphere, ocean and land systems, including the living systems of the marine and terrestrial biosphere. While ongoing climate change is driven in the first instance by change in carbon (as CO₂) in the atmosphere, this reservoir represents only approximately 1.4% of the total (pre-industrial) dynamic pool (Ciais et al., 2013), compared with 6.0% for land systems
40 (excluding permafrost) and 92.6% for ocean systems (excluding seafloor sediments). This dominance of the ocean reflects the solubility of inorganic carbon in seawater, and ultimately the majority fraction of these anthropogenic emissions is expected to be absorbed into the ocean (Archer, 2005). However, the magnitude of this, as well as the rate at which it occurs is dependent upon a raft of physico-chemical and biological processes, including surface solubility, deep ocean ventilation and circulation, and biological uptake and deep sequestration via sinking biogenic particles. Representing this uptake within an ESM requires
45 realistic performance across many aspects of its simulated ocean state, both physical and biogeochemical, and both surface and interior.

The situation is similar for heat, with observations over recent decades showing a clear upward trend in ocean heat content since the 1960s at the earliest, and accelerating since the 1990s (Levitus et al., 2012; Cheng et al., 2017). Approximately 90% of the anthropogenic imbalance in the Earth's heat content is stored within the ocean (Meysignac et al., 2019). Consequently,
50 and similarly to carbon, simulating this important property requires ESMs to accurately represent a broad range of physical phenomena, such as ocean circulation and mixing that distribute heat, as well as sea-ice that caps its exchange and affects albedo.

This manuscript is concerned with the realism of the ocean component of UKESM1, during CMIP6 Historical period simulations (1850-2014; Eyring et al., 2016). It has three primary goals:

- 55
- First, to evaluate the performance of UKESM1 against observational metrics, and identify biases in physical and biogeochemical properties.
 - Second, to identify the first-order causes of biases found, to elucidate where modelled processes may be less realistic.

– Third, to identify ~~specific weaknesses and avenues for future model improvement or development~~ avenues for addressing model limitations and weaknesses in future versions.

60 Model performance is evaluated across a broad range of properties to identify biases, with analysis focusing on the near-present period of 2000-2009 because of the greater availability of observational data in recent decades. Overall, the manuscript aims to facilitate subsequent more in-depth analyses of the model by identifying ocean states or processes where its representation is weaker. A summary analysis across all of UKESM1’s components can be found in Sellar et al. (2019).

65 The manuscript is structured as follows. A brief introduction to UKESM1 is presented, with an emphasis on its ocean components, followed by outlines of the model simulations used and the observational datasets selected for their evaluation. Results are then presented for the physical ocean, sea-ice and marine biogeochemistry components, with surface and interior bulk properties, dynamical and biogeochemical processes, and time-series examined. Discussion is focused on the major biases identified, proposals for reducing these in future model revisions, and an evaluation of UKESM1 in the context of peer (and precursor) CMIP models.

2.1 Earth system model

This study utilises UKESM1, a new state-of-the-art model built to ~~describe~~ simulate the coupled physical and biogeochemical dynamics of the Earth system including its atmosphere, ocean and land systems. UKESM1 uses the Hadley Centre Global Environment Model version 3 Global Coupled (GC) version 3.1 configuration, ~~HadGEM3-GC3~~ HadGEM3-GC3.1 (Williams et al., 2017; Kuhlbrodt et al., 2018), as its core physical climate model. This is then extended through the addition of interactive stratospheric–tropospheric trace gas chemistry, land biogeochemistry and ecosystem dynamics, and ocean biogeochemistry. In addition to the internal dynamics of these components, the resulting ESM includes couplings between them to represent potential feedback processes or interactions that may impact the time-evolution of the modelled climate. Sellar et al. (2019) provides an overview of UKESM1, including its development and tuning, while Yool et al. (2020) describes the spin-up of its pre-industrial control (piControl) state ahead of Historical period (1850-2014) simulations.

Supplementary Figure S1 shows a schematic overview of the constituent models of UKESM1. In outline, UKESM1 is comprised of ~~distinct~~ closely-coupled atmosphere and land ~~blocks that are coupled closely together, alongside an ocean and sea-ice block that is linked to the atmosphere~~ submodules that are linked through an explicit coupler module, OASIS3-MCT_3.0 (Valcke, 2013; Craig et al., 2017), to coupled ocean and sea-ice submodules. All three major ES components – atmosphere, land and ocean – are themselves built from submodels that separately represent domains such as physical dynamics, biogeochemistry and ecosystem dynamics.

The physical dynamics of the atmosphere of UKESM1 are represented by GA7.1 (Mulcahy et al., 2018; Walters et al., 2019), which includes processes such as mass transport, radiative transfer, thermodynamics and the water cycle. The UK Chemistry and Aerosols model (UKCA; Morgenstern et al., 2009; O’Connor et al., 2014) is coupled to GA7.1, and includes stratospheric and tropospheric chemistry together with separate aerosol (Mann et al., 2010) and dust schemes (Woodward, 2011). UKESM1 adds several couplings that are absent in GA7.1, including natural emissions of monoterpenes, dimethyl sulphide (DMS) and primary marine organic aerosols (PMOA), all of which are calculated dynamically from land and ocean components and which permit additional climate feedbacks. The atmosphere in UKESM1 also serves as a conduit for mineral dust, transferring this from bare soil on land into the ocean where it can fuel biological production and CO₂ uptake. Mulcahy et al. (2018), Sellar et al. (2019), Archibald et al. (2020) and Mulcahy et al. (2020) provide further details of the atmospheric chemistry and aerosol schemes in UKESM1.

Physics and biogeochemistry on land in UKESM1 is represented by the Joint UK Land Environment Simulator (JULES; Best et al., 2011; Clark et al., 2011). This is closely coupled to the Top-down Representation of Interactive Foliage and Flora Including Dynamics model (TRIFFID; Cox, 2001; Jones et al., 2011), which represents plant and soil dynamics on land. TRIFFID developments new to CMIP6 include updated plant parameterisations (Kattge et al., 2011), increased plant functional types (Harper et al., 2016), the production of volatile organic compounds (Pacífico et al., 2015), and nitrogen limitation of terrestrial primary production and carbon uptake (Wiltshire et al., *subm.*). TRIFFID represents land-use by agriculture by reserving grid

cell time-varying fractions for occupation by crops and pasture. For further details of UKESM1's land component, please refer to Sellar et al. (2019).

105 The physical ocean component in UKESM1 makes use of the Nucleus for European Modelling of the Ocean framework (NEMO; Madec et al., 2016). This is comprised of an ocean general circulation model, Océan PARallélisé version 9 (OPA9; Madec et al., 1998; Madec, 2008), and is coupled here to a separate sea-ice model, the Los Alamos Sea Ice Model version 5.1.2 (CICE; Hunke et al., 2015). OPA9 is a primitive equation model of ocean dynamics, and is used within UKESM1 at a horizontal resolution of approximately 1° on a tripolar grid (Madec and Imbard, 1996) with enhanced equatorial res-
110 olution (the extended ORCA1 grid, or eORCA1). This shared configuration of NEMO, dubbed “shaconemo”, is used by a number of European research groups, and many of its grid resolution-dependent settings are aligned with these other ESMs (NEMO v3.6_stable; available from <http://forge.ipsl.jussieu.fr/shaconemo>). Some other parameterisations (typically resolution-independent) are drawn from the GO6 configuration of NEMO developed in the UK (Storkey et al., 2018). More complete descriptions of the NEMO and CICE configurations used in UKESM1 (GO6, GSI8), including details of its sensitiv-
115 ity and resulting tuning, can be found in Storkey et al. (2018), Ridley et al. (2018) and Kuhlbrodt et al. (2018), while Kuhlbrodt et al. (subm.) investigates ocean heat uptake.

Marine biogeochemistry in UKESM1 is represented by the Model of Ecosystem Dynamics, nutrient Utilisation, Sequestration and Acidification (MEDUSA-2.1). MEDUSA-2.1 is “intermediate complexity” with a double size-class ecosystem that represents phytoplankton, zooplankton and particulate detrital pools, and which explicitly includes the biogeochemical cycles
120 of nitrogen, silicon and iron nutrients as well as the cycles of carbon, alkalinity and oxygen ([Supplementary Figure S2](#)). During its inclusion within UKESM1, a number of changes were introduced from its earlier predecessor model, MEDUSA-2, described in Yool et al. (2013), and the version used here is identified as MEDUSA-2.1, to distinguish it. These changes include ~~1. adoption of the MOCSY-2.0 carbonate chemistry scheme of Orr and Epitalon (2015); 2. inclusion of surface seawater updated carbonate chemistry Orr and Epitalon (2015), the addition of empirical submodels of dimethyl sulphide (DMS) concentration using the empirical submodel of (Anderson et al., 2001); 3. code improvements including:~~ [Anderson et al., 2001 and primary marine organic aerosol \(PMOA; Gantt et al., 2011; Gantt et al., 2012\), and code improvements such as](#) variable volume (VVL) and the XML Input-Output Server (XIOS) (Meurdesoif, 2013). [Within UKESM1, MEDUSA interacts with other model components via the following feedback connections: atmosphere-ocean exchange of CO₂; ocean-to-atmosphere fluxes of DMS and PMOA; deposition of terrestrial iron to the ocean via atmospheric dust transport.](#) A more complete description
125 of MEDUSA-2.1 can be found in ~~the~~ Appendix A.

[In addition to the biogeochemical tracers of MEDUSA-2.1, UKESM1 includes the chlorofluorocarbon tracer, CFC-11. This artificial tracer has an atmospheric time-history analogous to that of anthropogenic CO₂, and can be used as a marker for recently ventilated watermasses \(Key et al., 2004\). It can be measured from seawater samples with high accuracy, and provides an additional measure here for evaluating simulated circulation.](#)

135 [UKESM1 is the successor model to its CMIP5 predecessor, HadGEM2-ES \(Collins et al., 2011\). Many of its components are evolved versions of those in the earlier model, including its land surface, physical atmospheric core, and atmospheric chemistry components \(Sellar et al., 2019\). However, in the specific case of the ocean in UKESM1, its dynamical core, grid](#)

140 domain, sea-ice, and marine biogeochemistry are wholly new and replace the corresponding components in HadGEM2-ES. Consequently, there is no direct traceability between the oceans of the two generations of CMIP model. Nonetheless, as part of the assessment of UKESM1, elements of its performance relative to that of HadGEM2-ES are examined in Section 4.2.

2.2 ~~Model~~ CMIP6 simulations

This study utilises simulations of the UKESM1 model performed as part of the 6th phase of the Coupled Model Intercomparison Project (CMIP6). Model output is taken from the piControl and Historical simulations of CMIP6, and from an ensemble of 9 members, consistent with Sellar et al. (2019). Each ensemble member represents a branch at a different time point from
145 the piControl, after which the new simulation experiences time-varying changes in atmospheric and land-use properties characteristic of the Historical period from start-1850 to end-2014. Ensemble branch points were chosen selectively to span the variability in the model's multi-decadal behaviour (Sellar et al., 2019). To achieve this the model's behaviour across two major ocean modes was sampled: the Atlantic Multi-decadal Oscillation (AMO; Kerr, 2000), and the Inter-decadal Pacific Oscillation (IPO; Zhang et al., 1997; Power et al., 1999). Supplementary Table S1 lists the local run IDs of the simulations comprising the
150 ensemble, together with their branch times from the piControl. The mean of this 9 member ensemble is used throughout the following analysis, except where stated otherwise.

2.3 ~~Model~~ Datasets and evaluation

Model analysis in this study is focused on a subset of ocean properties. More complete evaluations of other UKESM1 components can be found in the dedicated studies of Mulcahy et al. (2018) and Mulcahy et al. (2020) (aerosols), Archibald et al.
155 (2020) (atmospheric chemistry) and Andrews et al. (2019) (radiative forcing, feedbacks and climate sensitivity). Sellar et al. (2019) provides a summary overview of the full model.

~~Evaluation uses the period 2000-2009 of the CMIP6 Historical simulation and compares to corresponding periods of observational data. Some evaluated properties are not as comprehensively sampled, but the same time period is likely to be representative of the ocean's state so we use this for consistency. The results shown make use of monthly climatologies of both model output and observational data (where available) for this period.~~

160

The specific observational datasets used for evaluation are as follows:

- World Ocean Atlas 2013, for ocean ~~physies~~ physical (interior; Locarnini et al., 2013; Zweng et al., 2013) and biogeochemistry (interior, surface; Garcia et al., 2014; Garcia et al., 2014) fields
- Hadley Centre Sea Ice and Sea Surface Temperature (HadISST.2.2; Titchner and Rayner, 2014), for ocean SST and
165 sea-ice fields
- National Sea Ice Data Centre for sea-ice thickness (Stroeve and Meier, 2016) and sea-ice index (Fetterer et al., 2017)
- Estimating the Circulation and Climate of the Ocean (ECCO) V4r4 (Forget et al., 2015; Fukumori et al., 2019), for ocean ~~circulation~~ hydrodynamic circulation state

- 170 – Smeed et al. (2018) for RAPID-MOCHA time-series measurements of the Atlantic meridional overturning circulation (AMOC) at 26°N
- SeaWiFS (O’Reilly et al., 1998), for surface ocean chlorophyll concentration
- Oregon State University Ocean Productivity group, for VGPM (Behrenfeld and Falkowski, 1997), Eppley-VGPM (Carr et al., 2006) and CbPM (Westberry et al., 2008) vertically-integrated primary production
- Rödenbeck et al. (2013) for observationally-derived global air-sea CO₂ flux and surface pCO₂
- 175 – Lana et al. (2011) for surface dimethyl sulfide (DMS) concentrations
- Global Ocean Data Analysis Project v1.1 (Key et al., 2004) and v2 (Key et al., 2015; Lauvset et al., 2016), for interior and surface carbonate biogeochemistry, including anthropogenic CO₂
- Moriarty and O’Brien (2013) for the COPEPOD dataset of gridded zooplankton biomass

Links to these datasets are given in Appendix D.

180 In addition, several derived variables are calculated from observational and model fields.

- 185 – Mixed layer depth (MLD) is calculated in the same way from both observed and modelled 3D fields of potential temperature. MLD is determined to be the depth at which the vertical profile of potential temperature is 0.5°C lower than that at the depth of 5 m. Alternative MLD schemes using similar thresholds in potential density (either fixed or variable with temperature) were also examined, but global coverage was less complete with these (especially in sea-ice regions), so the potential temperature criterion was favoured.
- Modelled integrated AMOC and Drake Passage transports are calculated here using the BGC-val toolkit (de Mora et al., 2018). In the case of AMOC, the calculations are based on those of Kuhlbrodt et al. (2007) and McCarthy et al. (2015) and use the cross-sectional area at the 26°N transect to calculate the maximum depth-integrated current. Drake Passage transport is calculated following Donohoe et al. (2016) as the total, depth-integrated current along a north-south transect between the South American continent and the Antarctic Peninsula. The methods for both transports are described in de Mora et al. (2018).
- 190 – Model anthropogenic CO₂ is estimated by differencing DIC fields from the Historical simulation of each ensemble member with the corresponding DIC field from the piControl at the same relative timepoint. For example, we estimate anthropogenic CO₂ in 1990 from a given Historical ensemble member as the difference between this member’s DIC field at this particular time and the DIC field from the piControl simulation from the same timepoint, i.e. the time that corresponds to 140 years (i.e. 1990 – 1850 = 140) after the Historical ensemble member branched from the piControl. This approach aims to account for drift in the simulations, although it omits changes driven by divergence in circulation and biogeochemistry between the Historical and piControl simulations. These are assumed to be small in this method.
- 195

200 Evaluation primarily uses the period 2000-2009 of the CMIP6 Historical simulation and compares to corresponding periods
of observational data. Some evaluated properties are not as comprehensively sampled, but we assume that the same time
period is likely to be representative of the ocean's state and use this for consistency. The results shown make use of monthly
climatologies of both model output and observational data (where available) for this period. A number of figures illustrate
observed and modelled properties (and the biases of the latter) for the June-July-August (JJA) and December-January-February
(DJF) meteorological seasons that correspond respectively to northern hemisphere summer and winter (and southern hemisphere
205 winter and summer).

Throughout, fields of observational and model properties are plotted on their original horizontal and vertical grids. Where
these properties are directly intercompared, for instance in difference plots, observational fields are first regridded to the model
grid (using the `scatteredInterpolant` function of Matlab v2020a). In Section 4.2, horizontal fields of UKESM1 output
are compared with those from fellow CMIP6 models, and here all models are regridded to a common, uniform 1° grid.

3.1 Surface physical ocean

Figure 1 shows observed (HadISST; Titchner and Rayner, 2014) and simulated global-scale sea surface temperature (SST) for summer and winter in both hemispheres, together with (model - observed) patterns of difference. The model reproduces the main observed features, including latitudinal and seasonal gradients, upwelling regimes and major fronts. A number of biases are also evident, including warm biases up to 4°C in upwelling regimes (especially the equatorial Pacific), a general warm bias in the Southern Ocean, cool biases of up to -2°C throughout the subtropics, and a marked cold bias in the North Atlantic of greater than -4°C. The former Pacific biases occur in December-January-February (DJF) when tropical atmospheric convection is primarily over the western Pacific warm pool, and the east-west pressure gradient is seasonally at a maximum. This gradient drives east-west wind stress and equatorial Ekman-induced upwelling, and a poor representation of this in UKESM1 likely leads to reduced upwelling and the warm SST bias. A warm bias close to the North American coastline and strong cold bias in the western North Atlantic occur due to resolution-dependent errors where the Gulf Stream separates too far north and then extends too zonally across the North Atlantic (Marzocchi et al., 2015; Hirschi et al., 2020). ~~Marzocchi et al., 2015, and~~ Similar but less marked biases occur in the Pacific in association with the Kuroshio Current. In general, surface temperature biases in the model have strong latitudinal patterns associated with major currents and patterns of upwelling and downwelling, and are persistent across the seasons. To illustrate the full seasonal cycle, Supplementary Figure S3 shows ~~Hovmöller~~ Hovmöller diagrams of latitudinal mean observed and simulated SST.

SST exhibits a number of major climate modes such as the Interdecadal Pacific Oscillation (IPO) and Atlantic Multidecadal Oscillation (AMO) that can introduce persistent and large-scale shifts in temperature that are of comparable magnitude to the model biases identified above. For instance, the IPO has a negative index (cooler than reference) during the time period shown in Figure 1, but a positive index (warmer than reference) during the preceding two decades (Salinger et al., 2001; Hu et al., 2018). Models also have climate modes, but these can be out of phase with those observed, and they may occlude or exaggerate biases. Supplementary Figure S4 partially addresses this by repeating the difference plot from Figure 1, but for the three preceding decades. The resulting patterns of model-observation difference are generally consistent between the decades and for both seasons, suggesting that they represent model biases rather than variability mismatch. In particular, persistent features include the strong cold bias in the western North Atlantic, warm biases in the equatorial Atlantic and Pacific basins (the latter seasonally), and a general warm bias in the Southern Ocean. As most other observational datasets used in the evaluation of UKESM1 properties are more restricted in the time periods they have available, similar analyses are more difficult. However, given the primary role of SST in many ocean processes, the apparent dominance of model bias in SST over its temporal variability is suggestive that mismatches in major climatic modes are of secondary importance in our analysis.

Supplementary Figure S5 parallels Figure 1, showing the observed (WOA, 2013; Zweng et al., 2013) and simulated sea surface salinity (SSS) for summer and winter, together with (model - observed) differences. UKESM1 shows a general negative bias in SSS (≈ 1 PSU), but with significant regions of positive bias in the tropical Atlantic and Indian oceans (< 1 PSU). There are also “hotspots” of bias in the Bay of Bengal (positive), off the west (negative) and east (positive) coastline of equatorial

245 South America, in the Yellow and East China seas (negative), and in the Arctic (both positive and negative). These regions are
mostly located close to major riverine inputs, and likely reflect model inaccuracies in the precise location and magnitude of
associated freshwater additions.

Remaining with the surface ocean but moving to high latitude regions, Figure 2 shows the observed and simulated sea-ice concentrations at the seasonal maxima, March in the Arctic and September in the Antarctic (HadISST; Titchner and Rayner, 2014). In general terms, the model reproduces the observed northern hemisphere sea-ice patterns, with complete ice cover in 250 the main Arctic basin, Baffin Bay down to Davis Strait, and Hudson Bay, cover on the eastern margins of Newfoundland and Greenland, and bounding the Barents Sea. In the Arctic, simulated maximum sea-ice area is $15.3 \times 10^6 \text{ km}^2$, compared with an observational maximum of $13.9 \times 10^6 \text{ km}^2$. This relationship is reversed in the Antarctic, with a simulated maximum of $11.8 \times 10^6 \text{ km}^2$ compared to $16.3 \times 10^6 \text{ km}^2$ observed. As the bottom row of Figure 2 shows, this general pattern of excess sea-ice in the Arctic and a deficit around Antarctica generally persists seasonally, with a modelled Arctic minimum of 8.7 255 compared to $4.7 \times 10^6 \text{ km}^2$ observed, and a model Antarctic minimum of 2.7 compared to $2.6 \times 10^6 \text{ km}^2$ observed. Modelled Arctic sea-ice also reaches its seasonal minimum slightly earlier than observed, in August rather than September. In the Arctic, sea-ice ~~is typically~~ typically persists for multi-year ~~;~~ and this positive bias in modelled area periods, such that this bias towards excess ice area in UKESM1 is accompanied by ~~excessively thick~~ sea-ice cover that is also excessively thick. Thicknesses are up to 5 m in the simulated “dome” of sea-ice over the north pole, compared to flatter observational estimates that are closer to 260 3 m (Supplementary Figure S6; Stroeve and Meier, 2016).

In response to ongoing climate change, Arctic sea-ice shows one of the most pronounced trends within the Earth system over recent decades (Brennan et al., 2020). Figure 3 shows simulated Arctic and Antarctic sea-ice extent over the full Historical period (1850–2014), together with observational estimates (HadISST, Titchner and Rayner, 2014; NSIDC, Fetterer et al., 2017) and for recent decades. Much as with sea-ice extent itself, UKESM1 performs ~~best~~ better in the Arctic, with similar negative 265 trends since 1980. In the Antarctic, however, the discrepancy in seasonal extent already noted is exacerbated by a negative trend in maximum sea-ice extent in UKESM1 opposite to the rising trend actually observed (although this observed trend may be reversing; Parkinson, 2019).

The Earth’s ocean and atmosphere interact principally at their interface, but turbulent mixing of the ocean ventilates its upper layer with both physical and biogeochemical consequences. As described in Section 2.3, this layer is characterised 270 from both observational and model fields of 3D potential temperature using a 0.5°C change criterion. Figure 4 shows the observed and modelled thickness of this mixed layer, together with (model - observed) patterns of difference. ~~Mixed-layer thickness (MLD) is estimated here as the depth at which the vertical profile of potential temperature changes by 0.5°C from its 5-m value (WOA; Locarnini et al., 2013).~~ Again, the model reproduces the main features of the ocean, including strong seasonality at high latitudes, deep mixed layers ($> 100 \text{ m}$) throughout the year in the Southern Ocean (away from sea-ice), and 275 shallow mixed layers ($< 50 \text{ m}$) in equatorial upwelling regions. When and where the mixed layer is shallow, the model tends to exaggerate this with even shallower mixed layers, most noticeably during the summer at temperate latitudes. At subpolar latitudes, in the Southern, Atlantic and Pacific oceans, deep mixing in the winter is more pronounced in the model, with larger areas experiencing mixing to deeper than 500 m. These model biases towards both shallower and deeper mixed layer depths

are more clearly visible in Figure 5, which shows the frequency at which different mixed layer depths occur seasonally. While median frequencies are similar between the model and those observation-derived, modelled summer and winter distributions can be seen to be shifted shallow and deep respectively.

Table 1 lists the global means (or mean integrals) of these surface physical properties across both the full Historical period and the corresponding piControl period. For both of these simulation ensembles, the variability and ranges of each of these properties are given, together with the simple linear trend over the full 165 y period.

285 3.2 Interior physical ocean

Switching to the ocean interior, Figures 6 and 7 respectively illustrate zonally-averaged depth profiles of temperature and salinity along so-called “thermohaline transects” of the Atlantic, Southern and Pacific oceans, for both UKESM1 and observations (Locarnini et al., 2013; Zweng et al., 2013). These transects are created from basin zonal means of the plotted properties. They track southward down the Atlantic into the Southern, before reversing direction to travel northward from the Southern into the Pacific, with the aim of broadly following watermass properties from young, freshly-ventilated North Atlantic Deep Water (NADW) through to much older North Pacific waters. For the purposes of this transect, the Arctic Ocean is considered a northern extension of the Atlantic, while the Indian Ocean – west of the Malay Archipelago, and including its sector of the Southern Ocean – is entirely omitted from consideration. In both cases, observed and modelled interior properties are shown, together with a difference plot to highlight biases.

295 For ocean temperature, while there are spots of cooler biases in the upper ocean (< 1000 m), temperature is generally positively biased in the upper 3000 m. This is more pronounced in the Atlantic basin, in particular at tropical latitudes, where midwater (100–1000 m) biases up to 4°C are found in the model. The bias in southward-moving NADW (> 1000 m) is consistent with the warm bias in SST shown in its subpolar source regions in Figure 1. Comparable Pacific biases are much lower, and tropical latitudes instead show a cold bias in the upper 500 m. At depth (> 3000 m), both basins show negative biases, which again are more pronounced in the Atlantic. Southern Ocean temperatures exhibit small positive biases, most clearly in the Atlantic sector, although these switch sign at depth into the Atlantic proper as already mentioned. Patterns of ocean salinity broadly mirror those of temperature in the Atlantic basin, with corresponding positive biases in the upper 3000 m, and negative biases below. The model’s Pacific basin is more uniformly fresh in the upper 1000 m, with smaller positive biases beneath, and negligible biases below 3000 m. Overall, temperature and salinity patterns indicate that the Atlantic is a warmer, more evaporative basin in the model, with its most positive upper ocean biases located there, as well as its largest negative biases in the deep ocean. Supplementary Figure S7 shows the corresponding patterns in potential density anomaly (σ_θ ; referenced to atmospheric pressure). These show the model ocean, particularly the Pacific basin, to be more stratified vertically compared to observations, with generally lower density surface waters (< 1000 m) overlying more dense deep waters. This bias suggests that the model’s parameterisation of vertical mixing may be insufficient, reducing the transfer of heat from the surface to deeper layers (and potentially weakening the deeper circulation; see below).

310 This pattern of biases in the zonal sections above indicates differences in the balance of interior watermasses in UKESM1 compared to that of the real ocean. Observationally, zonally-averaged North Atlantic circulation below 1000 m is dominated

by the transports associated with North Atlantic Deep Water (NADW) and the Antarctic Bottom Water (AABW). NADW is produced by the the subduction of cool, salty water at subpolar latitudes in north of the basin, and its southward-moving cell overlies a denser cell of Antarctic Bottom Water (AABW) travelling northward from its production in the Southern Ocean. To illustrate this, the upper panel of Figure 8 shows a reconstruction of the global streamfunction of the ocean’s meridional overturning circulation (MOC), produced by the Estimating the Circulation and Climate of the Ocean consortium (ECCO; Forget et al., 2015; Fukumori et al., 2019). This is an ocean reanalysis product in which the MOC is a result of a model simulation that has been constrained with observations (for a more complete overview, see Jackson et al., 2019). In this, the upper positive (clockwise) overturning cell extends its influence below 2000 m (in red; driven by circulation in the North Atlantic), overlying the negative overturning cell (in blue) of AABW. The lower panel of Figure 8 shows the corresponding MOC in UKESM1. In general, this follows the pattern shown in the ECCO reanalysis, although with a slightly stronger maximum MOC at 40°N, and a weaker AABW cell northward of the Antarctic Circumpolar Current (ACC). We note that the southernmost part of the overturning associated with AABW is stronger in UKESM1 than in ECCO (around 6 Sv against 4 Sv), suggesting that sinking around Antarctica is stronger in UKESM1. Stronger sinking [in UKESM1](#) around Antarctica, combined with a slightly weaker NADW [than observed, indicates a more dominant role for AABW in the model, and](#) is consistent with the colder and fresher [conditions shown for biases found in](#) the deep ocean (particularly the Atlantic) in Figures 6 and 7, as well as biases in biogeochemical fields (see later).

While Figure 8 shows a time- and zonally-averaged state of the MOC, ocean circulation exhibits significant variability (Majewski et al., 2009; Smeed et al., 2018). Annual mean observation-based estimates of the Atlantic MOC (AMOC) from the RAPID-MOCHA array at 26.5°N range from 14.6–19.3 Sv between 2004–2016 (Smeed et al., 2018). In the Southern Ocean, Drake Passage, the channel between the Antarctic Peninsula and South America, focuses the ACC that rings Antarctica, and from intermittent sampling has a transport estimated at 173 ± 11 Sv (Donohoe et al., 2016). Figure 9 shows time-series of both of these major transports across the full Historical period, for all 9 ensemble members [\(and includes RAPID-MOCHA observations of the AMOC\)](#). UKESM1’s pre-industrial AMOC is typically lower than that found by RAPID-MOCHA [\(Yool et al., 2020\)](#) [\(Yool et al., 2020;](#) consistent with the spatial displacement mentioned previously), but strengthens by approximately 3 Sv from 1850 to a maximum of around 17 Sv by the 1990s. This increase in AMOC strength, which ends in UKESM1 around 2000, is almost certainly causally linked to temporal trends in negative radiative forcing driven by anthropogenic aerosol emissions in the northern hemisphere over this period (Menary et al., 2020). Increases in these, driven by industrial activity, cool the north relative to the south, change the inter-hemispheric thermal gradient, and result in increasing AMOC strength in response. Although good observational data is absent prior to the construction of the RAPID-MOCHA array, this rise in AMOC strength is consistent with model reanalysis over this period (Jackson et al., 2016), although possibly overestimated in CMIP6 models such as UKESM1 (Menary et al., 2020). The subsequent decline during first decades of the 21st century matches that found by RAPID-MOCHA (Smeed et al., 2018) and reanalysis (Jackson et al., 2016). The modelled AMOC increase in UKESM1 is absent in the parallel segments of the piControl simulation that do not experience these anthropogenic changes (see the linear trends in Table 1).

Time-averaged over the Historical period (≈ 150 Sv), Drake Passage transport in UKESM1 is lower than that estimated (Donohoe et al., 2016), although across the full ensemble and its long-period variability, the model intermittently reaches the range observed (Figure 9). Throughout the Historical period the ensemble exhibits considerable multi-decadal to centennial scale variability in modelled ACC strength (135–173 Sv); see also Table 1). Unlike AMOC strength, where the ensemble shows a clear trend that all members follow, ACC strength is much less aligned across the ensemble, most clearly in the period 1850–1930. Between 1930–1980, however, the ensemble spread is reduced and most ensemble members exhibit a weak ACC. However, following this point most strengthen notably, recovering from this earlier minimum to reach higher values more consistent with the recent observations. The increase in ACC strength post-1970 is consistent with development of the Antarctic ozone hole and strengthened westerlies over the Southern ocean which then drives a stronger ACC (e.g. Li et al., 2016). Nonetheless, as Figure 9 shows, two of the nine members do not exhibit this minimum around 1970, suggesting that while a forced climate driver may be operating on ACC strength, it cannot completely override internal variability in the Southern Ocean.

3.3 Surface nutrient biogeochemistry

Figures 10 to 16 present model-observation intercomparisons for a range of key surface biogeochemical properties, showing seasonal geographical fields and zonal Hovmöller diagrams (where possible). Table 3 lists global-scale statistics for major biogeochemical properties, including variability and trends across both the full Historical period and the corresponding period of the piControl simulation. Table 3 compares global and regional means for the same properties on an annual-mean timescale with corresponding observational means for the 2000–2009 period. To summarise across these properties, Supplementary Figure S9 additionally shows seasonal and regional Taylor diagrams.

In terms of surface concentrations of the macronutrients that regulate biological productivity in the ocean, UKESM1 shows some shared and some divergent biases. For dissolved inorganic nitrogen (DIN; Figure 10), while the major, circulation-driven features occur (i.e. subtropical gyre lows, upwelling highs), the model is typically biased positive, with excess nutrients most obvious in the tropical Pacific and in the Arctic Ocean (see also Supplementary Figure S10). Globally, the model's mean is 7.8 compared to an observational mean of 5.2 mmol m^{-3} (+48%). However, in regions such as the North Atlantic, the model is biased negative with winter maximum concentrations much lower (≈ 5 vs. $\approx 10 \text{ mmol m}^{-3}$) in this important productive region. The North Pacific, by contrast, exhibits the year-round high nutrient concentrations that characterise this region (12 vs. 10 mmol m^{-3}). However, the spatial distribution of North Pacific DIN, particularly around the Bering Straits, biases inflow concentration to the Arctic Ocean and is responsible for the excess concentration in this region.

In MEDUSA, silicic acid is a key limiting factor for the growth of the model's large phytoplankton, the diatoms. As Figure 11 shows, away from the Southern Ocean where it is strongly biased positive (≈ 63 vs. $\approx 32 \text{ mmol m}^{-3}$; Table 23), the model is typically biased negative. Globally, the model's mean is 10.1 compared to an observational mean of 7.5 mmol m^{-3} (+50%). While silicic acid concentrations are generally low throughout the tropical and subtropical ocean (maxima $< 20 \text{ mmol m}^{-3}$), modelled concentrations are much more depleted throughout the year (maxima $< 5 \text{ mmol m}^{-3}$). In the North Pacific, un-

380 like with DIN, seasonal maximum silicic acid concentrations are significantly lower than observed in this region (4.5 vs. 21.3 mmol m⁻³).

Alongside nitrogen and silicon (the latter for diatoms only), phytoplankton productivity in MEDUSA is additionally limited by the micronutrient, iron. An important source of iron to the ocean is via deposition of aeolian dust that has been lifted from desiccated land surfaces and transported by winds (Tagliabue et al., 2017; Kok et al., 2018). MEDUSA represents this source
385 of iron to the ocean, and in UKESM1 this flux of dust is driven by dynamic land-atmosphere interactions (Woodward, 2011). Figure 12 compares the simulated flux of iron from dust with the observationally-derived dataset of Mahowald (2005). Following Yool et al. (2013), dust is scaled in UKESM1 such that total iron added to the ocean by deposited dust is approximately 2.6 Gmol Fe y⁻¹ (excluding the Mediterranean Sea), and the Mahowald (2005) panel is similarly scaled. In general, UKESM1 exhibits similar spatial patterns to the observational product, including high deposition downwind of arid regions such as the Sahara, and corresponding low deposition where airmasses do not intersect with land such as over the Southern Ocean. However, several key areas of low deposition are more pronounced in the model, including the Southern Ocean, the Peruvian upwelling and the Equatorial Pacific. These regions are also those where excess DIN occurs, indicating that at least one source for these biases may be excessively strong iron limitation on biological activity. To further illustrate this, Supplementary Figure S11 shows the dominant nutrient limitation for both phytoplankton types. Noticably, compared to other runs employing MEDUSA (Yool
390 et al., 2013), iron-stress is more pronounced in UKESM1, especially compared to nitrogen-stress, with the Southern Ocean and almost the whole of the Pacific iron-limited for non-diatom phytoplankton, and diatom phytoplankton iron-stressed across the Equatorial Pacific. Corresponding observational patterns of nutrient stress are more sparsely available (Moore et al., 2013). However, UKESM1's nutrient limitation overlaps the major observed patterns, including widespread nitrogen stress in the Atlantic Ocean, and iron stress throughout the Pacific and Southern oceans, as well as at high latitudes in the North Atlantic (Moore et al., 2013). Nonetheless, the simplicity of MEDUSA prevents it from representing the limitation of phytoplankton found by Moore et al. (2013) for the macronutrient, phosphorus, and the micronutrients, cobalt, zinc and vitamin B12.
400

Switching to the ~~biological community~~marine biology, Figure 13 presents surface chlorophyll, the main light-harvesting pigment used by phytoplankton. Again, the model exhibits both positive and negative biases relative to observations, but with a general positive bias (0.26 vs. 0.22 mg chl m⁻³). Most noticeably, modelled summer concentrations of chlorophyll in the
405 Southern Ocean are biased positive throughout the year, particularly so in the unproductive winter, when the model continues to simulate moderate concentrations even at high latitudes (although winter observations are less reliable or absent). In part, the ~~elevated positive bias of~~ chlorophyll concentrations in UKESM1 are driven by the reduced extent of winter sea-ice in this hemisphere, although, on the observational side, global satellite-based algorithms have also been shown to underestimate surface chlorophyll in this region (Johnson et al., 2013). At the equator, the model is biased positive in the Pacific (0.25 vs.
410 0.18 mg chl m⁻³), while strongly biased negative in the Atlantic (0.07 vs. 0.36 mg chl m⁻³). Meanwhile, in the subtropical gyres, the model simulates lower concentrations than observed throughout, particularly in the Atlantic Ocean, whereas the lowest observed concentrations occur in the southern Pacific subtropics. At high northern latitudes, maximum chlorophyll concentrations are typically slightly lower than those observed, although, much as in the southern hemisphere, moderate winter concentrations extend much further poleward than observed.

415 Figure 14 presents the corresponding distributions of net primary production, the process driving consumption of surface nutrients, biological uptake of dissolved CO₂, and the ultimate source of organic matter for the ocean's food web. The observations shown here are the simple mean of three observation-driven estimates of productivity modelsalgorithms: VGPM (Behrenfeld and Falkowski, 1997); Eppley-VGPM (Carr et al., 2006); and CbPM (Westberry et al., 2008). Generally, although with some of the same model biases already noted, simulated patterns clearly replicate those observed. Integrated globally, 420 modelled productivity across the UKESM1 ensemble averages 44.3 Pg C y⁻¹, compared with an average of 39.5 Pg C y⁻¹ estimated by the three modelsalgorithms. Regionally, the clearest bias lies, again, in the Southern Ocean, where modelled productivity is both greater and geographically more extensive, with a large summer bloom that extends further south towards Antarctica (0.31 vs. 0.10 g C m⁻² d⁻¹). Another discrepancy lies in the tropics, where modelled-UKESM1's productivity is more focused along the equator in the Pacific, and with generally lower productivity in the subtropical gyres. Also, while pro- 425 ductivity is focused in shelf regions in both observations and the model, in the model it extends further into the open ocean than observed, where productivity is generally restricted to a narrow band around the continents. Finally, in terms of seasonal extent, modelled productivity is typically broader, with positive biases extending further polewards during winter in both hemispheres.

Supplementary Figure S12 shows the time-series of net primary production, and its main driver, DIN, across the Historical period for all nine ensemble members. Earlier plots evaluated the geography and phenology of both fields in the early 21st 430 century, but this plot makes it clear that neither property is at equilibrium at this time. Global surface DIN shows a pronounced rise (approximately 5%) from 1950 to around 2000, consistent across the ensemble, but by 2014 this increase has been entirely reversed. Meanwhile, primary production has no clearly comparable 20th century trend, but from around 2000 declines (approximately 2%). In terms of the main production regions, the North Atlantic and the Southern Ocean drive these global signals, with production unsurprisingly lagging that of DIN (Supplementary Figure S12).

435 The critical role of primary production as the source of organic carbon (and chemical energy) on which marine ecology runs means that the realism of its representation in models has consequences across marine biogeochemistry. To illustrate this, Supplementary Figures S13 and S14 compare UKESM1 surface fields of a higher trophic level (mesozooplankton) and a climatically-active biogenic gas (DMS) with observational estimates. As would be expected, both properties scale closely with productivity, and share a number of the same geographical biases. While much of the ocean shows good model-observation 440 agreement, mesozooplankton biomass in the Southern Ocean is significantly elevated in both summer and winter compared with Moriarty and O'Brien (2013)'s dataset, as well as more focused around the Antarctic Polar Front (Supplementary Figure S13). The corresponding biomasses in both seasons in the northern hemisphere are better reproduced, although still with biases, including lower North Pacific mesozooplankton, a region where their abundance has long been known to play a role in seasonal dynamics (Steele and Henderson, 1992). Switching to DMS, the model actually shows pronounced negative biases in 445 the Southern Ocean, in contrast with other properties (Supplementary Figure S14). Elsewhere, regions of high concentration are also typically more geographically confined in the model, with maximum values lower than those observed. The relatively good general agreement with the observational (Lana et al., 2011) dataset in part relates to the tuning of the underlying (Anderson et al., 2001) DMS model, although the divergence where observed concentrations are high, especially the Southern Ocean, suggest this real-world property is a more complex function of primary production than modelled in UKESM1.

450 3.4 Surface carbon biogeochemistry

Figure 15 compares the annual mean surface concentrations of DIC and alkalinity, two key carbonate chemistry properties that constrain the ocean's exchange of CO₂ with the atmosphere. In both cases, the model reproduces the spatial patterns well, with main features such as elevated DIC at high latitudes, a strong Atlantic-Pacific alkalinity gradient, and generally lower concentrations of both at lower latitudes. Globally, both model mean DIC and alkalinity are slightly biased negative compared to observations, with implications for interior concentrations of DIC (see Section 3.5). Noticeable regional biases include positive biases for both properties in the Southern Ocean (particularly around Antarctica), and negative biases in alkalinity in the North Atlantic and (especially) the North Pacific.

Critically linked to surface DIC and alkalinity, Figure 16 shows the observed and modelled patterns of air-sea exchange of CO₂. This is a key Earth system property, as its integrated magnitude modulates the accumulation of anthropogenic CO₂ in the atmosphere with its absorption by sinks such as the ocean and the land. The observational product used here fits a simple ocean mixed layer biogeochemistry scheme to observations of surface ocean CO₂ partial pressure, and then extrapolates this globally (Rödenbeck et al., 2013). Much as with its surface carbonate chemistry, the model reproduces the main features of air-sea CO₂ exchange, including zonal bands of ingassing and outgassing, pronounced equatorial outgassing in the Pacific, and strong seasonal ingassing at high latitudes in the northern hemisphere. However, the model also exhibits a number of biases in its regional and seasonal patterns of flux. While observations suggest that the Southern Ocean is a complex mix of summer ingassing and winter outgassing, the model is biased towards ingassing, with weaker and more geographically limited outgassing in the southern winter. Further, though showing similar patterns to those observed, the model exaggerates seasonal ingassing in the northern hemisphere, particularly during late winter and spring at subtropical latitudes. Note that, again, the reliability of this observational product is lower in less sampled regimes, such as the ~~SO~~ [Southern Ocean](#) and during winter.

Overall, the ocean is a net sink for CO₂, with the model simulating total uptake of 2.05 Pg C y⁻¹ compared to an observational estimate of 1.60 Pg C y⁻¹ (although observational products differ on this quantity; see below). Supplementary Figure S15 shows corresponding plots of surface pCO₂, a function of surface DIC, alkalinity, temperature and salinity (Rödenbeck et al., 2013). Biases in these fields illuminate those in CO₂ flux, for instance much lower pCO₂ in the North Atlantic drives stronger uptake, while higher pCO₂ in the Southern Ocean damps down outgassing in this region.

To complement Figure 16's geographical snapshot, Figure 17 shows the time-series of CO₂ uptake across the Historical period for the UKESM1 ensemble, together with the observationally-derived estimate of Khatiwala et al. (2009). The plot shows the varying rate in the rise of oceanic uptake of CO₂ across this period, with growth from the 1850s until the 1930s, followed by stalling growth until the 1950s, and finally strong continuous growth to the present-day. With some variability, particularly in the early decades, the ensemble tracks the observationally-estimated uptake, reproducing the same pace and features, but with the ensemble estimating a slightly lower flux than estimated (88.5%; integrated 1850-2013). The plot also shows UKESM1's piControl simulation to illustrate the magnitude and period of variability with constant background atmospheric xCO₂. This shows CMIP6's Historical period beginning (and the piControl period ending) in 1850, approximately a century after the industrial revolution and significant fossil fuel CO₂ emissions began. This differs from the Khatiwala et al. (2009) product,

which estimates ocean CO₂ uptake over the more complete period of anthropogenic emissions. Note that the observational
485 estimate for the present-day here is more closely matched by the model than the preceding dataset of Rödenbeck et al. (2013),
although this is not unexpected given the large uncertainties involved in estimating this flux. The influx of CO₂ into the surface
ocean is also documented in Table 2's mean and trend statistics of surface DIC and air-sea flux, in particular how they compare
with the corresponding piControl period.

The air-sea flux is just the first stage of ocean storage of anthropogenic CO₂, and Figure 18 illustrates its fate once in the
490 ocean interior. The upper row shows estimated and simulated vertically-integrated anthropogenic CO₂ for the 1990s (the nor-
malised period for Key et al., 2004). ~~In the case of the model, this~~ As described in Section 2.3, model anthropogenic CO₂
is estimated by ~~differencing Historical DIC fields from each ensemble member with the corresponding DIC field from the same
relative timepoint from the piControl~~ subtracting the 3D fields of DIC from Historical and piControl simulations aligned in
time. In broad outline, UKESM1 reproduces most of the geographical patterns of storage, with Southern Ocean uptake dis-
495 tributed into the southern sectors of the Atlantic, Pacific and (especially) Indian basins, maximum column inventories in the
North Atlantic, and much lower storage at low latitudes, in the North Pacific and around Antarctica. However, the modelled
distributions of anthropogenic CO₂ also show some clear discrepancies with observational estimates. For instance, although
exhibiting high column inventories in the Greenland-Iceland-Norwegian Sea, the model ensemble does not simulate the corre-
sponding high observationally-estimated concentrations off Newfoundland in the west of the Atlantic. More significantly, the
500 pattern of anthropogenic CO₂ being transported southward at depth in the North Atlantic shows a strong east-west gradient
that does not correspond with that observed. To investigate this further, Figure 19 shows observational and model sections
across the Atlantic at 30°N for both anthropogenic CO₂ and CFC-11. The former is estimated from observations, while the
latter is measured directly. These show a general deficit in UKESM1 in tracer concentrations between approximately 1000 and
3000 m in depth west of the mid-Atlantic ridge. In the case of CFC-11, the model completely misses a distinctive watermass
505 with high concentrations immediately adjacent to the coast of North America at approximately 1800 m. As already noted for
the surface ocean in Section 3.1, grid resolution introduces errors into transport pathways, and UKESM1's poor representation
of Deep Western Boundary Current (DWBC) return flow may be an interior example of similar limitations, coupled potentially
to discrepancies in patterns in convection and deep mixing in the vicinity of the Labrador Sea (e.g. Handmann et al., 2018).

One major issue with the preceding estimate of anthropogenic CO₂ in the ocean is that it must be separated from the natural
510 background of DIC in the ocean. In the case of the model, this is straightforward (although there remain several ways of doing
so), but it is challenging observationally. The datasets used in this study, Key et al. (2004) and Lauvset et al. (2016) use different
methodologies (as well as different-sized underlying databases) to estimate and separate anthropogenic and natural CO₂. As
this complicates evaluation of the model's distributions, the lower row of Figure 18 shows the vertical inventory of CFC-11,
a conservative artificial tracer accumulating within the ocean similarly to anthropogenic CO₂. Relatively straightforward to
515 quantify to high precision, and without any natural background, this tracer serves as a loose proxy for anthropogenic CO₂
(Dutay et al., 2002; Doney et al., 2004). As such, it provides a second performance measure against which to compare the
interior redistribution of surface anthropogenic CO₂ uptake. Overall, UKESM1's CFC-11 distributions better match those of
the observational dataset than anthropogenic CO₂. However, the same differences also arise, particularly the east-west gradient

in Atlantic column inventory, likely for the same reasons suggested above. The model also exhibits more extensive coastal uptake of CFC-11 in the Weddell Sea.

3.5 Interior biogeochemistry

Figures 20 and 21 show intercomparisons of ocean interior DIN and DIC, with Supplementary Figures S16 to S19 showing the corresponding intercomparisons for other tracers. Per Figure 6, the plots use a thermohaline transect to illustrate the connection between young watermasses in the North Atlantic through to old watermasses in the North Pacific.

525 In the case of DIN, ~~a number of~~ the main observational features are reproduced, including the low concentrations in the Arctic and (especially) the surface oligotrophic gyres, generally lower concentrations within the NADW, a limb of elevated concentrations within the AAIW, intermediate concentrations within the Southern Ocean, and the highest concentrations in the North Pacific, particularly at midwater depths. However, despite this agreement on the main patterns of features, the model also exhibits a number of pronounced biases. In the Atlantic basin, ~~near-surface positive biases~~ small positive biases in
530 near-surface waters overlie strong negative biases in the upper 3 km, where maximum concentration differences of more than 10 mmol N m⁻³ occur, ~~while~~. In the South Atlantic, these negative biases occur in association with the northward-moving limb of AAIW (approximately 1 km depth) that supplies DIN to the North Atlantic, but which can be seen to be less pronounced in UKESM1 (and also in the salinity field of Figure 7). Meanwhile, in deeper waters the bias is reversed to strong positive, as the more sluggish AABW circulation shown in Figure 8 is ventilated less efficiently, accumulating excess DIN while accruing an
535 oxygen deficit (Supplementary Figure S17). This split of biases is generally aligned with the NADW and AABW watermasses in this basin. In the Pacific basin this pattern is broadly repeated, although with stronger positive bias in the upper 1 km, and less pronounced, but similar sign, biases at depth. More clearly than in the Atlantic basin, the model shows a shallow focused layer of maximum DIN concentration in the upper 1 km, while observations indicate a more gradual change in DIN concentration with depth. An indication of its source lies in Supplementary Figure S17, which shows the corresponding transect
540 for dissolved oxygen. Oxygen concentrations are typically highest at the surface where they are replenished by the atmosphere, and progressively lower in older watermasses as oxygen is consumed by remineralisation of sinking organic matter driven by the biological pump. Based on these fields, UKESM1 exhibits a bias towards shallower remineralisation, with less nitrogen reaching the deep ocean interior through sinking particles, and corresponding overconsumption of oxygen in shallower waters, and underconsumption at depth.

545 Supplementary Figure S16 shows the corresponding situation for silicic acid, a nutrient which is primarily consumed by diatom phytoplankton in the ocean (and by diatom phytoplankton only in UKESM1). Unlike nitrogen, which is incorporated in organic matter and widely used in cellular biochemistry, silicic acid is polymerised to make protective shells (frustules), and is returned to solution principally by physicochemical dissolution rather than active remineralisation (Kamatani, 1982). Consequently, its biological turnover is slower, and a greater proportion of biogenic silica (opal) reaches the deep ocean than
550 nitrogen. Coupled to the current mode of the thermohaline circulation, which has deep-water formation in the Atlantic and the oldest watermasses in the Pacific, this results in a deep nutrient distribution where the highest concentrations occur in the Pacific basin. UKESM1 generally reproduces the differences in the nitrogen and silicon distributions, although with a number

of biases. Principally, the AABW cell in the Atlantic is a more significant reservoir of silicic acid, while the North Pacific maxima is decreased. The silicon nutricline in the North Pacific is also deeper, although this is shallower in the South Pacific. Overall, the model shows a less skewed silicon cycle, with a greater fraction of total silicon stored in the Atlantic than observed.

As Figure 21 shows, the biological coupling between nitrogen and carbon means that the distribution of DIC in the ocean shares a number of common patterns with nitrogen (albeit against a high background concentration driven by CO₂ solubility). As a consequence, UKESM1's DIC distribution also shares a number of the same biases, including NADW negative biases and AABW positive biases. However, modelled DIC has an additional negative bias across the global domain, indicating that the ocean of UKESM1 has a lower mean DIC concentration than observed, 2289.4 mmol C m⁻³ as compared to 2337.6 mmol C m⁻³ (-2.1%; Key et al., 2004). Figure 22 shows the corresponding profiles of modelled and observed DIC, together with corresponding profiles of alkalinity, anthropogenic CO₂ and CFC-11. As shown by Figures 16 and 17, this difference in DIC concentration does not prevent the model from realistically simulating the rate of ocean exchange and uptake of anthropogenic CO₂ over the Historical period, but it alters the model ocean's carbonate chemistry system including ocean pH, potentially with consequences (see Section 4.3).

Supplementary Figure S18 shows the corresponding distributions of alkalinity. While patterns of surface alkalinity are primarily driven by the hydrological cycle (evaporation, precipitation and runoff), interior alkalinity is affected by marine biogeochemistry. In UKESM1 a simplified alkalinity cycle is represented with only the net production of calcium carbonate (CaCO₃; calcite polymorph) affecting alkalinity distributions (i.e. "hard tissues pump" only, no "soft tissues pump"; cf. Marinov and Sarmiento, 2004). As this production of CaCO₃ is ultimately tied to the production of organic material, the patterns of bias in alkalinity overlap with those already seen. However, a significant mismatch in model alkalinity is a general negative surface bias. As alkalinity balances dissolved CO₂, bicarbonate and carbonate, it regulates total DIC concentration, with a negative bias in alkalinity acting to reduce total DIC concentration. Such a negative DIC bias at the surface preconditions the interior ocean to lower DIC, consistent with Figure 21. To further illustrate this model bias, Supplementary Figure S20 shows the observed and simulated relationships between salinity and alkalinity. Each data point is a surface alkalinity versus surface salinity, and the plot shows the linear relationship between these properties (c.f. Lee et al., 2006) and the offset from the observed relationship exhibited by UKESM1. The calculated regressions intersect at a salinity of 35 PSU, although model alkalinity generally lies below that observed even above this value.

4 Discussion

580 For ESMs to deliver reliable estimates of future global change, including quantification of key feedbacks, it is important that the states of their component submodels are realistic, in particular for climate-relevant time-mean distributions and temporal trends of material (carbon) and energy (heat). Here we have examined the state of the ocean component of the UKESM1 model, a new state-of-the-art ESM, and participant in CMIP6. We have evaluated the performance of both physical and biogeochemical aspects of the ocean submodel in the context of diverse observational datasets. As well as the model’s “present-day” state, we
585 have additionally examined trends in key model properties across the Historical period (1850–2014). Kuhlbrodt et al. (subm.) presents a complementary analysis of ocean heat uptake. ~~Seperate~~Separate ensemble members have been used to understand the consistency of these temporal trends, but where comparing with observational fields, we have used model output averaged across the Historical ensemble (per Supplementary Table S1).

In terms of the physical performance of UKESM1’s ocean, its state is broadly realistic, but with a number of biases. At the
590 ocean’s surface, temperature is well-reproduced globally, but with biases including a warm Southern Ocean driven by receipt of too much shortwave radiation (Sellar et al., 2019), and a marked North Atlantic “cold spot” associated with poor Gulf Stream separation and North Atlantic Current pathway. Model upper ocean mixing also reproduces the geographical and seasonal patterns observed, with a bias towards exaggeration of extreme low and high mixing. UKESM1’s sea-ice distribution captures much of the seasonal cycle in both hemispheres, although is biased positive (and thicker) throughout the year in the north
595 (driven primarily by excessively cooling aerosol forcing), while falling short of its maximum extent in the south (in part owing to the ~~SO~~Southern Ocean warm bias). The excess in Arctic sea-ice is driven by a general cool bias in surface temperature in the northern hemisphere in UKESM1, a product of aerosol or land-use forcing (Sellar et al., 2019). In the ocean interior, compensating biases in temperature and salinity are found, related to the deficiencies in the overturning circulation mentioned in Section 3.1 (Figures 6 to 8), as well as a cumulative warming bias produced during forced ocean-only spin-up (Yool et al.,
600 2020).

Biogeochemical performance of UKESM1 largely traces to previous applications of the model (e.g. Yool et al., 2013) despite a significantly longer-duration spin-up as part of UKESM1 (Yool et al., 2020). Regarding the ocean’s nutrient cycles and biological activity, the model displays a pattern of general agreement but with biases that are sometimes large. In the surface ocean, while retaining major nutrient boundaries, the model also exhibits excessive nitrogen and silicon in the Southern
605 Ocean, excess nitrogen in the Equatorial Pacific, and depletion of silicon in the North Pacific. Upper ocean productivity in the model also follows major observed patterns, though with biases including a excessively productive Southern Ocean (both geographically and temporally), and insufficiently productive oligotrophic gyres. These biases are also mirrored in other important biological fields such as zooplankton and in the surface concentration of dimethyl sulphide. Meanwhile, in the ocean interior, biases in mesopelagic nitrogen and oxygen indicate that remineralisation of sinking biogenic material in the model occurs too
610 shallow, with compensating opposite-sense biases below. In the deep Atlantic, the model’s sluggish AABW cell accumulates more nutrients than observed, both nitrogen and silicon, while losing more oxygen.

Regarding the ocean's carbon cycle, the model represents patterns of surface carbon properties well, although with general negative biases in both DIC and alkalinity concentrations. Spatial and temporal patterns of air-sea CO₂ exchange are broadly in agreement with those estimated from observations, though the model does not well represent Southern Ocean outgassing, and
615 simulates excessively strong North Atlantic ingassing. Despite these discrepancies, the model falls within the uncertainty in the observationally-estimated temporal patterns of net ocean CO₂ uptake over the Historical period. Storage of anthropogenic CO₂ in the model ocean generally matches that estimated from observations, with high amounts in the Southern and North Atlantic oceans. However, the model exhibits a spatial discrepancy in storage in the North Atlantic, with southward transport down the western side of the basin in NADW noticeably lower than observed. Within the ocean interior, because of the role of
620 the biological pump, the spatial pattern in DIC biases tracks those of nitrogen. However, the surface bias towards lower DIC also imposes a general negative bias throughout the ocean interior, with the model ocean storing less carbon than observed in the Earth system.

On the spatial and temporal scales analysed here (i.e. global and centennial), the main fields and time-series analysed show good consistency across the UKESM1 ensemble. For higher time frequencies (e.g. decadal in the Southern Ocean; interannual for the El Niño–Southern Oscillation), or for smaller regions with significant dynamics (e.g. the Arctic), cross-
625 ensemble variability will be more important, and will be considered for detailed future studies.

4.1 Biogeochemistry biases

As already described in Sections 3.1 and 3.2, UKESM1 has a number of physical biases. Examining these biases within UKESM1 forms a component of a number of parallel studies, including on circulation and Gulf Stream ~~seperation~~separation
630 (Kuhlbrodt et al., 2018), sea-ice thickness (SIMIP Community, 2020), ocean heat uptake (Kuhlbrodt et al., *subm.*), and AMOC trends (Menary et al., 2020). Consequently, in the following, we focus on explaining the biogeochemical biases found within UKESM1.

As described above, although UKESM1 reproduces the broad patterns observed in marine biogeochemistry, it also includes a number of significant biases in properties. In the following, we consider the likely underlying causes as well as potential
635 actions to address them in future versions of UKESM1.

Vertical profiles of nitrogen, oxygen and carbon display matching patterns driven by the action of the biological pump. Nitrogen and carbon consumed by phytoplankton growth in the upper ocean are transported as organic material by this pump into the ocean interior where they are released back to dissolved inorganic forms in parallel with the consumption of oxygen. In UKESM1, the profile of this process is skewed, with remineralisation of organic matter occurring too shallow in the water
640 column, resulting in excess nitrogen and carbon in the mesopelagic, a corresponding deficit of oxygen, and reversal of these biases in deeper waters that less sinking material reaches. In MEDUSA, the organic material reaching the deep interior does so primarily as “fast-sinking” particles, coupled to a ballast model in which biominerals (opal and calcite) “protect” this organic flux. Extending the remineralisation lengthscale of these sinking particles, or affording them greater biomineral protection, are both means of addressing this bias to first order.

645 Significantly for ocean productivity, UKESM1's ocean displays strong positive biases in the surface concentration of nitrogen
nutrient in a number of ocean regions, including the Southern Ocean, the Equatorial Pacific and the Peruvian Upwelling. Such
biases can indicate oversupply of nutrients, or insufficient consumption by phytoplankton. On the first, as all three regions
experience significant upwelling of interior waters, any biases in the nitrogen supply from these watermasses will play a role.
For instance, shallow remineralisation bias noted above will contribute toward the positive biases in surface waters in these
650 regions. On the second, an additional issue lies with the availability of the micronutrient iron in these regions. Although this
is also supplied by upwelling watermasses, its availability is also dependent on deposition of iron from aeolian dust, and this
deposition is biased negative in these regions in UKESM1. While atmosphere-land aspects of the deposition flux may ultimately
be important here (e.g. location of desert source regions, patterns of wind dispersal), within the ocean model itself, parameter
changes to reduce iron stress (more iron from dust, lower iron quotients in phytoplankton) could assist here. However, by
655 relieving iron stress in this uniform way, there may be consequences elsewhere in the model.

Leaving aside the interior biases described above, carbon in UKESM1 is more generally negatively biased throughout the
ocean, with implications for the ocean's role as the largest reservoir of carbon in the Earth system. As noted previously, surface
alkalinity plays a role in interior carbon by buffering the surface carbonate system and regulating the surface DIC concentrations
that ultimately ventilate the ocean interior. Modelled surface alkalinity has a general negative bias, and a different relationship
660 with surface salinity than that observed (cf. Lee et al., 2006). In the model, aside from hydrological cycle processes, only net
calcium carbonate production (and its subsequent dissolution at depth) affects alkalinity, and this acts to decrease its upper
concentration and increase its interior concentration below the calcite compensation depth (CCD). As such, this bias could
be addressed in MEDUSA simply by decreasing calcium carbonate production (and its export) to increase the retention of
alkalinity in the surface ocean. However, while alkalinity is generally lower across the upper ocean, calcite production is
665 not uniform, with a latitudinal gradient in which most net calcification occurs in the tropics. A broader point is that calcium
carbonate production in MEDUSA is highly simplified and only concerns the fraction export to the ocean interior, whereas
other models treat it in more complex ways (e.g. Kvale et al., 2015; Butenschön et al., 2016; Buitenhuis et al., 2019) that
potentially offer more realistic solutions than simple parameter scaling.

Another clear surface bias in UKESM1, and one which is easy to discern because of the ready availability of synoptic, high
670 quality observational data, is its field of surface chlorophyll. In the Southern Ocean in particular, the seasonal spring-summer
bloom has higher chlorophyll concentrations that persist longer and extend further polewards. Even in winter, anomalously
high chlorophyll concentrations ($> 0.1 \text{ mg m}^{-3}$) extend southward to the tip of the Antarctic Peninsula. This bias is strongly
associated with a corresponding productivity bias, although at the highest latitudes (in both hemispheres) there is a degree of
decoupling. This bias is particularly significant in UKESM1 because simulated chlorophyll is used in its empirical submodels
675 of DMS and PMOA, both of which are climatically-active compounds (cf. Quinn and Bates, 2011). Noticeably, the high
concentrations of chlorophyll simulated at high latitudes also persist beyond the peak of productivity. In part these biases are
related to negative sea-ice biases that allow more light to penetrate into the high latitude ocean, but their excess extent and
persistence also suggest that the chlorophyll submodel may be too responsive under low light conditions. At lower latitudes,
where light is less limiting and nutrient stress more important, sensitivity to the chlorophyll model is less pronounced.

680 Separate from these biogeochemical biases, the model exhibits several physical biases, including a general warm bias throughout the ocean, warm and cool biases regionally, some hemisphere-specific ice biases, and issues with interior circulation. These all affect the realism of the physical regime in which MEDUSA's biogeochemistry is embedded and introduce biases independently of those arising from its deficiencies. For instance, the ~~weaker-deep-overturning-weak deep overturning~~ AABW cell north of the ACC combined with a NADW cell which is slightly too weak (Figure 8) ~~leads to a colder and~~
685 ~~fresher deep ocean. At the same time, these circulation features favour~~ reduces the ventilation rate of the abyssal Atlantic, and contributes to the build-up of nutrients and the corresponding depletion of oxygen.

Finally, the corrective measures outlined above are proposed independently without any consideration of their full impacts. For instance, decreased calcium carbonate production is proposed as a countermeasure to decrease the negative bias in surface alkalinity. However, this change will also decrease the quantity of sinking organic material "protected" by this mineral, allowing
690 it to be remineralised more rapidly, shoaling the remineralisation horizon of the biological pump, and worsening the biases in nitrogen, carbon and oxygen profiles. This interdependency of model biogeochemical processes and states – and their dependency on the ocean physical state – significantly complicates model tuning, particularly given the long timescales of ventilation and three-dimensional connectivity in the ocean. Optimisation techniques such as the Transport Matrix Method (TMM; Khatiwala et al., 2005) are increasingly being used to address this (e.g. Kriest, 2017), although the resulting solutions
695 are also found to be sensitive to the physical framework (Kriest et al., accepted).

4.2 CMIP intercomparison

Figure 23 and Supplementary Figures S21 to S26 illustrate the performance of UKESM1 alongside a series of CMIP6 models for the same suite of key surface biogeochemical properties already shown. Annual mean fields for each property for each model are shown, together with the corresponding observational field (data missing from the CMIP6 archive is denoted by a blank
700 field). Fields are also shown from UKESM1's CMIP5 predecessor, HadGEM2-ES (Totterdell, 2019), to illustrate improvement between CMIP generations, together with those from MEDUSA-2.0 (Yool et al., 2013) to demonstrate the traceability of UKESM1's MEDUSA-2.1 to prior work (note that this latter work is ocean-only rather than fully-coupled).

The CMIP6 models included in this analysis are: CESM2-FV2 (Danabasoglu et al., 2020); CNRM-ESM2-1 (Voldoire et al., 2019); CanESM5 (Swart et al., 2019); IPSL-CM6A-LR (Boucher et al., 2020); MIROC-ES2L (Hajima et al., 2020); MPI-
705 ESM1-2-LR (Mauritsen et al., 2020); MRI-ESM2-0 (Yukimoto et al., 2019); NorESM2-LM (Tjiputra et al., 2020). While the full configurations of these models are diverse, the CNRM-ESM2-1, CanESM5, and IPSL-CM6A-LR models share a common NEMO physical ocean with UKESM1, though they diverge on other components, including marine biogeochemistry.

Supplementary Figures S21 and S22 show patterns of surface nitrogen and silicon nutrients. Reassuringly, most of the models capture the main geographical features of availability, including high abundance in the Southern Ocean and the subpolar north,
710 low availability throughout the subtropics, and elevated concentrations in upwelling regions (less prominently in the case of silicon). All of the models do display biases, however, differing in over- or under-estimation of Southern Ocean concentrations (UKESM1 consistently over-estimates), and in how low subtropical concentrations are drawn down to.

The patterns in surface chlorophyll shown in Supplementary Figure S23, however, are more diverse. As already noted, UKESM1 exhibits both excess concentrations in regions such as the Southern Ocean and Equatorial Pacific, and negative biases in its oligotrophic gyre regions. Other CMIP6 models exhibit both similar and different biases. For instance, several models share UKESM1's positive biases in major productive regions (MPI, MRI), while others reverse its pattern in oligotrophic region and instead have excessive chlorophyll concentrations (MIROC, NorESM). In general, while patterns of surface nutrient are broadly shared by models, chlorophyll patterns are instead somewhat divergent.

Figure 23 shows a similarly diverse pattern for ocean productivity, with models estimating both much higher and much lower global totals. While all of the models show biases, they agree on the focusing of productivity in key biomes such as the temperate high latitudes and upwelling regimes, although the biases found are not always aligned with those in chlorophyll. Excessive productivity in the Southern Ocean is significant problem in UKESM1, although it is noticeable that models using PISCES marine biogeochemistry (IPSL, CNRM) do much better in this regard.

Supplementary Figures S24 and S25 respectively show surface DIC and alkalinity across the suite of models. As already suggested from the results of UKESM1, biases in surface alkalinity are important in setting biases in DIC, with several models showing matching positive biases in both (CanESM, MRI). Interestingly, while UKESM1's institutional precursor model, HadGEM2-ES, shares neither its ocean physics nor its marine biology (Totterdell, 2019), the models share biases, particularly in alkalinity, a field strongly governed by atmospheric freshwater interactions, and a component where the models do share submodels. This underscores the role that other Earth system components may play in shaping model marine biogeochemistry.

Similarly, there is generally strong agreement in patterns of air-sea CO₂ flux shown in Supplementary Figure S26. The models broadly reproduce the latitudinal patterns of flux observed, outgassing in the tropics and (generally) ingassing at high latitudes. The models differ in detail, with variation in the magnitude of CO₂ uptake in regions such as the North Atlantic, its release along the Equatorial Pacific, and in the magnitude and geographical extent of outgassing regions in the Southern Ocean. Interestingly, a marked bias in UKESM1, strong outgassing along the west coast of South America, is reproduced in several models (CanESM, MRI), while being absent in others (MIROC, CESM).

Figure 24 summarises the performances of this suite of models using Taylor diagrams Taylor (2001). In each case, the panels indicate spatial variability normalised to that of observations (radial axis) and model-observation correlation (circular axis), both at the global, annual mean scale used in the preceding figures. In such diagrams, proximity to the red and black circle on the x-axis indicates agreement with the observational field. Overall, UKESM1 performs comparably with other ESMs, particularly well for DIC and alkalinity, and less well for DIN. The panels also show that no one model is superior in all properties, with the "best" model differing between properties, and that the various models tend to perform similarly across properties. Chlorophyll, in particular, is a property that all of the models perform badly at, while DIN is something they all perform relatively well at.

Note that this cross-CMIP6 analysis overlooks the role played by the duration of spin-up prior to Historical simulations in the magnitude of model biases. The analysis Séférian et al. (2016) found that spin-up duration of CMIP5 models ranged widely from 200 years up to almost 12000 years, and that this duration could explain the magnitude of biases. Essentially, the longer that a model is spun-up, the greater its drift from the observationally-derived initial conditions that also typically serve

as performance targets (as they do here). In the specific case of UKESM1, its ocean component was spun-up for approximately 5300 years to equilibrate its net air-sea CO₂ flux below a target of 0.1 Pg C y⁻¹ (Yool et al., 2020). This duration was also
750 sufficient for other physical and biogeochemical properties to approach quasi-equilibrium, and UKESM1's performance is unlikely to be significantly affected by drift.

In terms of performance between CMIP generations, UKESM1 shows improved representation across almost all properties relative to HadGEM2-ES, with the exception of surface DIN (where excess concentrations in the Equatorial Pacific impact UKESM1's global realism). UKESM1 improves on the marked biases in silicic acid and chlorophyll in particular, as well as a
755 generally better representation of the ocean's role in CO₂ exchange. Séférian et al. (2020) provides a more complete view of the improvements achieved in marine biogeochemistry modelling from CMIP5 to CMIP6, including between HadGEM2-ES and UKESM1.

Finally, the preceding figures show good traceability in the marine biogeochemistry performance of UKESM1 with previous instances of its use (e.g. Yool et al., 2013). For better and for worse, UKESM1 and MEDUSA-2 perform similarly across all of
760 the properties examined. With the exception of DIN, where UKESM1's geographical biases are similar but clearly larger than those of MEDUSA-2, UKESM1's performance in Figure 24 is marginally better (and despite a much longer spin-up period; 5300 vs. 120 years).

4.3 Future projection

As described above, when compared to observational metrics, UKESM1 performs well over a large number of diverse physical
765 and biogeochemical properties. However, the model displays a number of biases in the present-day state that have implications for its future behaviour under different climate scenarios.

UKESM1's Arctic sea-ice is biased positive in both seasonal extent and, in particular, thickness. In the absence of biases in the other direction, these aspects will enable it to persist longer under climate change, with a range of likely consequences for the Arctic environment (Thackeray and Hall, 2019).

770 Decreased productivity is a common ecosystem response under climate change, as ocean warming enhances ocean stratification, reduces nutrient resupply from mixing and depletes surface concentrations (Kwiatkowski et al., 2020). The positive nitrogen nutrient biases across UKESM1's ocean may (at least temporarily) stave off this depletion dampening the response of its marine ecosystem. In particular, the excess nutrient bias in the Arctic may result in unrealistic future responses as the Arctic continues to thaw (cf. Popova et al., 2012; Vancoppenolle et al., 2013).

775 In terms of surface DIC and alkalinity, UKESM1 performs best in the CMIP6 ensemble examined here (Figure 24). However, as already noted, UKESM1 exhibits a negative bias in surface alkalinity which drives a corresponding bias in surface DIC (and within the ocean interior more generally; Figure 21). This bias reduces the buffering capacity of the surface ocean (Egleston et al., 2010) and impacts the long-term capacity of the model ocean to act as a reservoir for carbon (cf. Archer, 2005).

780 Staying with the carbon cycle, although the simulated uptake of anthropogenic CO₂ by the ocean is comparable to that estimated at the global scale (Figure 17, its spatial pattern within the ocean interior exhibits circulation-driven biases (Figure

22). Invasion of anthropogenic CO₂ into shallow or rapidly ventilated water masses will lead to its more rapid return to the surface ocean and atmosphere, potentially reducing future uptake by the ocean.

5 Conclusions

- 785 – Physical and biogeochemical properties of the ocean component of the UKESM1 model have been evaluated against observations for the Historical period.
- Examined properties indicate that the model generally reproduces the main geographical and temporal features of the ocean, but with a number of marked biases.
- Physically, model biases include a global warm bias, resolution-dependent surface biases, excessive northern sea-ice, sluggish AABW circulation and weak DWBC flow.
- 790 – Biogeochemically, model biases include nutrients skewed by remineralisation and iron availability, corresponding productivity biases, and surface chemistry causing reduced carbon storage.
- Temporally, the ocean shows a number of secular trends including aerosol-driven strengthening Atlantic MOC transport, associated sea-ice and productivity changes, and realistic carbon uptake.
- 795 – The UKESM1 ensemble shows consistent behaviour across ocean properties, it performs well in key metrics compared to CMIP6 peers, and improves on that of its CMIP5 predecessor, HadGEM2-ES.
- Though overall performance is good, UKESM1’s biases have implications for its response to climate change, including the sea-ice loss rate, future productivity changes and ocean carbon uptake.

Code and data availability. All simulations used in this work were performed using version 10.9 of the Unified Model (UM), version 5.0 of JULES, NEMO version 3.6, CICE version 5.1.2 and OASIS3-MCT version 3.0. Model output from the NEMO ocean model was handled
800 using the XML Input-Output Server (XIOS) library (Meurdesoif, 2013).

Guidance concerning the availability and use of UKESM1 is available from a dedicated website:

<http://cms.ncas.ac.uk/wiki/UM/Configurations/UKESM>

(last access: 30 October 2020)

Due to intellectual property rights restrictions, neither the source code nor documentation papers for the UM or JULES can be provided.
805 However, the Met Office UM is available for use under licence, and further information on how to apply for a licence is available here:

<https://www.metoffice.gov.uk/research/approach/modelling-systems/unified-model/>

(last access: 30 October 2020)

JULES is also available under licence, free of charge, with further information on obtaining access for research purposes here:

http://jules-lsm.github.io/access_req/JULES_access.html

810 (last access: 30 October 2020)

The simulation data used in this study are archived on the Earth System Grid Federation (ESGF) node:

<https://esgf-index1.ceda.ac.uk/projects/cmip6-ceda/>

(last access: 30 October 2020)

The model Source ID for UKESM1 is UKESM1-0-LL, and simulations are identified by the following Variant Labels: r5i1p1f2, r6i1p1f2,
815 r7i1p1f2, r4i1p1f2, r8i1p1f2, r1i1p1f2, r2i1p1f2, r3i1p1f2, and r9i1p1f2 (see Supplementary Table S1 for more details). The simulation data are also archived at the Met Office and are available for research purposes through the JASMIN platform (www.jasmin.ac.uk). For further details please contact UM_collaboration@metoffice.gov.uk referencing this paper.

The Matlab scripts used for analysis and plotting are available in this Zenodo archive:

<https://doi.org/10.5281/zenodo.4155210>

820 (last access: 30 October 2020)

Appendix A: MEDUSA-2.1

MEDUSA's dual size-structure resolves small (nanophytoplankton and microzooplankton) and large (microphytoplankton and mesozooplankton) components. Similar to its living components, MEDUSA's detrital components are also split into two size classes, with small, slow-sinking detrital particles represented explicitly as separate nitrogen and carbon tracers, and large, fast-sinking particles represented implicitly. At the seafloor, MEDUSA resolves 4 reservoirs to temporarily store organic (nitrogen and carbon) and inorganic (opal and CaCO_3) material reaching the sediment via both slow- and fast-sinking particles (iron is slaved to nitrogen in these reservoirs). Supplementary Figure S2 presents a schematic outline of MEDUSA's components and the process connections between them.

The model's nitrogen, silicon and alkalinity cycles are closed and conservative (e.g. no riverine inputs), while the ~~other~~ three-eyes-cycles of iron, carbon and oxygen are open. The ocean's iron cycle includes additions from aeolian and benthic sources, and is depleted by scavenging based on local iron availability (and an assumed fixed binding ligand concentration). The ocean's carbon cycle exchanges CO_2 with the atmosphere based on local carbonate chemistry, atmospheric $x\text{CO}_2$ and ambient winds. The ocean's oxygen cycle exchanges with the atmosphere (which has an assumed fixed oxygen concentration), and dissolved oxygen is additionally created by primary production and depleted by remineralisation throughout the ocean. The various elemental cycles include both fixed and variable stoichiometry. Iron is slaved to nitrogen throughout, while nitrogen and carbon have fixed (but different) ratios in phytoplankton and zooplankton, and variable ratios in detritus. Diatom silicon has a variable ratio with nitrogen, dependent on nutrient availability and growth rate. Calcium carbonate is produced at a geographically-variable rate relative to organic carbon according to the ambient calcite saturation state, and consumes both dissolved inorganic carbon (DIC) and alkalinity in a ratio of 1:2 respectively. Oxygen production and consumption reflects the C:N ratio of organic matter produced and consumed.

Yool et al. (2013) extensively describes the structure, differential equations, functional forms and parameterisation of the MEDUSA-2.0 model in an earlier, ocean-only configuration. As part of the development cycle of UKESM1, a number of changes were made to the model, and the resulting version used here is denoted as MEDUSA-2.1 for clarity. These specific developments are listed below.

- 845 – The carbonate chemistry submodel used in MEDUSA-2, Blackford et al. (2007) (also Artoli et al., 2012) has been replaced by the MOCSY-2.0 scheme of Orr and Epitalon (2015). See Appendix B for more details.
- Since UKESM1 represents atmospheric chemistry, including elements of the sulphur cycle, MEDUSA now includes several empirical submodels of surface dimethyl sulphide (DMS) concentration to permit this Earth system feedback. See Appendix C for more details.
- 850 – In addition to DMS, the atmospheric chemistry submodel of UKESM1 includes the emission of primary marine organic aerosol (PMOA). This utilises MEDUSA's simulated surface chlorophyll coupled to the PMOA parameterisation of Gantt et al. (2011) and Gantt et al. (2012).
- During development and testing, a small number of changes have been made to MEDUSA parameter values.

- MEDUSA’s underlying model code has been extensively reorganised into small subroutines with discrete functionality, to facilitate better code management and to adopt newer Fortran conventions. Its code has also been reorganised to reflect changes within the host NEMO code, for example around model restarting.
- Throughout MEDUSA, processes involving the model’s representation of vertical space, including the explicit sinking of slow detritus and the time-stepping of material fluxes into and out of the benthic reservoirs, have been revised to reflect the adoption of variable volume (VVL) by the host NEMO model.
- Diagnostic output in MEDUSA has been upgraded to utilise the XML Input-Output Server (XIOS) adopted by NEMO. Available output from MEDUSA has been extended to include additional diagnostics, including those requested by CMIP6.

Appendix B: MOCSY-2

As indicated above, MEDUSA-2.1 replaces an existing carbonate chemistry submodel with that of MOCSY-2.0 Orr and Epitalon (2015). This includes an improved iterative solver, applicability over a wider range of ambient conditions, as well as revised parameterisations that avoid several approximations in the earlier scheme. MOCSY-2.0 is primarily used to calculate surface ocean carbonate chemistry and air-sea CO₂ exchange, but it is additionally used on a periodic basis (monthly) to calculate ocean interior carbon chemistry. In MEDUSA-2.1, the latter is used to determine the dissolution depth of sinking biogenic calcite (via its normalised saturation state, $\Omega_{calcite}$). MEDUSA-2.1 principally passes bulk ocean temperature, salinity and concentrations of dissolved inorganic carbon and alkalinity, together with atmospheric pressure, gas transfer velocity (calculated from wind speed), and xCO₂ (i.e. mole fraction; ppm) to MOCSY-2.0. Ocean concentrations of ambient silicic acid and estimated phosphate (= DIN ÷ 16) are additionally passed to MOCSY-2.0 for use in secondary coefficients, and for interior carbonate chemistry, depth and latitude are used to calculate pressure. MOCSY-2.0 has been implemented within MEDUSA in a “plug-and-play” manner to permit easy replacement with future revisions.

Per the guidance of Orr et al. (2017), the updated gas exchange scheme of Wanninkhof (2014) is used to calculate gas transfer velocity.

Appendix C: DMS concentration

As already mentioned, one addition to MEDUSA-2.1 is a representation of surface dimethyl sulphide concentration. This concentration is passed to the atmospheric chemistry component, UKCA, where it is used in UKESM1’s sulphur cycle.

MEDUSA-2.1 includes four empirical calculations for surface DMS: Anderson et al. (2001), Simo and Dachs (2002), Aranami and Tsunogai (2004) and Halloran et al. (2010). After evaluation (Sellar et al., 2019), the formulation of Anderson et al. (2001) selected for use in UKESM1 simulations. This calculates DMS from three fields provided by MEDUSA-2: surface chlorophyll, C (mg chl m⁻³), surface daily average shortwave radiation, J (W m⁻²), and surface nutrient limitation, Q (-). Surface chlorophyll, C , is the sum of contributions of the two phytoplankton types, while J is provided by UKESM1’s

885 atmospheric component. The Q term is a conventional hyperbolic function of nitrogen nutrient concentration and uptake half-saturation concentration, here using the lower half-saturation concentration of the non-diatom phytoplankton, $k_{N, Pn}$ (which has the same numerical value as that originally used in Anderson et al., 2001). Anderson et al. (2001) used these terms in a “broken stick” regression:

if $\log_{10}(C \cdot J \cdot Q) \leq s$
890 DMS = a
else
 DMS = $b \cdot [\log_{10}(C \cdot J \cdot Q) - s] + a$

Parameters a , b and s were originally fitted using an observational dataset (Kettle et al., 1999), and were tuned during the development of UKESM1 to balance the top-of-atmosphere (TOA) radiation (Sellar et al., 2019). Parameter a was lowered to 895 1.0 (from 2.29), in line with Anderson et al. (2001)’s own assessment of likely high-biased observations; parameter b was left unchanged (at 8.24); while parameter s was linearly extended to 1.56 (from 1.72) to align with the reduced a .

Appendix D: Observational data sources

The following weblinks are to sources of the observational data used in the evaluation of UKESM1.

- World Ocean Atlas 2013: temperature, salinity, nutrients, oxygen
900 <https://www.nodc.noaa.gov/OC5/woa13/>
- Hadley Centre Sea Ice and Sea Surface Temperature (HadISST.2.2): SST, sea-ice
 <https://www.metoffice.gov.uk/hadobs/hadisst2/>
- National Sea Ice Data Centre: sea-ice thickness and index
 <https://nsidc.org/data/G10006/versions/1>
905 <https://nsidc.org/data/G02135/versions/3>
- Estimating the Circulation and Climate of the Ocean (ECCO) V4r4: ocean circulation
 <https://ecco-group.org/products-ECCO-V4r4.htm>
- RAPID-MOCHA array: AMOC strength
 <https://www.rapid.ac.uk/data.php>
- 910 – Oregon State University Ocean Productivity group: chlorophyll and productivity
 <http://orca.science.oregonstate.edu/1080.by.2160.monthly.hdf.chl.seawifs.php>
 <http://orca.science.oregonstate.edu/1080.by.2160.monthly.hdf.vgpm.m.chl.m.sst.php>
 <http://orca.science.oregonstate.edu/1080.by.2160.monthly.hdf.eppley.s.chl.a.sst.php>
 <http://orca.science.oregonstate.edu/1080.by.2160.monthly.hdf.cbpm2.s.php>

- 915 – Rödenbeck et al. (2013): ocean pCO₂ and air-sea CO₂ flux
<https://www.bgc-jena.mpg.de/CarboScope/?ID=oc>
- Lana et al. (2011): surface DMS
https://www.bodc.ac.uk/solas_integration/implementation_products/group1/dms/documents/dms-1degrex1degree.zip
- 920 – Global Ocean Data Analysis Project: carbon chemistry and CFC-11
<https://www.ncei.noaa.gov/access/ocean-carbon-data-system/oceans/glodap/>
<https://www.glodap.info/index.php/mapped-data-product/>
- Moriarty and O'Brien (2013): zooplankton biomass
<http://dx.doi.org/10.1594/PANGAEA.777398>

925 *Author contributions.* AY led the analysis and wrote the manuscript. CGJ, LdeM, TK, EEP, AJGN, JH, and EWB provided significant text, analysis and/or figures for the manuscript. JP, AAS, ATB, and ACC made significant contributions to the development and implementation of marine biogeochemistry and physics within the ocean component of UKESM1. All co-authors contributed invaluable comments during the drafting of the manuscript.

Competing interests. The authors declare that they have no conflict of interest.

930 *Acknowledgements.* The authors are extremely grateful for the input and support provided by colleagues within the UKESM core team at the Met Office and associated NERC research centres. The development of UKESM1, as well as its simulation for CMIP6, has involved an extended period of close collaboration by this team over several years, and the work presented here would not have been possible without this support.

935 The authors are additionally grateful to S. Khatiwala (air-sea CO₂ flux time-series estimates), S. Lauvset (GLODAPv2), R. Lumpkin (overturning streamfunction), and T. Tyrrell (carbonate chemistry) for their assistance with data products and analysis.

The work was facilitated through the use of the Monsoon2 system, a collaborative facility supplied under the Joint Weather and Climate Research Programme (JWCRP), a strategic partnership between the Met Office and the Natural Environment Research Council.

The authors are also grateful to two anonymous referees for their comments and suggestions on an earlier draft of this manuscript. These identified a number of omissions in content and analysis, as well as lapses in clarity, and greatly assisted in improving its revised form.

940 AY, JP, CGJ, LdM, EEP, and TK were supported by the National Environmental Research Council (NERC) National Capability Science Multi-Centre (NCSMC) funding for the UK Earth System Modelling project: CGJ and TK through grant NE/N017978/1; AY, JP, LdM and EEP through grant NE/N018036/1. AS was supported by the Met Office Hadley Centre Climate Programme funded by BEIS and Defra. ATB, AGN, ACC and JH were supported by NERC project CLASS (NE/R015953/1). CGJ, TK, AY, JP and EEP additionally acknowledge the EU Horizon 2020 CRESCENDO project, grant number 641816. AY further acknowledges support from the ARISE (NE/P006000/1) and
945 Arctic PRIZE (NE/P006078/1) projects as part of the Changing Arctic Ocean program, funded by NERC.

~~The authors are grateful for the use of the Monsoon2 system, a collaborative facility supplied under the Joint Weather and Climate Research Programme (JWCRP), a strategic partnership between the Met Office and the Natural Environment Research Council.~~

950 ~~The authors are additionally grateful to S. Khatiwala (air-sea CO₂ flux time-series estimates), S. Lauvset (GLODAPv2), R. Lumpkin (overturning streamfunction)~~ “rainbow” palette used in Figures within this manuscript has been produced using the “Turbo” colourmap developed by Mikhailov (2019). This colourmap is perceptually neutral (i.e. continuous variation without sharp discontinuities), and T. Tyrrell (carbonate chemistry) for their assistance with data products and analysis ~~its design has included a formal evaluation of its colour blindness accessibility.~~

References

- Anderson, T.R., Spall, S.A., Yool, A., Cipollini, P., Challenor, P.G., and Fasham, M.J.R.: Global fields of sea surface dimethylsulphide
955 predicted from chlorophyll, nutrients and light, *J. Mar. Syst.*, 30, 1–20, 2001.
- Andrews, T., Andrews, M.B., Bodas-Salcedo, A., Jones, G.S., Kuhlbrodt, T., Manners, J., Menary, M.B., Ridley, J., Ringer, M.A., Sellar, A.A., Senior, C.A., and Tang, Y.: Forcings, feedbacks, and climate sensitivity in HadGEM3-GC3.1 and UKESM1, *J. Adv. Mod. Earth Syst.*, 11, doi:10.1029/2019MS001866, 2019.
- Aranami, K. and Tsunogai, S.: Seasonal and regional comparison of oceanic and atmospheric dimethylsulfide in the northern North Pacific:
960 Dilution effects on its concentration during winter, *J. Geophys. Res.*, 109, doi:10.1029/2003JD004288, 2004.
- Archer, D.: Fate of fossil fuel CO₂ in geologic time, *J. Geophys. Res.*, 110, C09S05, doi:10.1029/2004JC002625, 2005.
- Archibald, A.T., O'Connor, F.M., Abraham, N.L., Archer-Nicholls, S., Chipperfield, M.P., Dalvi, M., Folberth, G.A., Dennison, F., Dhomse, S.S., Griffiths, P.T., Hardacre, C., Hewitt, A.J., Hill, R.S., Johnson, C.E., Keeble, J., Köhler, M.O., Morgenstern, O., Mulcahy, J.P., Ordóñez, C., Pope, R. J., Rumbold, S.T., Russo, M.R., Savage, N.H., Sellar, A., Stringer, M., Turnock, S.T., Wild, O., and Zeng, G.: Description and evaluation of the UKCA stratospheretroosphere chemistry scheme (StratTrop vn 1.0) implemented in UKESM1, *Geosci. Model Dev.*, 13, 1223–1266, doi:10.5194/gmd-13-1223-2020, 2020.
- Artoili, Y., Blackford, J. C., Butenschon, M., Holt, J. T., Wakelin, S. L., Thomas, H., Borges, A. V., and Allen, J. I.: The carbonate system of the North Sea: Sensitivity and model validation, *J. Mar. Syst.* 102, 1–13, 2012.
- Behrenfeld, M. J. and Falkowski, P. G.: Photosynthetic rates derived from satellite-based chlorophyll concentration, *Limnol. Oceanogr.*, 42,
970 1–20, 1997.
- Best, M. J., Pryor, M., Clark, D.B., Rooney, G.G., Essery, R.L.H., Ménard, C.B., Edwards, J.M., Hendry, M.A., Porson, A., Gedney, N., Mercado, L.M., Sitch, S., Blyth, E., Boucher, O., Cox, P.M., Grimmond, C.S.B., and Harding, R.J.: The Joint UK Land Environment Simulator (JULES), model description – Part 1: Energy and water fluxes, *Geosci. Model Dev.*, 4, 677–699, doi: 10.5194/gmd-4-677-2011, 2011.
- 975 Blackford, J. C. and Gilbert, F. J.: pH variability and CO₂ induced acidification in the North Sea, *J. Mar. Syst.*, 64, 229–241, doi: 10.1016/j.jmarsys.2006.03.016, 2007.
- [Bock, J., Michou, M., Nabat, P., Abe, M., Mulcahy, J. P., Olivieri, D. J. L., Schwinger, J., Suntharalingam, P., Tjiputra, J., van Hulst, M., Watanabe, M., Yool, A., and Séférian, R.: Evaluation of ocean dimethylsulfide concentration and emission in CMIP6 models, *Biogeosciences Discuss.* \[preprint\], doi: 10.5194/bg-2020-463, in review, 2021.](#)
- 980 Boucher, O., Servonnat, J., Albright, A. L., Aumont, O., Balkanski, Y., Bastrikov, V., Bekki, S., Bonnet, R., Bony, S., Bopp, L., Braconnot, P., Brockmann, P., Cadule, P., Caubel, A., Cheruy, F., Codron, F., Cozic, A., Cugnet, D., D'Andrea, F., Davini, P., de Lavergne, C., Denvil, S., Deshayes, J., Devillers, M., Ducharne, A., Dufresne, J.-L., Dupont, E., Éthé, C., Fairhead, L., Falletti, L., Flavoni, S., Foujols, M.-A., Gardoll, S., Gastineau, G., Ghattas, J., Grandpeix, J.-Y., Guenet, B., Guez, L.E., Guilyardi, E., Guimberteau, M., Hauglustaine, D., Hourdin, F., Idelkadi, A., Joussaume, S., Kageyama, M., Khodri, M., Krinner, G., Lebas, N., Levvasseur, G., Lévy, C., Li, L., Lott,
985 F., Lurton, T., Luyssaert, S., Madec, G., Madeleine, J.-B., Maignan, F., Marchand, M., Marti, O., Mellul, L., Meurdesoif, Y., Mignot, J., Musat, I., Ottlé, C., Peylin, P., Planton, Y., Polcher, J., Rio, C., Rochetin, N., Rousset, C., Sepulchre, P., Sima, A., Swingedouw, D., Thiéblemont, R., Traore, A.K., Vancoppenolle, M., Vial, J., Vialard, J., Viovy, N., and Vuichard, N.: Presentation and evaluation of the IPSL-CM6A-LR climate model, *J. Adv. Mod. Earth Syst.*, 12, e2019MS002010, doi:, 2020.

- Brennan, M.K., Hakim, G.J., and Blanchard-Wrigglesworth, E.: Arctic sea-ice variability during the instrumental era, *Geophysical Research Letters*, 47, e2019GL086843, doi:10.1029/2019GL086843, 2020.
- 990 Buitenhuis, E.T., Le Quéré, C., Bednaršek, N., and Schiebel, R.: Large contribution of pteropods to shallow CaCO₃ export, *Global Biogeochemical Cycles*, 33, 458–468, doi: 10.1029/2018GB006110, 2019.
- Butenschön, M., Clark, J., Aldridge, J.N., Allen, J.I., Artioli, Y., Blackford, J., Bruggeman, J., Cazenave, P., Ciavatta, S., Kay, S., Lessin, G., van Leeuwen, S., van der Molen, J., de Mora, L., Polimene, L., Sailley, S., Stephens, N., and Torres, R.: ERSEM 15.06: a generic
995 model for marine biogeochemistry and the ecosystem dynamics of the lower trophic levels, *Geosci. Model Dev.*, 9, 1293–1339, doi: 10.5194/gmd-9-1293-2016, 2016.
- Carr, M.-E., Friedrichs, M. A. M., Schmeltz, M., Aita, M. N., Antoine, D., Arrigo, K. R., Asanuma, I., Aumont, O., Barber, R., Behrenfeld, M., Bidigare, R., Buitenhuis, E. T., Campbell, J., Ciotti, A., Dierssen, H., Dowell, M., Dunne, J., Esaias, W., Gentili, B., Gregg, W., Groom, S., Hoepffner, N., Ishizaka, J., Kameda, T., Le Quéré, C., Lohrenz, S., Marra, J., Mélin, F., Moore, K., Morel, A., Reddy, T. E.,
1000 Ryan, J., Scardi, M., Smyth, T., Turpie, K., Tilstone, G., Waters, K., and Yamanaka, Y.: A comparison of global estimates of marine primary production from ocean color, *Deep-Sea Res. Pt. II*, 53, 741–770, 2006.
- Ciais, P., Sabine, C., Bala, G., Bopp, L., Brovkin, V., Canadell, J., Chhabra, A., DeFries, R., Galloway, J., Heimann, M., Jones, C., Le Quéré, C., Myneni, R.B., Piao, S., and Thornton, P.: Carbon and Other Biogeochemical Cycles, in: *Climate Change 2013: The Physical Science Basis, Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by: Stocker, T.F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S.K., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P.M., Cambridge University Press, Cambridge, UK and New York, NY, USA, 465–570, doi:10.1017/CBO9781107415324.015, 2013.
- 1005 Cheng, L., Trenberth, K. E., Fasullo, J., Boyer, T., Abraham, J., and Zhu, J.: Improved estimates of ocean heat content from 1950 to 2015, *Sci. Adv.*, 3, 1–11, doi:10.1126/sciadv.1601545, 2017.
- Clark, D.B., Mercado, L.M., Sitch, S., Jones, C.D., Gedney, N., Best, M.J., Pryor, M., Rooney, G.G., Essery, R.L.H., Blyth, E., Boucher, O.,
1010 Harding, R.J., Huntingford, C., and Cox, P.M.: The Joint UK Land Environment Simulator (JULES), model description – Part 2: Carbon fluxes and vegetation dynamics, *Geosci. Model Dev.*, 4, 701–722, doi:10.5194/gmd-4-701-2011, 2011.
- Collins, W. J., Bellouin, N., Doutriaux-Boucher, M., Gedney, N., Halloran, P., Hinton, T., Hughes, J., Jones, C. D., Joshi, M., Liddicoat, S., Martin, G., O’Connor, F., Rae, J., Senior, C., Sitch, S., Totterdell, I., Wiltshire, A., and Woodward, S.: Development and evaluation of an Earth-system model – HadGEM2, *Geosci. Model Dev. Discuss.*, 4, 997–1062, doi:10.5194/gmdd-4-997-2011, 2011.
- 1015 Cox, P.M.: Description of the “TRIFFID” Dynamic Global Vegetation Model, Hadley Centre Technical Note 24, Met Office, Bracknell, UK, 2001.
- Craig A., Valcke S., Coquart L., 2017: Development and performance of a new version of the OASIS coupler, OASIS3-MCT_3.0, *Geoscientific Model Development*, 10, pp. 3297-3308, doi:10.5194/gmd-10-3297-2017.
- [de Mora, L., Yool, A., Palmieri, J., Sellar, A., Kuhlbrodt, T., Popova, E., Jones, C., and Allen, J. I.: BGC-val: a model- and grid-independent
1020 Python toolkit to evaluate marine biogeochemical models, *Geosci. Model Dev.*, 11, 4215-4240, doi:10.5194/gmd-11-4215-2018, 2018.](#)
- Danabasoglu, G., Lamarque, J.-F., Bacmeister, J., Bailey, D.A., DuVivier, A.K., Edwards, J., Emmons, L.K., Fasullo, J., Garcia, R., Gettelman, A., Hannay, C., Holland, M.M., Large, W.G., Lauritzen, P.H., Lawrence, D.M., Lenaerts, J.T.M., Lindsay, K., Lipscomb, W.H., Mills, M.J., Neale, R., Oleson, K.W., Otto-Bliesner, B., Phillips, A.S., Sacks, W., Tilmes, S., van Kampenhout, L., Vertenstein, M., Bertini, A., Dennis, J., Deser, C., Fischer, C., Fox-Kemper, B., Kay, J.E., Kinnison, D., Kushner, P.J., Larson, V.E., Long, M.C., Mickelson, S.,
1025 Moore, J.K., Nienhouse, E., Polvani, L., Rasch, P.J., and Strand, W.G.: The Community Earth System Model Version 2 (CESM2), *J. Adv. Mod. Earth Syst.*, 12, e2019MS001916, doi:10.1029/2019MS001916, 2020.

- Doney, S.C., Lindsay, K., Caldeira, K., Campin, J.-M., Drange, H., Dutay, J.-C., Follows, M., Gao, Y., Gnanadesikan, A., Gruber, N., Ishida, A., Joos, F., Madec, G., Maier-Reimer, E., Marshall, J.C., Matear, R.J., Monfray, P., Mouchet, A., Najjar, R., Orr, J.C., Plattner, G.-K., Sarmiento, J., Schlitzer, R., Slater, R., Totterdell, I.J., Weirig, M.-F., Yamanaka, Y., and Yool, A.: Evaluating global ocean carbon models: The importance of realistic physics, *Global Biogeochem. Cycles*, 18, GB3017, doi:10.1029/2003GB002150, 2004.
- 1030 Donohue, K.A., Tracey, K.L., Watts, D.R., Chidichimo, M.P., and Chereskin, T.K.: Mean Antarctic Circumpolar Current transport measured in Drake Passage, *Geophys. Res. Lett.*, 43, 11,760-11,767, doi:10.1002/2016GL070319, 2016.
- Dutay, J.-C., Bullister, J.L., Doney, S.C., Orr, J.C., Najjar, R., Caldeira, K., Campin, J.-M., Drange, H., Follows, M., Gao, Y., Gruber, N., Hecht, M.W., Ishida, A., Joos, F., Lindsay, K., Madec, G., Maier-Reimer, E., Marshall, J.C., Matear, R.J., Monfray, P., Mouchet, A., Plattner, G.-K., Sarmiento, J., Schlitzer, R., Slater, R., Totterdell, I.J., Weirig, M.-F., Yamanaka, Y., and Yool, A.: Evaluation of ocean model ventilation with CFC-11: comparison of 13 global ocean models, *Ocean Modelling*, 4, 89-120, doi: 10.1016/S1463-5003(01)00013-0, 2002.
- 1035 Egleston, E.S., Sabine, C.L., and Morel, F.M.M.: Revelle revisited: Buffer factors that quantify the response of ocean chemistry to changes in DIC and alkalinity, *Global Biogeochem. Cycles*, 24, GB1002, doi:10.1029/2008GB003407, 2010.
- 1040 Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E.: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, *Geosci. Model Dev.*, 9, 1937-1958, doi:10.5194/gmd-9-1937-2016, 2016.
- Fetterer, F., Knowles, K., Meier, W.N., Savoie, M., and Windnagel, A.K.: Sea Ice Index, Version 3, [1980–2014], National Snow and Ice Data Center, Boulder, Colorado, USA, doi:10.7265/N5K072F8, [last access: 3 September 2020], 2017.
- 1045 Forget, G., Campin, J.-M., Heimbach, P., Hill, C. N., Ponte, R. M., and Wunsch, C.: ECCO version 4: an integrated framework for non-linear inverse modeling and global ocean state estimation, *Geosci. Model Dev.*, 8, 3071–3104, doi:10.5194/gmd-8-3071-2015, 2015.
- Fukumori, I., Wang, O., Fenty, I., Forget, G., Heimbach, P., and Ponte, R.M.: ECCO Version 4 Release 4, available at https://ecco.jpl.nasa.gov/drive/files/Version4/Release4/doc/v4r4_synopsis.pdf, 2019.
- 1050 [Gantt, B., Meskhidze, N., Facchini, M.C., Rinaldi, M., Ceburnis, D., and O'Dowd, C.D.: Wind speed dependent size-resolved parameterization for the organic mass fraction of sea spray aerosol, *Atmospheric Chemistry and Physics*, 11, 8777-8790, doi: 10.5194/acp-11-8777-2011, 2011.](#)
- [Gantt, B., Johnson, M.S., Meskhidze, N., Sciare, J., Ovadnevaite, J., Ceburnis, D., and O'Dowd, C.D.: Model evaluation of marine primary organic aerosol emission schemes, *Atmospheric Chemistry and Physics*, 12, 8553-8566, doi: 10.5194/acp-12-8553-2012, 2012.](#)
- Garcia, H.E., Locarnini, R.A., Boyer, T.P., Antonov, J.I., Baranova, O.K., Zweng, M.M., Reagan, J.R., and Johnson, D.R.: World Ocean Atlas 2013, Volume 3: Dissolved Oxygen, Apparent Oxygen Utilization, and Oxygen Saturation. S. Levitus, Ed., A. Mishonov Technical Ed.; NOAA Atlas NESDIS 75, 27 pp., 2014.
- 1055 Garcia, H.E., Locarnini, R.A., Boyer, T.P., Antonov, J.I., Baranova, O.K., Zweng, M.M., Reagan, J.R., and Johnson, D.R.: World Ocean Atlas 2013, Volume 4: Dissolved Inorganic Nutrients (phosphate, nitrate, silicate). S. Levitus, Ed., A. Mishonov Technical Ed.; NOAA Atlas NESDIS 76, 25 pp., 2014.
- 1060 Gent, P.R., and McWilliams, J.C.: Isopycnal mixing in ocean circulation models, *J. Phys. Oceanogr.*, 20, 150-155, doi: 10.1175/1520-0485(1990)020<0150:IMIOCM>2.0.CO;2, 1990.
- Gent, P.R., Willebrand, J., McDougall, T.J., and McWilliams, J.C.: Parameterizing Eddy-Induced Tracer Transports in Ocean Circulation Models, *J. Phys. Oceanogr.*, 25, 463-474, doi:10.1175/1520-0485(1995)025<0463:PEITTI>2.0.CO;2, 1995.

- Gnanadesikan, A., Slater, R. D., Gruber, N., and Sarmiento, J. L.: Oceanic vertical exchange and new production: a comparison between
1065 models and observations, *Deep-Sea Res. II*, 49, 363–401, 2002.
- Hajima, T., Watanabe, M., Yamamoto, A., Tatebe, H., Noguchi, M. A., Abe, M., Ohgaito, R., Ito, A., Yamazaki, D., Okajima, H., Ito, A.,
Takata, K., Ogochi, K., Watanabe, S., and Kawamiya, M.: Development of the MIROC-ES2L Earth system model and the evaluation of
biogeochemical processes and feedbacks, *Geosci. Model Dev.*, 13, 21972244, doi:10.5194/gmd-13-2197-2020, 2020.
- Halloran, P. R., Bell, T. G., and Totterdell, I. J.: Can we trust empirical marine DMS parameterisations within projections of future climate?,
1070 *Biogeosciences*, 7, 1645–1656, doi:10.5194/bg-7-1645-2010, 2010.
- Handmann, P., Fischer, J., Visbeck, M., Karstensen, J., Biastoch, A., Böning, C., and Patara, L.: The deep western boundary current in the
Labrador Sea from observations and a high-resolution model, *J. Geophys. Res.*, 123, 2829–2850, doi: 10.1002/2017JC013702, 2018.
- Harper, A.B., Cox, P.M., Friedlingstein, P., Wiltshire, A.J., Jones, C.D., Sitch, S., Mercado, L.M., Groenendijk, M., Reich, P.B.,
Soudzilovskaia, N.A., and van Bodegom, P.: Improved representation of plant functional types and physiology in the Joint UK Land
1075 Environment Simulator (JULES v4.2) using plant trait information, *Geosci. Model Development*, 9, 2415–2440, doi:10.5194/gmd-9-
2415-2016, 2016.
- Hirschi, J.J.-M., Barnier, B., Böning, C., Biastoch, A., Blaker, A.T., Coward, A., Danilov, S., Drijfhout, S., Getzlaff, K., Griffies, S.M., Ha-
sumi, H., Hewitt, H., Iovino, D., Kawasaki, T., Kiss, A.E., Koldunov, N., Marzocchi, A., Mecking, J.V., Moat, B., Molines, J.-M., Myers,
P.G., Penduff, T., Roberts, M., Treguier, A.-M., Sein, D.V., Sidorenko, D., Small, J., Spence, P., Thompson, L., Weijer, W., Xu, X.: The At-
1080 lantic meridional overturning circulation in high-resolution models, *J. Geophys. Res.*, 125, e2019JC015522, doi: 10.1029/2019JC015522,
2020.
- [Hu, Z., Hu, A., and Hu, Y.: Contributions of Interdecadal Pacific Oscillation and Atlantic Multidecadal Oscillation to Global Ocean Heat
Content Distribution, *Journal of Climate*, 31, 1227-1244, doi: 10.1175/JCLI-D-17-0204.1, 2018.](#)
- Hunke, E.C., Lipscomb, W.H., Turner, A.K., Jeffery, N., and Elliott, S.: CICE: the Los Alamos sea ice model documentation and software
1085 user’s manual version 5.1, (LA-CC-06-012), Los Alamos National Laboratory, N.M., 2015.
- Jackson, L.C., Peterson, K.A., Roberts, C.D., and Woods, R.A.: Recent slowing of Atlantic overturning circulation as a recovery from earlier
strengthening, *Nature Geosci.* 9, 518–522, doi:10.1038/ngeo2715, 2016.
- Jackson, L.C., Dubois, C., Forget, G., Haines, K., Harrison, M., Iovino, D., Köhl, A., Mignac, D., Masina, S., Peterson, K.A., Piecuch, C.G.,
Roberts, C.D., Robson, J., Storto, A., Toyoda, T., Valdivieso, M., Wilson, C., Wang, Y., and Zuo, H.: The mean state and variability of the
1090 North Atlantic circulation: A perspective from ocean reanalyses, *J. Geophys. Res.*, 124, ~~9141–9170~~[9141–9170](#), doi:10.1029/2019JC015210,
2019.
- [Johnson, R., Strutton, P.G., Wright, S.W., McMinn, M., and Meiners, K.M.: Three improved satellite chlorophyll algorithms for the Southern
Ocean, *J. Geophys. Res. Oceans*, 118, 3694-3703, doi:10.1002/jgrc.20270, 2013.](#)
- Jones, C. D., Hughes, J. K., Bellouin, N., Hardiman, S. C., Jones, G. S., Knight, J., Liddicoat, S., O’Connor, F. M., Andres, R. J., Bell, C.,
1095 Boo, K.-O., Bozzo, A., Butchart, N., Cadule, P., Corbin, K. D., Doutriaux-Boucher, M., Friedlingstein, P., Gornall, J., Gray, L., Halloran,
P. R., Hurtt, G., Ingram, W. J., Lamarque, J.-F., Law, R. M., Meinshausen, M., Osprey, S., Palin, E. J., Parsons Chini, L., Raddatz, T.,
Sanderson, M. G., Sellar, A. A., Schurer, A., Valdes, P., Wood, N., Woodward, S., Yoshioka, M., and Zerroukat, M.: The HadGEM2-ES
implementation of CMIP5 centennial simulations, *Geosci. Model Dev.*, 4, 543–570, doi:10.5194/gmd-4-543-2011, 2011.
- Kamatani, A.: Dissolution rates of silica from diatoms decomposing at various temperatures, *Mar. Biol.*, 68, 91–96, 1982.
- 1100 Kattge, J., Díaz, S., Lavorel, S., Prentice, I. C., Leadley, P., Bönsch, G., Garnier, E., Westoby, M., Reich, P.B., Wright, I.J., Cornelissen, J.H.,
Violle, C., Harrison, S.P., Van Bodegom, P.M., Reichstein, M., Enquist, B.J., Soudzilovskaia, N.A., Ackerly, D.D., Anand, M., Atkin,

- O., Bahn, M., Baker, T.R., Baldocchi, D., Bekker, R., Blanco, C.C., Blonder, B., Bond, W.J., Bradstock, R., Bunker, D.E., Casanoves, F., Cavender-Bares, J., Chambers, J.Q., Chapin III, F.S., Chave, J., Coomes, D., Cornwell, W.K., Craine, J.M., Dobrin, B.H., Duarte, L., Durka, W., Elser, J., Esser, G., Estiarte, M., Fagan, W.F., Fang, J., Fernández-Méndez, F., Fidelis, A., Finegan, B., Flores, O., Ford, H., Frank, D., Freschet, G.T., Fyllas, N.M., Gallagher, R.V., Green, W.A., Gutierrez, A.G., Hickler, T., Higgins, S.I., Hodgson, J.G., Jalili, A., Jansen, S., Joly, C.A., Kerkhoff, A.J., Kirkup, D., Kitajima, K., Kleyer, M., Klotz, S., Knops, J.M., Kramer, K., Kühn, I., Kurokawa, H., Laughlin, D., Lee, T.D., Leishman, M., Lens, F., Lenz, T., Lewis, S.L., Lloyd, J., Llusà, J., Louault, F., Ma, S., Mahecha, M.D., Manning, P., Massad, T., Medlyn, B.E., Messier, J., Moles, A.T., Müller, S.C., Nadrowski, K., Naeem, S., Niinemets, Ü., Nöllert, S., Nüske, A., Ogaya, R., Oleksyn, J., Onipchenko, V.G., Onoda, Y., Ordoñez, J., Overbeck, G., Ozinga, W.A., Patiño, S., Paula, S., Pausas, J.G., Peñuelas, J., Phillips, O.L., Pillar, V., Poorter, H., Poorter, L., Poschold, P., Prinzing, A., Proulx, R., Rammig, A., Reinsch, S., Reu, B., Sack, L., Salgado-Negret, B., Sardans, J., Shiodera, S., Shipley, B., Siefert, A., Sosinski, E., Soussana, J., Swaine, E., Swenson, N., Thompson, K., Thornton, P., Waldram, M., Weiher, E., White, M., White, S., Wright, S.J., Yguel, B., Zaehle, S., Zanne, A.E., and Wirth, C.: TRY – a global database of plant traits, *Global Change Biology*, 17, 2905–2935, doi:10.1111/j.1365-2486.2011.02451.x, 2011.
- 1105 Kelley, M., Schmidt, G.A., Nazarenko, L.S., Bauer, S.E., Ruedy, R., Russell, G.L., Ackerman, A.S., Aleinov, I., Bauer, M., Bleck, R., Canuto, V., Cesana, G., Cheng, Y., Clune, T.L., Cook, B.I., Cruz, C.A., Del Genio, A.D., Elsaesser, G.S., Faluvegi, G., Kiang, N.Y., Kim, D., Lacis, A.A., Leboissetier, A., LeGrande, A.N., Lo, K.K., Marshall, J., Matthews, E.E., McDermid, S., Mezzuman, K., Miller, R.L., Murray, L.T., Oinas, V., Orbe, C., Pérez, C., Pando, G., Perlwitz, J.P., Puma, P.J., Rind, D., Romanou, A., Shindell, D.T., Sun, S., Tausnev, N., Tsigaridis, K., Tselioudis, G., Weng, E., Wu, J., and Yaoet, M.-S.: GISS-E2.1: Configurations and climatology, *J. Adv. Mod. Earth Syst.*, 12, e2019MS002025, doi:10.1029/2019MS002025, 2020.
- 1110 Kerr, R.A.: A north atlantic climate pacemaker for the centuries, *Science*, 288, 1984–1985, doi:10.1126/science.288.5473.1984, 2000.
- 1120 Kettle, A.J., Andreae, M.O., Amouroux, D., Andreae, T.W., Bates, T.S., Berresheim, H., Bingemer, H. and Boniforti, R., Curran, M.A.J., DiTullio, G.R., Helas, G., Jones, G.B., Keller, M.D., Kiene, R.P., Leck, C., Lefebvre, M., Malin, G., Maspero, M., Matrai, P., McTaggart, A.R., Mihalopoulos, N., Nguyen, B.C., Novo, A., Putaud, J.P., Rapsomanikis, S., Roberts, G., Schebeske, G., Sharma, S., Simó, R., Staubes, R., Turner, S., Uher, G.: A global database of sea surface dimethylsulfide DMS. measurements and a procedure to predict sea surface DMS as a function of latitude, longitude, and month. *Global Biogeochem. Cycles* 13, 399–444, doi:10.1029/1999GB900004, 1999.
- 1125 Key, R. M., Kozyr, A., Sabine, C. L., Lee, K., Wanninkhof, R., Bullister, J. L., Feely, R. A., Millero, F. J., Mordy, C., and Peng, T.-H.: A global ocean carbon climatology: results from Global Data Analysis Project (GLODAP), *Global Biogeochem. Cy.*, 18, GB4031, doi:10.1029/2004GB002247, 2004.
- 1130 Key, R.M., Olsen, A., van Heuven, S., Lauvset, S.K., Velo, A., Lin, X., Schirnick, C., Kozyr, A., Tanhua, T., Hoppema, M., Jutterström, S., Steinfeldt, R., Jeansson, E., Ishi, M., Perez, F.F. and Suzuki, T.: Global Ocean Data Analysis Project, Version 2 (GLODAPv2), ORNL/CDIAC-162, ND-P093, Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, US Department of Energy, Oak Ridge, Tennessee, doi:10.3334/CDIAC/OTG.NDP093_GLODAPv2, 2015.
- Khatiwala, S., Visbeck, M., and Cane, M.A.: Accelerated simulation of passive tracers in ocean circulation models, *Ocean Modelling*, 9, 51–69, doi:10.1016/j.ocemod.2004.04.002, 2005.
- 1135 Khatiwala, S., Primeau, F., and Hall, T.: Reconstruction of the history of anthropogenic CO₂ concentrations in the ocean, *Nature*, 462, 346–349, doi:10.1038/nature08526, 2009.
- Kok, J.F., Ward, D.S., Mahowald, N.M., and Evan, A.T.: Global and regional importance of the direct dust-climate feedback, *Nature Communications*, 9, 241, doi:10.1038/s41467-017-02620-y, 2018.

- 1140 Kriest, I.: Calibration of a simple and a complex model of global marine biogeochemistry, *Biogeosciences*, 14, 4965-4984, doi:10.5194/bg-14-4965-2017, 2017.
- Kriest, I., Kähler, P., Koeve, W., Kvale, K., Sauerland, V., and Oeschlies, A.: One size fits all? – Calibrating an ocean biogeochemistry model for different circulations, *Biogeosciences Discuss.*, doi:10.5194/bg-2020-9, accepted, 2020.
- Kuhlbrodt, T., Jones, C.G., Sellar, A., Storkey, D., Blockley, E., Stringer, M., Hill, R., Graham, T., Ridley, J., Blaker, A., Calvert, D.,
1145 Copesey, D., Ellis, R., Hewitt, H., Hyder, P., Ineson, S., Mulcahy, J., Siahahan, A., and Walton, J.: The low-resolution version of HadGEM3 GC3.1: Development and evaluation for global climate. *Journal of Advances in Modeling Earth Systems*, 10, 2865-2888, doi:10.1029/2018MS001370, 2018.
- [Kuhlbrodt, T., Griesel, A., Montoya, M., Levermann, A., Hofmann, M., and Rahmstorf, S.: On the driving processes of the Atlantic meridional overturning circulation, *Rev. Geophys.*, 45, 132, doi: 10.1029/2004RG000166, 2007.](#)
- 1150 Kuhlbrodt, T., Voldoire, A., Killick, R.E., and Palmer, M.D.: Historical ocean heat uptake in CMIP6 Earth System models: global and regional perspectives, *J. Clim.*, *subm.*
- Kvale, K.F., Meissner, K.J., Keller, D.P., Eby, M., and Schmittner, A.: Explicit Planktic Calcifiers in the University of Victoria Earth System Climate Model, Version 2.9, *Atmosphere-Ocean*, 53, 332-350, doi: 10.1080/07055900.2015.1049112, 2015.
- Kwiatkowski, L., Torres, O., Bopp, L., Aumont, O., Chamberlain, M., Christian, J.R., Dunne, J.P., Gehlen, M., Ilyina, T., John, J. G., Lenton,
1155 A., Li, H., Lovenduski, N.S., Orr, J.C., Palmieri, J., Santana-Falcón, Y., Schwinger, J., Séférian, R., Stock, C.A., Tagliabue, A., Takano, Y., Tjiputra, J., Toyama, K., Tsujino, H., Watanabe, M., Yamamoto, A., Yool, A., and Ziehn, T.: Twenty-first century ocean warming, acidification, deoxygenation, and upper-ocean nutrient and primary production decline from CMIP6 model projections, *Biogeosciences*, 17, 3439–3470, doi:10.5194/bg-17-3439-2020, 2020.
- Lana, A., Bell, T.G., Simó, R., Vallina, S.M., Ballabrera-Poy, J., Kettle, A.J., Dachs, J., Bopp, L., Saltzman, E.S., Stefels, J., Johnson,
1160 J.E., and Liss, P.S.: An updated climatology of surface dimethylsulfide concentrations and emission fluxes in the global ocean, *Global Biogeochemical Cycles*, 25, GB1004, doi:10.1029/2010GB003850, 2011.
- Lauvset, S.K., Key, R.M., Olsen, A., van Heuven, S., Velo, A., Lin, X., Schirnick, C., Kozyr, A., Tanhua, T., Hoppema, M., Jutterström, S., Steinfeldt, R., Jeansson, E., Ishii, M., Pérez, F.F., Suzuki, T., and Watelet, S.: A new global interior ocean mapped climatology: the $1^\circ \times 1^\circ$ GLODAP version 2, *Earth Syst. Sci. Data*, 8, 325–340, doi:10.5194/essd-8-325-2016, 2016.
- 1165 Lee, K., Tong, L.T., Millero, F.J., Sabine, C.L., Dickson, A.G., Goyet, C., Park, G.-H., Wanninkhof, R., Feely, R.A., Key, R.M.: Global relationships of total alkalinity with salinity and temperature in surface waters of the world's oceans, *Geophys. Res. Lett.*, 33, L19605, doi:10.1029/2006GL027207, 2006.
- Levitus, S., Antonov, J.I., Boyer, T.P., Baranova, O.K., Garcia, H.E., Locarnini, R.A., Mishonov, A.V., Reagan, J.R., Seidov, D., Yarosh, E.S., and Zweng, M.M.: World ocean heat content and thermocline sea level change (0–2000 m), 1955–2010, *Geophys. Res. Lett.*, 39, L10603,
1170 doi:10.1029/2012GL051106, 2012.
- Li, F., Vikhliav, Y.V., Newman, P.A., Pawson, S., Perlwitz, J., Waugh, D.W., and Douglass, A.R.: Impacts of Interactive Stratospheric Chemistry on Antarctic and Southern Ocean Climate Change in the Goddard Earth Observing System, Version 5 (GEOS-5), *J. Climate*, 29, 3199–3218, doi: 10.1175/JCLI-D-15-0572.1, 2016.
- Locarnini, R.A., Mishonov, A.V., Antonov, J.I., Boyer, T.P., Garcia, H.E., Baranova, O.K., Zweng, M.M., Paver, C.R., Reagan, J.R., Johnson,
1175 D.R., Hamilton, M., and Seidov, D.: World Ocean Atlas 2013, Volume 1: Temperature. S. Levitus, Ed., A. Mishonov Technical Ed.; NOAA Atlas NESDIS 73, 40 pp., 2013.

- Madec, G., and Imbard, M.: A global ocean mesh to overcome the North Pole singularity, *Clim. Dyn.*, 12, 381–388, doi: 10.1007/BF00211684, 1996.
- 1180 Madec, G., Delecluse, P., Imbard, M., and Lévy, C.: OPA 8.1 ocean general circulation model reference manual, Note du Pôle de Modélisation de l'Institut Pierre-Simon Laplace, France, Technical Report 11, 91 pp., 1998.
- Madec, G.: NEMO reference manual, ocean dynamic component: NEMO–OPA, Note du Pôle de Modélisation, Institut Pierre-Simon Laplace, Technical Report 27, Note du Pole de modélisation de l'Institut Pierre-Simon Laplace, France, No. 27, ISSN No. 1288–1619, 2008.
- Madec, G., and the NEMO Team: NEMO ocean engine, version 3.6 stable, Note du Pôle de Modélisation de l'Institut Pierre-Simon Laplace, 1185 France, doi:10.5281/zenodo.1472492, 2016.
- Mahowald, N. M., Baker, A. R., Bergametti, G., Brooks, N., Duce, R. A., Jickells, T. D., Kubilay, N., Prospero, J. M., and Tegen, I.: Atmospheric global dust cycle and iron inputs to the ocean, *Global Biogeochem. Cy.*, 19, GB4025, doi:10.1029/2004GB002402, 2005.
- Mayewski, P.A., Meredith, M.P., Summerhayes, C.P., Turner, J., Worby, A., Barrett, P.J., Casassa, G., Bertler, N.A.N., Bracegirdle, T., Naveira Garabato, A.C., Bromwich, D., Campbell, H., Hamilton, G.S., Lyons, W.B., Maasch, K.A., Aoki, S., Xiao, C., van Ommenet T., State of 1190 the Antarctic and Southern Ocean climate system, *Rev. Geophys.*, 47, RG1003, doi:10.1029/2007RG000231, 2009.
- Mann, G.W., Carslaw, K.S., Spracklen, D.V., Ridley, D.A., Manktelow, P.T., Chipperfield, M.P., Pickering, S.J., and Johnson, C.E.: Description and evaluation of GLOMAP-mode: a modal global aerosol microphysics model for the UKCA composition-climate model, *Geosci. Model Dev.*, 3, 519–551, doi:10.5194/gmd-3-519-2010, 2010.
- Marinov I., Sarmiento J.L.: The Role of the Oceans in the Global Carbon Cycle: An Overview, In: Follows M., Oguz T. (eds) *The Ocean Carbon Cycle and Climate*, NATO Science Series (Series IV: Earth and Environmental Sciences), 40, Springer, Dordrecht, Netherlands, 1195 doi:10.1007/978-1-4020-2087-2_8, 2004.
- Marzocchi, A., Hirschi, J.J.-M., Holliday, N.P., Cunningham, S.A., Blaker, A.T., and Coward, A.C.: The North Atlantic subpolar circulation in an eddy-resolving global ocean model, *J. Mar. Syst.*, 142, 2015, 126–143, doi:10.1016/j.jmarsys.2014.10.007, 2015.
- Mauritsen, T., Bader, J., Becker, T., Behrens, J., Bittner, M., Brokopf, R., Brovkin, V., Claussen, M., Crueger, T., Esch, M., Fast, I., Fiedler, S., 1200 Fläschner, D., Gayler, V., Giorgetta, M., Goll, D.S., Haak, H., Hagemann, S., Hedemann, C., Hohenegger, C., Ilyina, T., Jahns, T., Jimenez-de-la-Cuesta, D., Jungclaus, J., Kleinen, T., Kloster, S., Kracher, D., Kinne, S., Kleberg, D., Lasslop, G., Kornbluh, L., Marotzke, J., Matei, D., Meraner, K., Mikolajewicz, U., Modali, K., Möbis, B., Müller, W.A., Nabel, J.E.M.S., Nam, C.C.W., Notz, D., Nyawira, S.-S., Paulsen, H., Peters, K., Pincus, R., Pohlmann, H., Pongratz, J., Popp, M., Jürgen, T., Sebastian, R., Redler, R.R., Reick, C.H., Rohrschneider, T., Schemann, V., Schmidt, H., Schnur, R., Schulzweida, U., Six, K.D., Stein, L., Stemmler, I., Stevens, B., von Storch, J.-S., Tian, F., 1205 Voigt, A., Vrese, P., Wieners, K.-H., Wilkenskjaeld, S., Winkler, A., and Roeckner, E.: Developments in the MPI-M Earth System Model version 1.2 (MPI-ESM1.2) and its response to increasing CO₂, *J. Adv. Mod. Earth Syst.*, 11, 998–1038, doi:10.1029/2018MS001400, 2020.
- [McCarthy, G.D., Smeed, D.A., Johns, W.E., Frajka-Williams, E., Moat, B.I., Rayner, D., Baringer, M.O., Meinen, C.S., Collins, J., and Bryden, H.L.: Measuring the Atlantic Meridional Overturning Circulation at 26°N, *Prog. Oceanogr.*, 130, 91–111, doi: 10.1016/j.pocean.2014.10.006, 2015.](#)
- Menary, M.B., Robson, J., Allan, R., Booth, B.B.B., Cassou, C., Gastineau, G., Gregory, J., Hodson, D., Jones, C., Mignot, J., Ringer, M., Sutton, R., Wilcox, L., and Zhang, R.: Aerosol-forced AMOC changes in CMIP6 historical simulations, *Geophys. Res. Lett.*, 47, e2020GL088166, doi:10.1029/2020GL088166, 2020.

- Meurdesoif Y. Xios: an efficient and highly configurable parallel output library for climate modelling, The Second Workshop on Coupling Technologies for Earth System Models, Boulder, Colorado, USA, 20–22 February 2013, CW2013, 2013.
- 1215 Meysignac, B., Boyer, T., Zhao, Z., Hakuba, M.Z., Landerer, F.W., Stammer, D., Köhl, A., Kato, S., L'Ecuyer, T., Ablain, M., Abraham, J.P., Blazquez, A., Cazenave, A., Church, J.A., Cowley, R., Cheng, L., Domingues, C.M., Giglio, D., Gouretski, V., Ishii, M., Johnson, G.C., Killoch, R.E., Legler, D., Lovell, W., Lyman, J., Palmer, M.D., Piotrowicz, S., Purkey, S.G., Roemmich, D., Roca, R., Savita, A., Schuckmann, K. von, Speich, S., Stephens, G., Wang, G., Wijffels, S.E., Zilberman, N.: Measuring Global Ocean Heat Content to Estimate the Earth Energy Imbalance, *Front. Mar. Sci.*, 6, doi:10.3389/fmars.2019.00432, 2019.
- 1220 [Mikhailov, A.: Turbo, An Improved Rainbow Colormap for Visualization, Google AI Blog, https://ai.googleblog.com/2019/08/turbo-improved-rainbow-colormap-for.html, last access: 09 February 2021, 2019.](https://ai.googleblog.com/2019/08/turbo-improved-rainbow-colormap-for.html)
- Monterey, G. and Levitus, S.: Seasonal Variability of Mixed Layer Depth for the World Ocean, 96 pp. NOAA Atlas NESDIS 14, U.S. Gov. Printing Office, Washington, D.C., 1997.
- 1225 [Moore, C.M., Mills, M.M., Arrigo, K.R., Berman-Frank, I., Bopp, L., Boyd, P.W., Galbraith, E.D., Geider, R.J., Guieu, C., Jaccard, S.L., Jickells, T.D., La Roche, J., Lenton, T.M., Mahowald, N.M., Marañón, E., Marinov, I., Moore, J.K., Nakatsuka, T., Oschlies, A., Saito, M.A., Thingstad, T.F., Tsuda, A., and Ulloa, O.: Processes and patterns of oceanic nutrient limitation, *Nature Geosci.*, 6, 701-710, doi: 10.1038/ngeo1765, 2013.](https://doi.org/10.1038/ngeo1765)
- Morgenstern, O., Braesicke, P., O'Connor, F.M., Bushell, A.C., Johnson, C.E., Osprey, S.M., and Pyle, J.A.: Evaluation of the new UKCA climate-composition model – Part 1: The stratosphere, *Geosci. Model Dev.*, 2, 43-57, doi:10.5194/gmd-2-43-2009, 2009.
- Moriarty, R. and O'Brien, T. D.: Distribution of mesozooplankton biomass in the global ocean, *Earth Syst. Sci. Data*, 5, 45-55, doi:10.5194/essd-5-45-2013, 2013.
- Mulcahy, J. P., Jones, C., Sellar, A., Johnson, B., Boutle, I. A., Jones, A., et al.: Improved aerosol processes and effective radiative forcing in HadGEM3 and UKESM1, *Journal of Advances in Modeling Earth Systems*, 10, 2786-2805, doi:10.1029/2018MS001464, 2018.
- 1235 Mulcahy, J.P., Johnson, C., Jones, C.G., Povey, A.C., Scott, C.E., Sellar, A., Turnock, S.T., Woodhouse, M.T., Abraham, N.L., Andrews, M.B., Bellouin, N., Browse, J., Carslaw, K.S., Dalvi, M., Folberth, G.A., Glover, M., Grosvenor, D.P., Hardacre, C., Hill, R., Johnson, B., Jones, A., Kipling, Z., Mann, G., Mollard, J., O'Connor, F.M., Palmieri, J., Reddington, C., Rumbold, S.T., Richardson, M., Schutgens, N.A.J., Stier, P., Stringer, M., Tang, Y., Walton, J., Woodward, S., and Yool, A.: Description and evaluation of aerosol in UKESM1 and HadGEM3-GC3.1 CMIP6 historical simulations, *Geosci. Model Dev. Discuss.*, 13, 63836423, doi:10.5194/gmd-2019-357, in review, 2020.
- 1240 O'Connor, F.M., Johnson, C.E., Morgenstern, O., Abraham, N.L., Braesicke, P., Dalvi, M., Folberth, G.A., Sanderson, M.G., Telford, P.J., Voulgarakis, A., Young, P.J., Zeng, G., Collins, W.J., and Pyle, J.A.: Evaluation of the new UKCA climate-composition model – Part 2: The Troposphere, *Geosci. Model Dev.*, 7, 41-91, doi:10.5194/gmd-7-41-2014, 2014.
- O'Neill, B.C., Tebaldi, C., van Vuuren, D.P., Eyring, V., Friedlingstein, P., Hurtt, G., Knutti, R., Kriegler, E., Lamarque, J.-F., Lowe, J., 1245 Meehl, G.A., Moss, R., Riahi, K., and Sanderson, B.M.: The Scenario Model Intercomparison Project (ScenarioMIP) for CMIP6, *Geosci. Model Dev.*, 9, 3461–3482, doi:10.5194/gmd-9-3461-2016, 2016.
- O'Reilly, J. E., Maritorena, S., Mitchell, B. G., Siegal, D. A., Carder, K. L., Garver, S. A., Kahru, M., and McClain, C.: Ocean color chlorophyll algorithms for SeaWiFS, *J. Geophys. Res.*, 103, 24937–24953, 1998.
- Oka, A.: Ocean carbon pump decomposition and its application to CMIP5 earth system model simulations, *Prog. Earth Planet. Sci.*, 7, 25, 1250 doi:10.1186/s40645-020-00338-y, 2020.

- Orr, J. C. and Epitalon, J.-M.: Improved routines to model the ocean carbonate system: mocsy 2.0, *Geosci. Model Dev.*, 8, 485–499, doi:10.5194/gmd-8-485-2015, 2015.
- 1255 Orr, J.C., Najjar, R.G., Aumont, O., Bopp, L., Bullister, J.L., Danabasoglu, G., Doney, S.C., Dunne, J.P., Dutay, J.-C., Graven, H., Griffies, S.M., John, J.G., Joos, F., Levin, I., Lindsay, K., Matear, R.J., McKinley, G. A., Mouchet, A., Oschlies, A., Romanou, A., Schlitzer, R., Tagliabue, A., Tanhua, T., and Yool, A.: Biogeochemical protocols and diagnostics for the CMIP6 Ocean Model Intercomparison Project (OMIP), *Geosci. Model Dev.*, 10, 2169–2199, doi:10.5194/gmd-10-2169-2017, 2017.
- 1260 Pacifico, F., Harrison, S. P., Jones, C. D., Arneth, A., Sitch, S., Weedon, G. P., Barkley, M. P., Palmer, P. I., Serça, D., Potosnak, M., Fu, T.-M., Goldstein, A., Bai, J., and Schurgers, G.: Evaluation of a photosynthesis-based biogenic isoprene emission scheme in JULES and simulation of isoprene emissions under present-day climate conditions, *Atmos. Chem. Phys.*, 11, 4371–4389, doi:10.5194/acp-11-4371-2011, 2011.
- Parkinson, C.L.: A 40-y record reveals gradual Antarctic sea ice increases followed by decreases at rates far exceeding the rates seen in the Arctic, *Proc. Nat. Acad. Sci.*, 116, 14414–14423, doi: 10.1073/pnas.1906556116, 2019.
- 1265 Popova E., Yool, A., Coward, A., Dupont, F., Deal, C., Elliot, S., Hunke, E., Jin, M., Steele, M., and Zhang, J.: What controls primary production in the Arctic Ocean? Results from an ecosystem model intercomparison, *J. Geophys. Res.*, 117, C00D12, doi:10.1029/2011JC007112, 2012.
- Power, S., Casey, T., Folland, C., Colman, A., and Mehta, V.: Inter-decadal modulation of the impact of ENSO on Australia. *Clim. Dyn.*, 15, 319–324, doi:10.1007/s003820050284, 1999.
- Quinn, P., and Bates, T.S.: The case against climate regulation via oceanic phytoplankton sulphur emissions, *Nature*, 480, 51, doi:10.1038/nature10580, 2011.
- 1270 Rae, J.G.L., Hewitt, H.T., Keen, A.B., Ridley, J.K., West, A.E., Harris, C.M., Hunke, E.C., and Walters, D.N.: Development of the Global Sea Ice 6.0 CICE configuration for the Met Office Global Coupled model, *Geosci. Model Dev.*, 8, 2221–2230, doi:10.5194/gmd-8-2221-2015, 2015.
- Ridley, J. K., Blockley, E. W., Keen, A. B., Rae, J. G. L., West, A. E., and Schroeder, D.: The sea ice model component of HadGEM3-GC3.1, *Geosci. Model Dev.*, 11, 713–723, doi:10.5194/gmd-11-713-2018, 2018.
- 1275 Rödenbeck, C., Keeling, R.F., Bakker, D.C.E., Metzl, N., Olsen, A., Sabine, C., and Heimann, M.: Global surface-ocean pCO₂ and sea–air CO₂ flux variability from an observation-driven ocean mixed-layer scheme, *Ocean Sci.*, 9, 193–216, doi:10.5194/os-9-193-2013, 2013.
- [Salinger, M., Renwick, J. and Mullan, A.: Interdecadal Pacific Oscillation and South Pacific climate, *Int. J. Climatol.*, 21, 1705–1721, doi: 10.1002/joc.691, 2001.](https://doi.org/10.1002/joc.691)
- 1280 Schröder, D., Feltham, D.L., Tsamados, M., Ridout, A., and Tilling, R.: New insight from CryoSat-2 sea ice thickness for sea ice modelling, *The Cryosphere*, 13, 125139, doi:10.5194/tc-13-125-2019, 2019.
- Séférian, R., Gehlen, M., Bopp, L., Resplandy, L., Orr, J.C., Marti, O., Dunne, J.P., Christian, J.R., Doney, S.C., Ilyina, T., Lindsay, K., Halloran, P.R., Heinze, C., Segsneider, J., Tjiputra, J., Aumont, O., and Romanou, A.: Inconsistent strategies to spin up models in CMIP5: implications for ocean biogeochemical model performance assessment, *Geosci. Model Dev.*, 9, 1827–1851, doi:10.5194/gmd-9-1827-2016, 2016.
- 1285 Séférian, R., Berthet, S., Yool, A., Palmiéri, J., Bopp, L., Tagliabue, A., Kwiatkowski, L., Aumont, O., Christian, J., Dunne, J., Gehlen, M., Ilyina, T., John, J.G., Li, H., Long, M., Luo, J. Y., Nakano, H., Romanou, A., Schwinger, J., Stock, C., Santana-Falcón, Y., Takano, Y., Tjiputra, J., Tsujino, H., Watanabe, M., Wu, T., Wu, F., Yamamoto, A.: Tracking improvement in simulated marine biogeochemistry between CMIP5 and CMIP6, *Current Climate Change Reports*, doi:10.1007/s40641-020-00160-0, 2020.

- Sellar, A.A., Jones, C.G., Mulcahy, J., Tang, Y., Yool, A., Wiltshire, A. O'Connor, F.M., Stringer, M., Hill, R., Palmiéri, J., Woodward, S.,
1290 de Mora, L., Kuhlbrodt, T., Rumbold, S., Kelley, D.I., Ellis, R., Johnson, C.E., Walton, J., Abraham, N.L., Andrews, M.B., Andrews,
T., Archibald, A.T., Berthou, S., Burke, E., Blockley, E., Carslaw, K., Dalvi, M., Edwards, J., Folberth, G.A., Gedney, N., Griffiths,
P.T., Harper, A.B., Hendry, M.A., Hewitt, A.J., Johnson, B., Jones, A., Jones, C.D., Keeble, J., Liddicoat, S., Morgenstern, O., Parker,
R.J., Predoi, V., Robertson, E., Siahann, A., Smith, R.S., Swaminathan, R., Woodhouse, M., Zeng, G., and Zerroukat, M.: UKESM1:
Description and evaluation of the UK Earth System Model, *Journal of Advances in Modeling Earth Systems*, in press.
- 1295 SIMIP Community: Arctic sea ice in CMIP6, *Geophys. Res. Lett.*, 47, e2019GL086749, doi: 10.1029/2019GL086749, 2020.
- Simo, R. and Dachs, J.: Global ocean emission of dimethylsulfide predicted from biogeophysical data, *Global Biogeochem. Cy.*, 16,
doi:10.1029/2001GB001829, 2002.
- Smeed, D.A., Josey, S.A., Beaulieu, C., Johns, W.E., Moat, B.I., Frajka-Williams, E., Rayner, D., Meinen, C.S., Baringer, M.O., Bryden,
H.L., and McCarthy, G.D.: The North Atlantic Ocean is in a state of reduced overturning, *Geophysical Research Letters*, 45, 1527-1533,
1300 doi:10.1002/2017GL076350, 2018.
- Steele, J.H., and Henderson, E.W.: The role of predation in plankton models, *J. Plankton Res.*, 14, 157-172, doi:10.1093/plankt/14.1.157,
1992.
- Storkey, D., Blaker, A. T., Mathiot, P., Megann, A., Aksenov, Y., Blockley, E. W., Calvert, D., Graham, T., Hewitt, H. T., Hyder, P., Kuhlbrodt,
T., Rae, J. G. L., and Sinha, B.: UK Global Ocean GO6 and GO7: a traceable hierarchy of model resolutions, *Geosci. Model Dev.*, 11,
1305 3187-3213, doi:10.5194/gmd-11-3187-2018, 2018.
- Stroeve, J. and Meier, W.N.: Gridded Observational Sea Ice Thickness Products, Version 1, NSIDC: National Snow and Ice Data Center,
Boulder, Colorado USA, 2016.
- Swart, N. C., Cole, J. N. S., Kharin, V. V., Lazare, M., Scinocca, J. F., Gillett, N. P., Anstey, J., Arora, V., Christian, J. R., Hanna, S., Jiao, Y.,
Lee, W. G., Majaess, F., Saenko, O. A., Seiler, C., Seinen, C., Shao, A., Sigmond, M., Solheim, L., von Salzen, K., Yang, D., and Winter,
1310 B.: The Canadian Earth System Model version 5 (CanESM5.0.3), *Geosci. Model Dev.*, 12, 48234873, doi:10.5194/gmd-12-4823-2019,
2019.
- Tagliabue, A., Bowie, A., Boyd, P., Buck, K.N., Johnson, K.S., and Saito, M.A.: The integral role of iron in ocean biogeochemistry, *Nature*,
543, 51–59, doi:10.1038/nature21058, 2017.
- Taylor, K.E.: Summarizing multiple aspects of model performance in a single diagram, *J. Geophys. Res.*, 106, 7183–7192,
1315 doi:10.1029/2000JD900719, 2001.
- Thackeray, C.W., Hall, A.: An emergent constraint on future Arctic sea-ice albedo feedback, *Nat. Clim. Chang.*, 9, 972–978,
doi:10.1038/s41558-019-0619-1, 2019.
- Titchner, H.A., and Rayner, N.A.: The Met Office Hadley Centre sea ice and sea surface temperature data set, version 2: 1. Sea ice concen-
trations, *J. Geophys. Res. Atmos.*, 119, 2864-2889, doi: 10.1002/2013JD020316, 2014.
- 1320 Tjiputra, J. F., Schwinger, J., Bentsen, M., Morée, A. L., Gao, S., Bethke, I., Heinze, C., Goris, N., Gupta, A., He, Y.-C., Olivié, D., Seland, Ø.,
and Schulz, M.: Ocean biogeochemistry in the Norwegian Earth System Model version 2 (NorESM2), *Geosci. Model Dev.*, 13, 23932431,
doi:10.5194/gmd-13-2393-2020, 2020.
- Totterdell, I.J.: Description and evaluation of the Diat-HadOCC model v1.0: the ocean biogeochemical component of HadGEM2-ES, *Geosci.*
Model Dev., 12, 4497–4549, doi:10.5194/gmd-12-4497-2019, 2019.
- 1325 Valcke, S.: The OASIS3 coupler: a European climate modelling community software, *Geosci. Model Dev.*, 6, 373-388, doi:10.5194/gmd-6-
373-2013, 2013.

- Vancoppenolle, M., Bopp, L., Madec, G., Dunne, J., Ilyina, T., Halloran, P. R., and Steiner, N.: Future Arctic Ocean primary productivity from CMIP5 simulations: Uncertain outcome, but consistent mechanisms, *Global Biogeochem. Cycles*, 27, 605–619, doi:10.1002/gbc.20055, 2013.
- 1330 Voldoire, A., Saint-Martin, D., Sénési, S., Decharme, B., Alias, A., Chevallier, M., Colin, J., Guérémy, J.-F., Michou, M., Moine, M.-P., Nabat, N., Roehrig, R., Salas y Mélia, D., Séférian, R., Valcke, S., Beau, I., Belamari, S., Berthet, S., Cassou, C., Cattiaux, J., Deshayes, J., Douville, H., Ethé, C., Franchistéguy, L., Geoffroy, O., Lévy, C., Madec, G., Meurdesoif, Y., Msadek, R., Ribes, A., Sanchez-Gomez, E., Terray, L., and Waldman, R.: Evaluation of CMIP6 DECK experiments with CNRM-CM6-1, *J. Adv. Mod. Earth Syst.*, 11, 2177–2213, doi:10.1029/2019MS001683, 2019.
- 1335 Walters, D., Baran, A. J., Boutle, I., Brooks, M., Earnshaw, P., Edwards, J., Furtado, K., Hill, P., Lock, A., Manners, J., Morcrette, C., Mulcahy, J., Sanchez, C., Smith, C., Stratton, R., Tennant, W., Tomassini, L., Van Weverberg, K., Vosper, S., Willett, M., Browse, J., Bushell, A., Carslaw, K., Dalvi, M., Essery, R., Gedney, N., Hardiman, S., Johnson, B., Johnson, C., Jones, A., Jones, C., Mann, G., Milton, S., Rumbold, H., Sellar, A., Ujiie, M., Whittall, M., Williams, K., and Zerroukat, M.: The Met Office Unified Model Global Atmosphere 7.0/7.1 and JULES Global Land 7.0 configurations, *Geosci. Model Dev.*, 12, 1909–1963, doi:10.5194/gmd-12-1909-2019, 1340 2019.
- Wanninkhof, R.: Relationship between wind speed and gas exchange over the ocean revisited, *Limnol. Oceanogr. Methods*, 12, doi:10.4319/lom.2014.12.351, 2014.
- Westberry, T., Behrenfeld, M. J., Siegel, D. A., and Boss, E.: Carbon-based primary productivity modeling with vertically resolved photoacclimation, *Global Biogeochem. Cy.*, 22, GB2024, doi:10.1029/2007GB003078, 2008.
- 1345 Williams, K.D., Copsey, D., Blockley, E.W., Bodas-Salcedo, A., Calvert, D., Comer, R., Davis, P., Graham, T., Hewitt, H.T., Hill, R., Hyder, P., Ineson, S., Johns, T.C., Keen, A.B., Lee, R.W., Megann, A., Milton, S.F., Rae, J.G.L., Roberts, M.J., Scaife, A.A., Schiemann, R., Storkey, D., Thorpe, L., Watterson, I.G., Walters, D.N., West, A., Wood, R.A., Woollings, T., and Xavier, P. K.: The Met Office Global Coupled model 3.0 and 3.1 (GC3.0 and GC3.1) configurations. *Journal of Advances in Modeling Earth Systems*, 10, 357–380, doi:10.1002/2017, 2017.
- 1350 Wiltshire, A.J., Burke, E.J., Chadburn, S.E., Jones, C.D., Cox, P.M., Davies-Barnard, T., Friedlingstein, P., Harper, A.B., Liddicoat, S., Sitch, S.A., and Zaehle, S.: JULES-CN: a coupled terrestrial Carbon-Nitrogen Scheme (JULES vn5.1), *Geosci. Model Dev. Discuss.*, doi:10.5194/gmd-2020-205, in review, 2020.
- Woodward, S.: Mineral dust in HadGEM2, Hadley Centre Technical Report 87, Met Office, Exeter, UK, 2011.
- Yool, A., Popova, E. E., and Anderson, T. R.: MEDUSA–2.0: an intermediate complexity biogeochemical model of the marine carbon cycle 1355 for climate change and ocean acidification studies, *Geosci. Model Dev. Discuss.*, 6, 1259–1365, doi:10.5194/gmdd-6-1259-2013, 2013.
- Yool, A., Palmiéri, J., Jones, C.G., Sellar, A.A., de Mora, L., Kuhlbrodt, T., Popova, E.E., Mulcahy, J., Wiltshire, A., Rumbold, S., Stringer, M., Hill, R., Tang, Y., Walton, J., Blaker, A., Nurser, A.G., Coward, A., Hirschi, J., Woodward, S., Kelley, D.I., Ellis, R., and Rumbold-Jones, S.: Spin-up of UK Earth System Model 1 (UKESM1) for CMIP6, *Journal of Advances in Modeling Earth Systems*, 12, e2019MS001933, doi:10.1029/2019MS001933, 2020.
- 1360 Yukimoto, S., Kawai, H., Koshiro, T., Oshima, N., Yoshida, K., Urakawa, S., Tsujino, H., Deushi, M., Tanaka, T., Hosaka, M., Yabu, S., Yoshimura, H., Shindo, E., Mizuta, R., Obata, A., Adachi, Y., and Ishii, M.: The Meteorological Research Institute Earth System Model Version 2.0, MRI-ESM2.0: Description and Basic Evaluation of the Physical Component, *J. Met. Soc. Japan. Ser. II*, 2019, Volume 97, 931–965, doi:10.2151/jmsj.2019-051, 2019.

Zhang, Y., Wallace, J.M., Battisti, D.S., Zhang, Y., Wallace, J.M., and Battisti, D.S.: ENSO-like Interdecadal Variability: 1900-93, *J. Clim.*, 10, 1004–1020, doi:10.1175/1520-0442(1997)010<1004:ELIV>2.0.CO;2, 1997.

1365

Zweng, M.M., Reagan, J.R., Antonov, J.I., Locarnini, R.A., Mishonov, A.V., Boyer, T.P., Garcia, H.E., Baranova, O.K., Johnson, D.R., Seidov, D., and Biddle, M.M.: *World Ocean Atlas 2013, Volume 2: Salinity*. S. Levitus, Ed., A. Mishonov Technical Ed.; NOAA Atlas NESDIS 74, 39 pp., 2013.

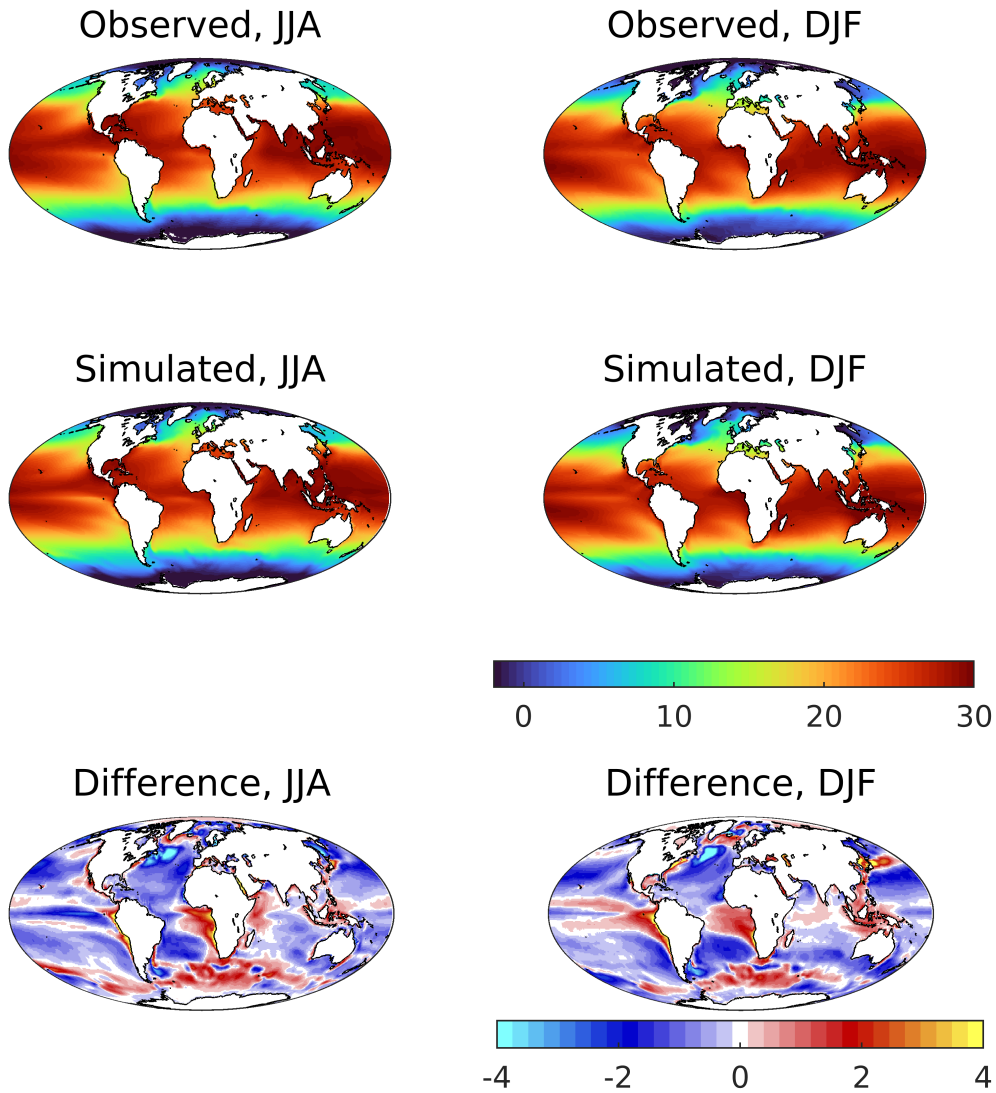


Figure 1. Observational (top; HadISST) and simulated (middle) sea surface temperature for northern (topleft; JJA) and southern (mediumright; DJF) summer. Differences (simulated - observed) for both seasons shown in the bottom row. Temperature (and difference in temperature) in °C.

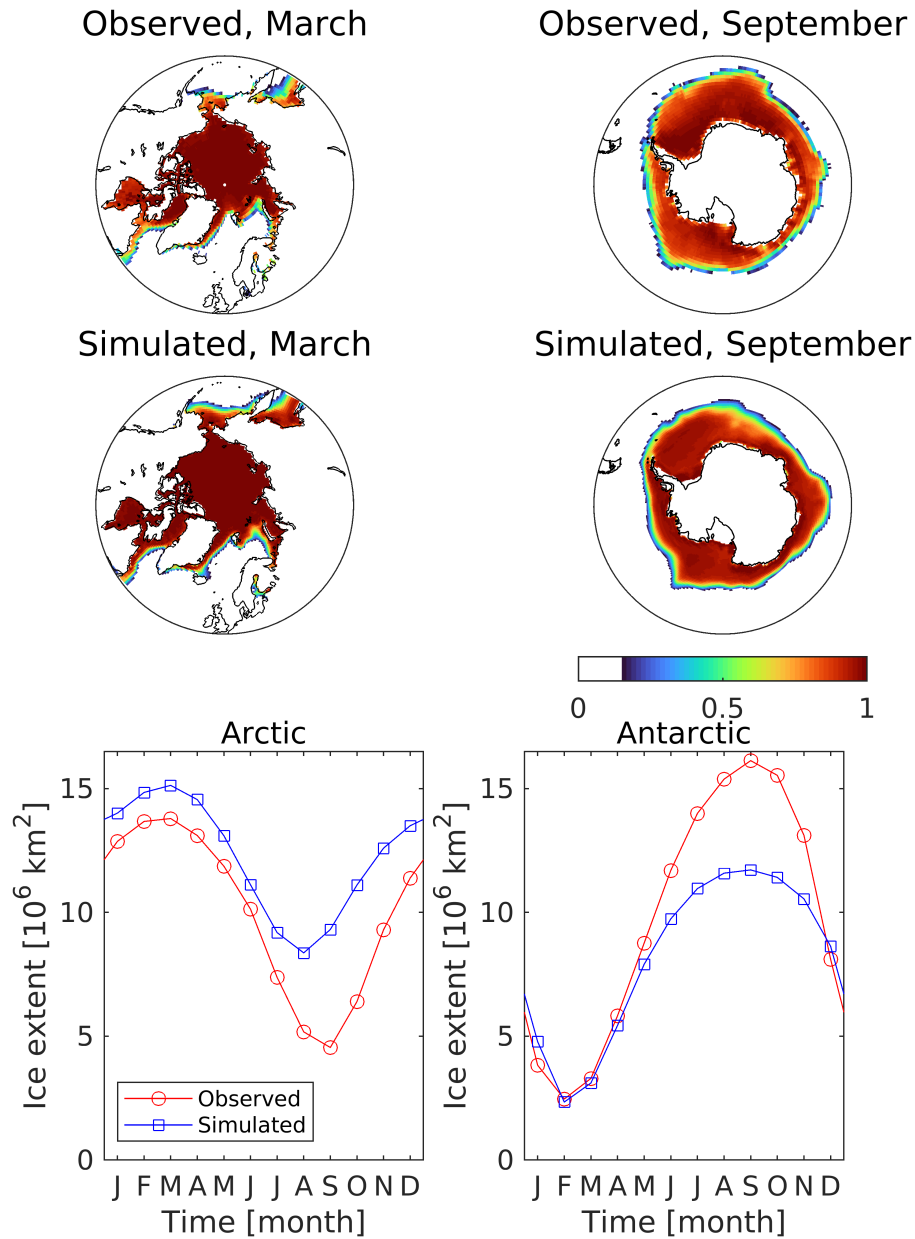


Figure 2. Observational ([lefttop](#); HadISST) and simulated ([rightmiddle](#)) maximum annual sea-ice cover for the Arctic (March; [left](#)) and Antarctic (September; [right](#)). Sea-ice cover is non-dimensional, and values less than 0.15 have been masked. The bottom row shows the seasonal sea-ice extent ($> 15\%$ cover; in 10^6 km^2) for the polar regions of each hemisphere.

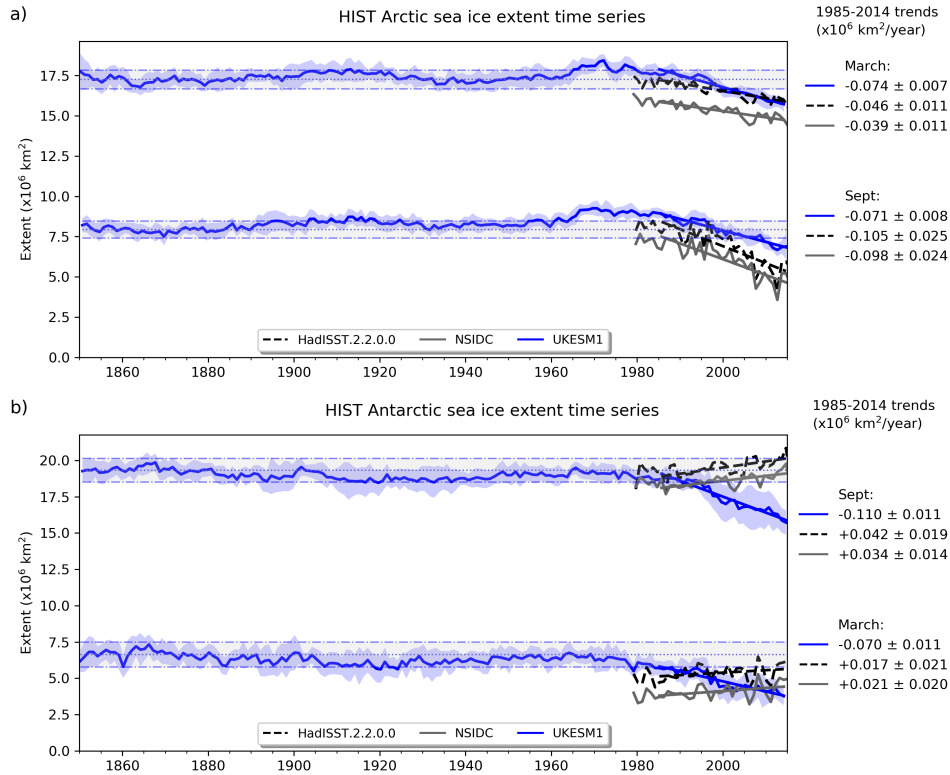


Figure 3. Observational (black, HadISST; grey, NSIDC) and simulated (blue) sea-ice extent in the Arctic (top) and Antarctic (bottom) across the Historical period (1850–2014). Panels show extent at seasonal minima and maxima, with recent (1985–2014) trends shown. All model ensemble members are shown, with the model ensemble mean is shown, with the span of their variability ± 1 standard deviation shaded in blue to show their variability. Panels show extent for September and March, which roughly correspond to the seasonal minima and maxima.

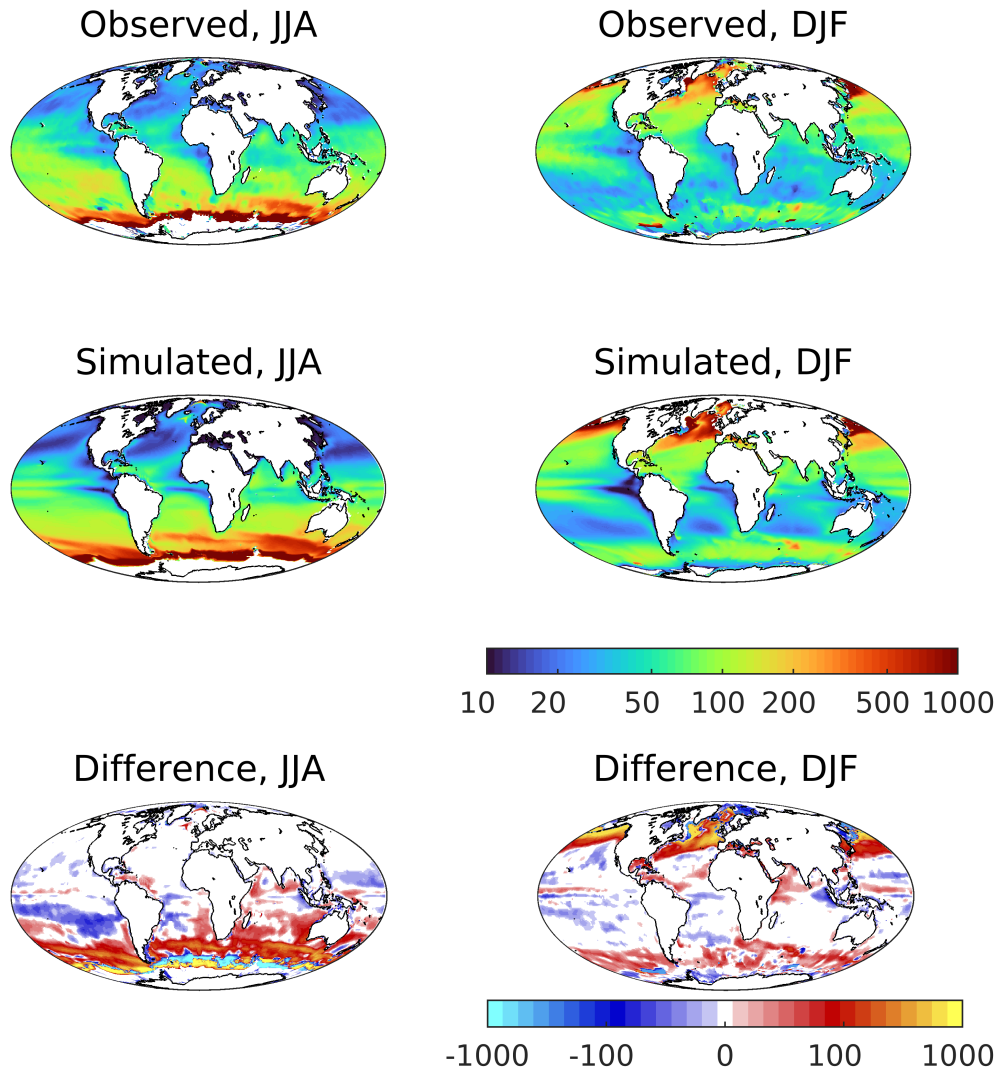


Figure 4. Observationally-derived ([left top](#); World Ocean Atlas) and simulated ([right middle](#)) mixed layer depth for northern summer ([top left](#); JJA) and southern summer ([bottom right](#); DJF). [Differences \(simulated - observed\) for both seasons shown in the bottom row.](#) Mixed layer depth derived [using a from full three-dimensional fields of potential temperature, using a temperature difference criterion \(Monterey and Levitus, 1997\).](#) In this, mixed layer depth is the depth at which potential temperature differs from that at 5 m [temperature criterion \(by 0.5°C\) and full three-dimensional fields of potential temperature \(Monterey and Levitus, 1997\).](#) White regions are those where this criterion fails (i.e. [temperature below 5 m ocean interior temperature is never cooler by than that at 5 m by the 0.5°C criterion; typically sea-ice covered regions](#)). Mixed layer depth in m, and shown on a logarithmic scale.

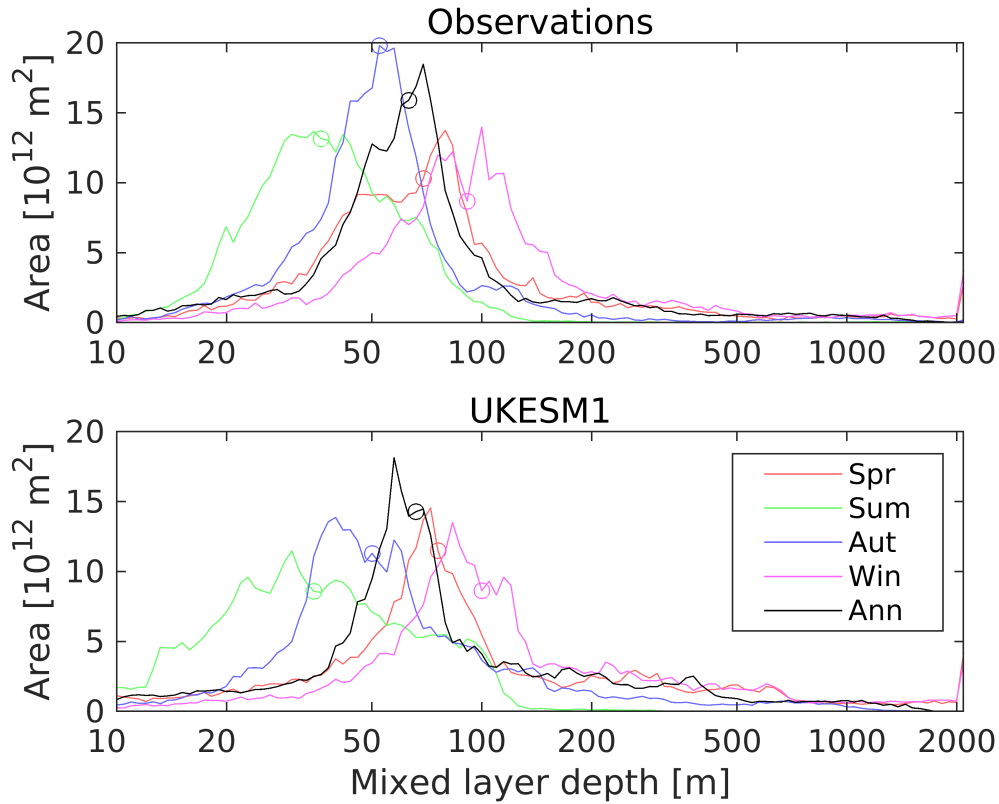


Figure 5. Frequency (in areal terms) of observation-derived (top; WOA) and simulated (bottom) seasonal mixed layer depths. Mixed layer depth derived here using a 5 m temperature criterion ($0.5\text{ }^{\circ}\text{C}$) and full three-dimensional fields of potential temperature (Monterey and Levitus, 1997). Hemispheres have been temporally-aligned so that seasons co-occur (i.e. summer is JJA for the north and DJF for the south). Circles indicate the medians for each seasonal period (i.e. the 50% of ocean area mark).

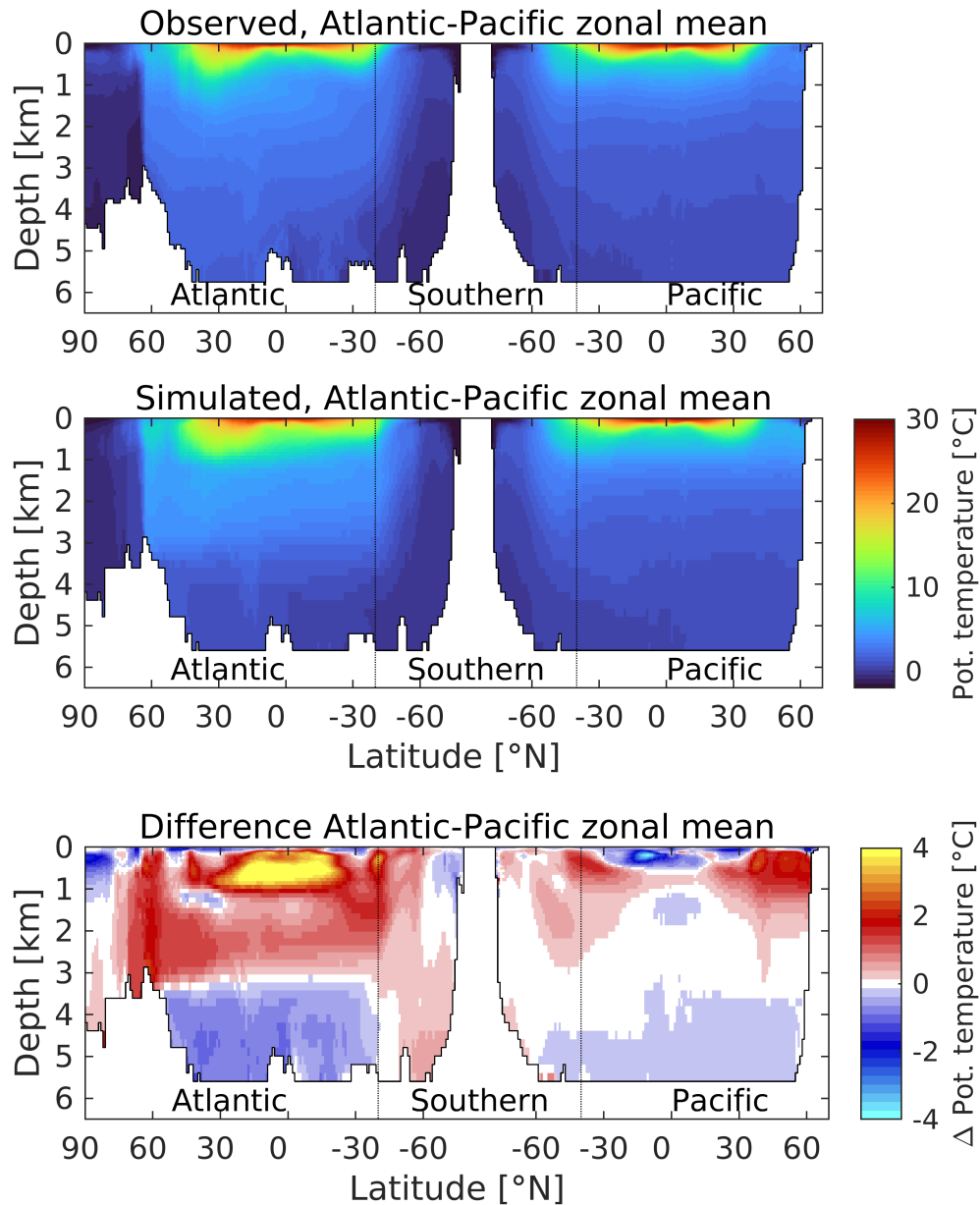


Figure 6. A “thermohaline circulation” section of observed (top) and modelled (bottom middle) zonal average potential temperature. [Difference \(simulated - observed\) is shown in the bottom panel.](#) The section tracks southwards “down” the Atlantic basin from the Arctic to the Southern Ocean, before tracking northwards “up” the Pacific basin from the Southern Ocean to the Bering Straits. The aim is to capture the stereotypical transport of deep water from its formation as a “young” water mass in the high North Atlantic through to end as an “old” water mass in the North Pacific. Dotted lines mark the “boundaries” of the Southern Ocean at 40 $^{\circ}$ S in each basin. Potential temperature in $^{\circ}$ C.

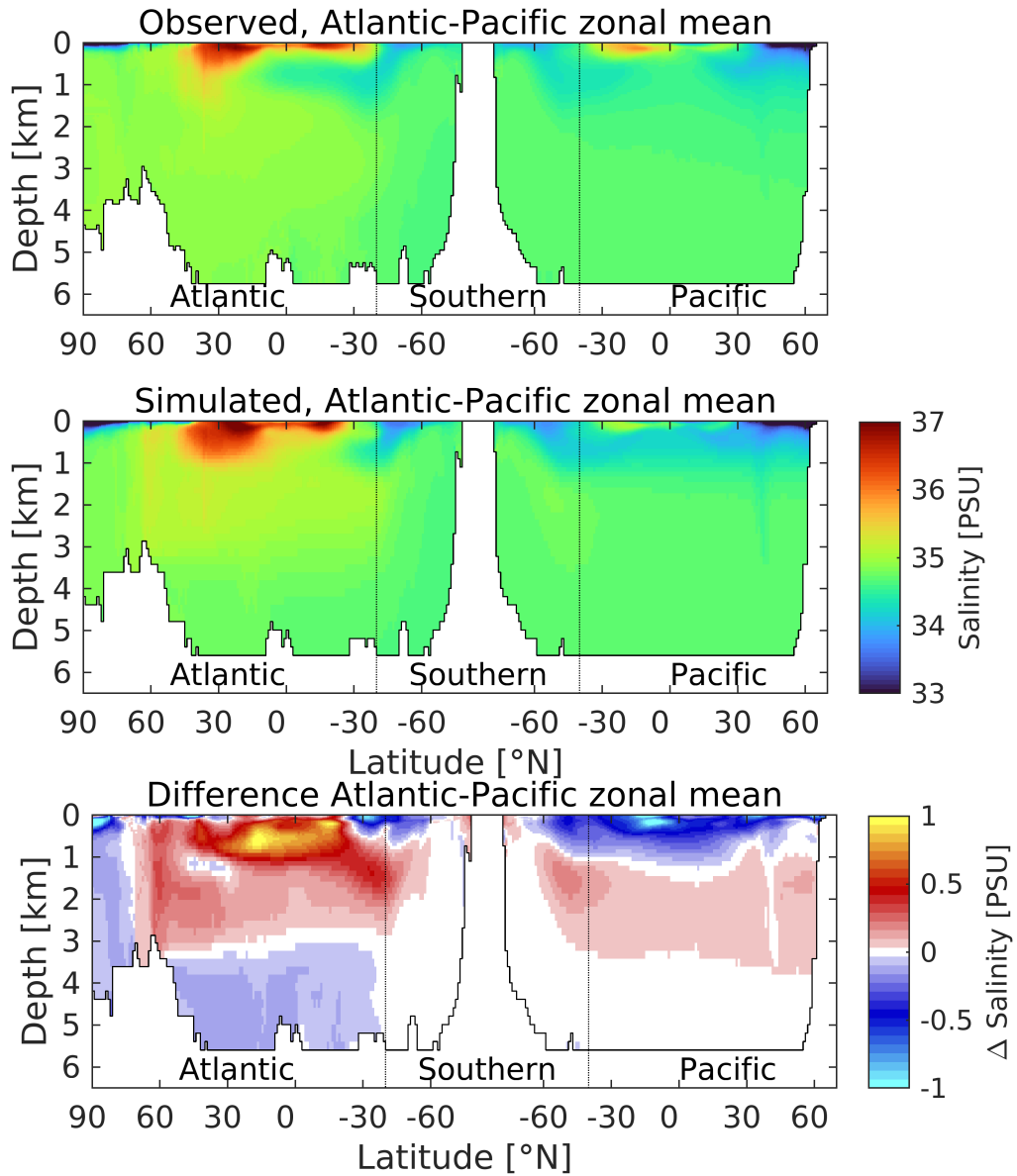


Figure 7. A “thermohaline circulation” section of observed (top) and modelled (bottom middle) zonal average salinity. [Figure 6 explains the format of this section](#) Difference (simulated - observed) is shown in the bottom panel. Salinity in practical salinity units (PSU). [Figure 6 explains the format of this section.](#)

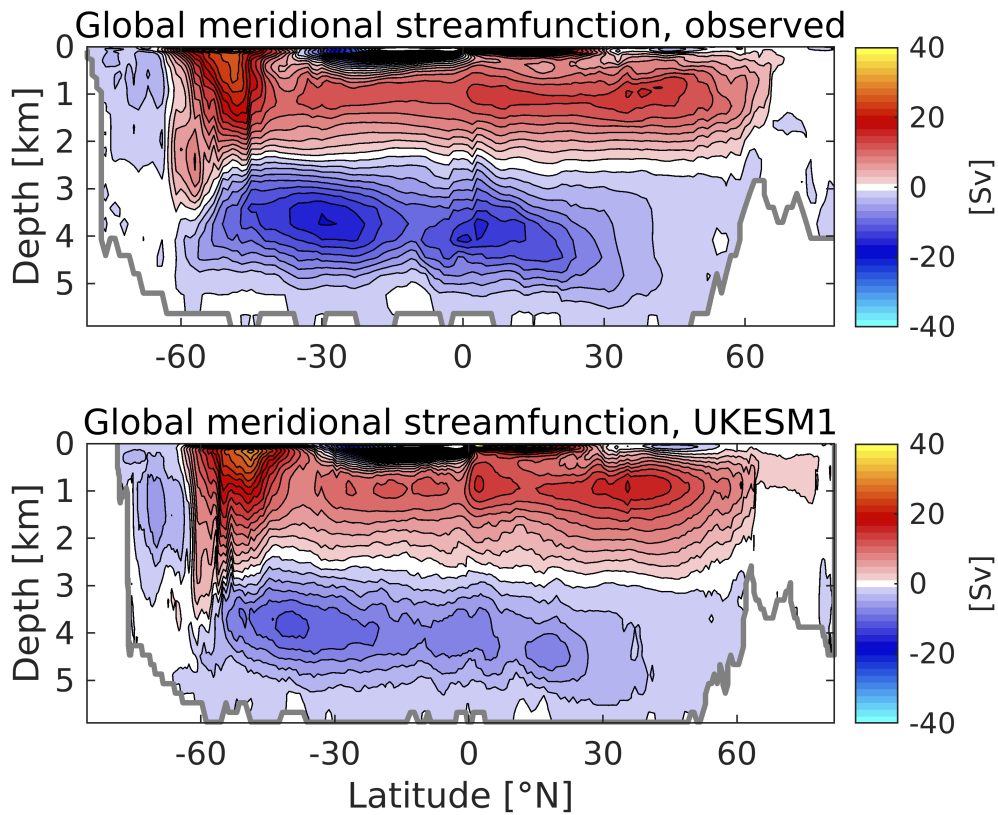


Figure 8. Observationally-derived (top) and simulated (bottom) meridional overturning circulation (MOC) for the global ocean. The [observational circulation is derived from the ECCO V4r4 ocean circulation reanalysis for the period 1992-2017](#). The model circulation shown is based on the decadal-averaged streamfunction. ~~MOC in Sv with both plots including~~, [2000-2009](#). Both plots include the components from parameterised mesoscaled eddies (Gent and McWilliams, 1990; Gent et al., 1995). ~~Observational data from the ECCO V4r4 ocean circulation reanalysis for the period 1992-2017~~ [MOC is in Sv with a contour interval of 2 Sv](#).

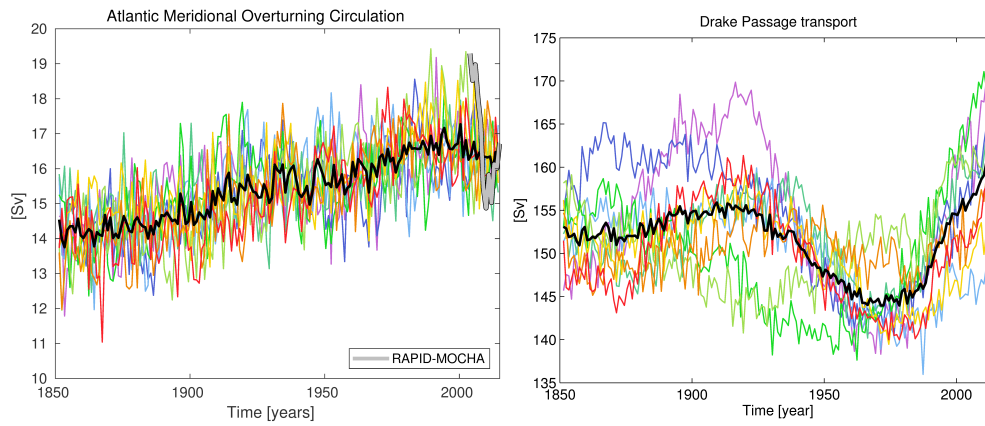


Figure 9. Time-series plots of the ocean circulation during the Historical period from 1850 to 2015. Panels show annual averages of AMOC (left) and Drake Passage (right) transport for all 9 ensemble members (coloured lines) and the ensemble mean (solid black line). [Observational data of AMOC transport from the RAPID-MOCHA array is shown in grey for the period 2003-2015. For additional clarity, Supplementary Figure S8 replots this panel to focus on this recent period.](#)

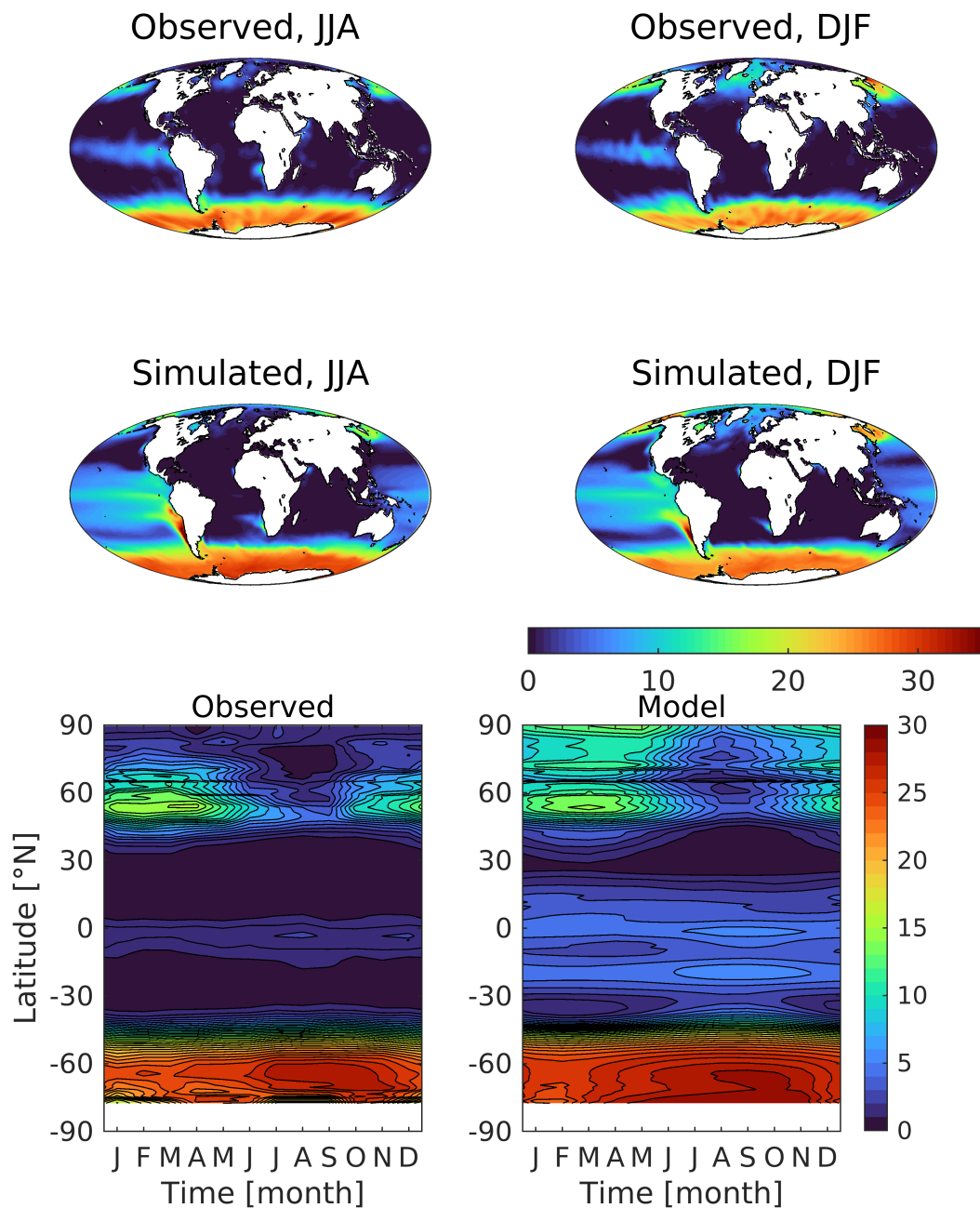


Figure 10. Observational ([lefttop](#); World Ocean Atlas) and simulated ([rightmiddle](#)) surface dissolved inorganic nitrogen, shown geographically for northern ([topleft](#); JJA) and southern summer ([middleright](#); DJF), and as zonal Hovmöller diagrams (bottom). Concentrations in mmol N m^{-3} .

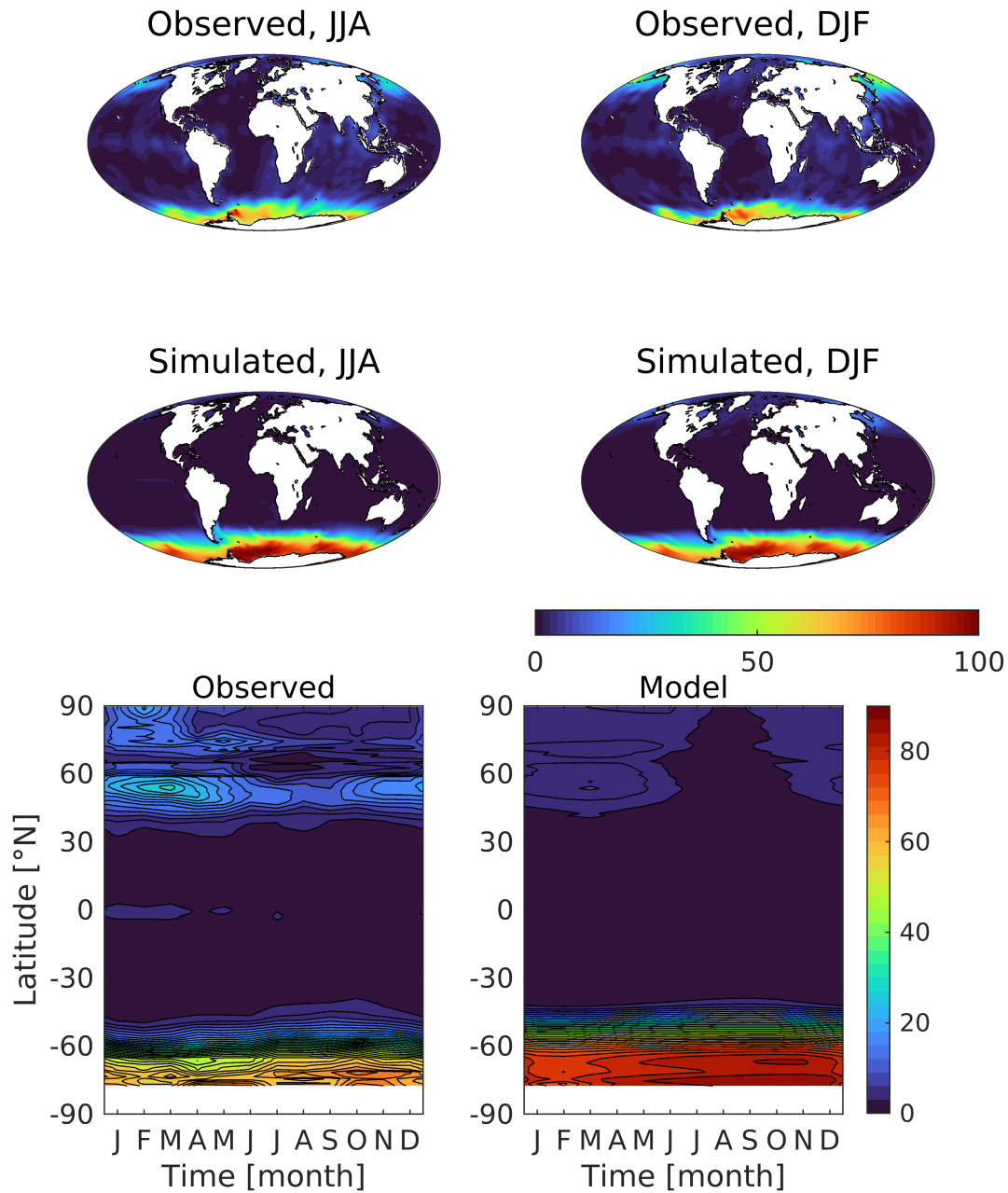
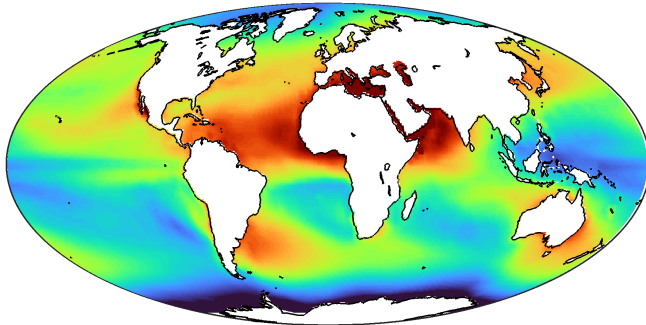


Figure 11. Observational (lefttop; World Ocean Atlas) and simulated (rightmiddle) surface dissolved silicic acid, shown geographically for northern (topleft; JJA) and southern summer (middleright; DJF), and as zonal Hovmöller diagrams (bottom). Concentrations in mmol Si m^{-3} .

Mahowald (2005)



UKESM1

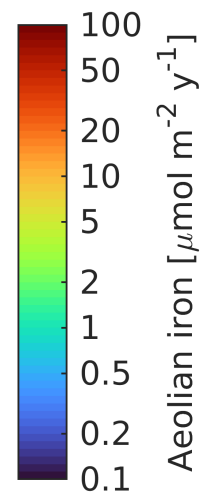
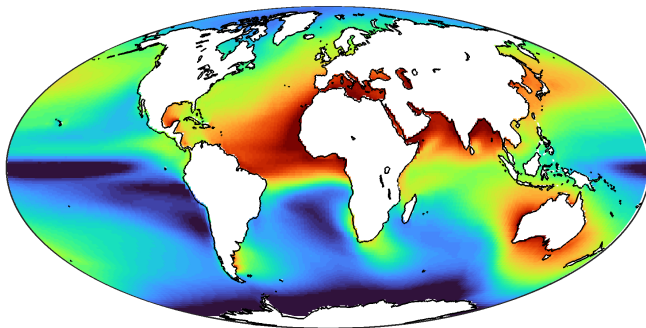


Figure 12. Observational (top; Mahowald (2005)) and simulated ensemble mean (middle) aeolian deposition of iron. Due to its large dynamic range, deposition flux is shown on a logarithmic scale. Deposition is in $\mu\text{mol Fe m}^{-2} \text{y}^{-1}$.

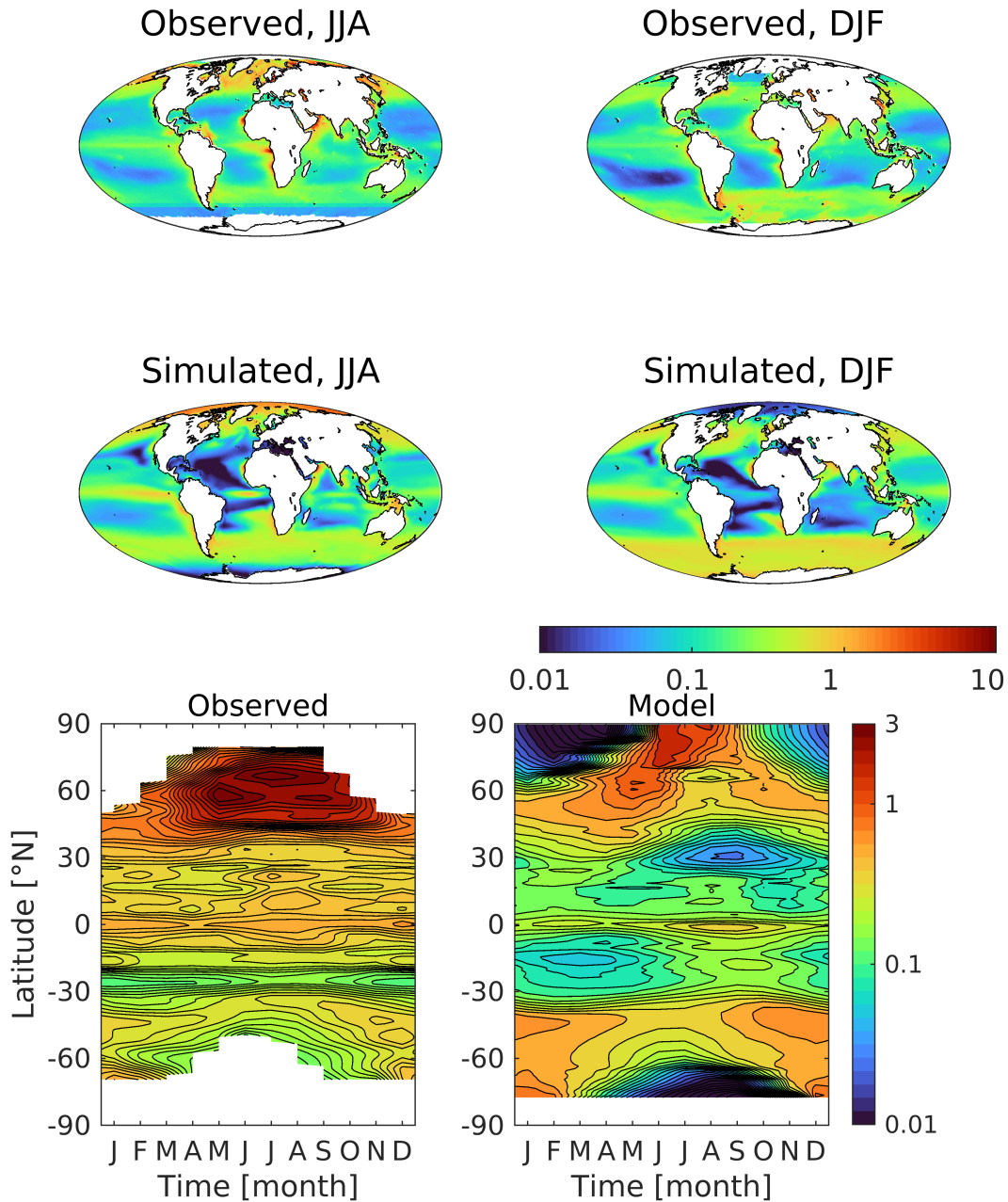


Figure 13. Observational ([lefttop](#); SeaWiFS) and simulated ([rightmiddle](#)) surface chlorophyll, shown geographically for northern ([topleft](#); JJA) and southern summer ([middleright](#); DJF), and as zonal Hovmöller diagrams (bottom). Missing observational data at high latitudes [are shown as near-zero](#) because of polar night / sea-ice [appear as white regions in both geographical and Hovmöller panels](#). Concentrations in mg chl m⁻³.

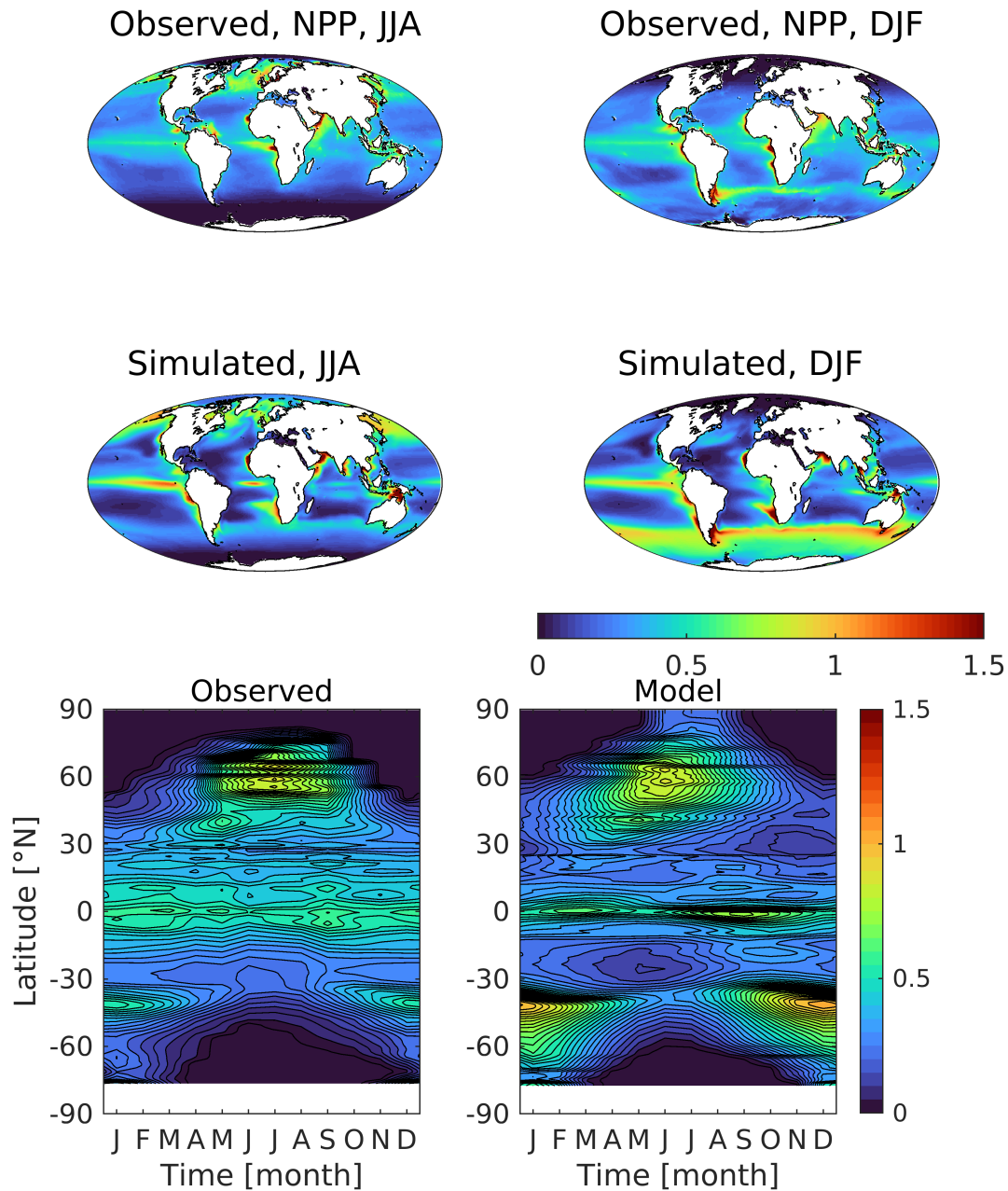


Figure 14. Observational ([left](#)[top](#)) and simulated ([right](#)[middle](#)) vertically-integrated net primary production, shown geographically for northern ([top](#)[left](#); JJA) and southern summer ([middle](#)[right](#); DJF), and as zonal Hovmöller diagrams (bottom). Missing observational data at high latitudes are shown as zero because of polar night / sea-ice. Primary production in $\text{g C m}^{-2} \text{d}^{-1}$.

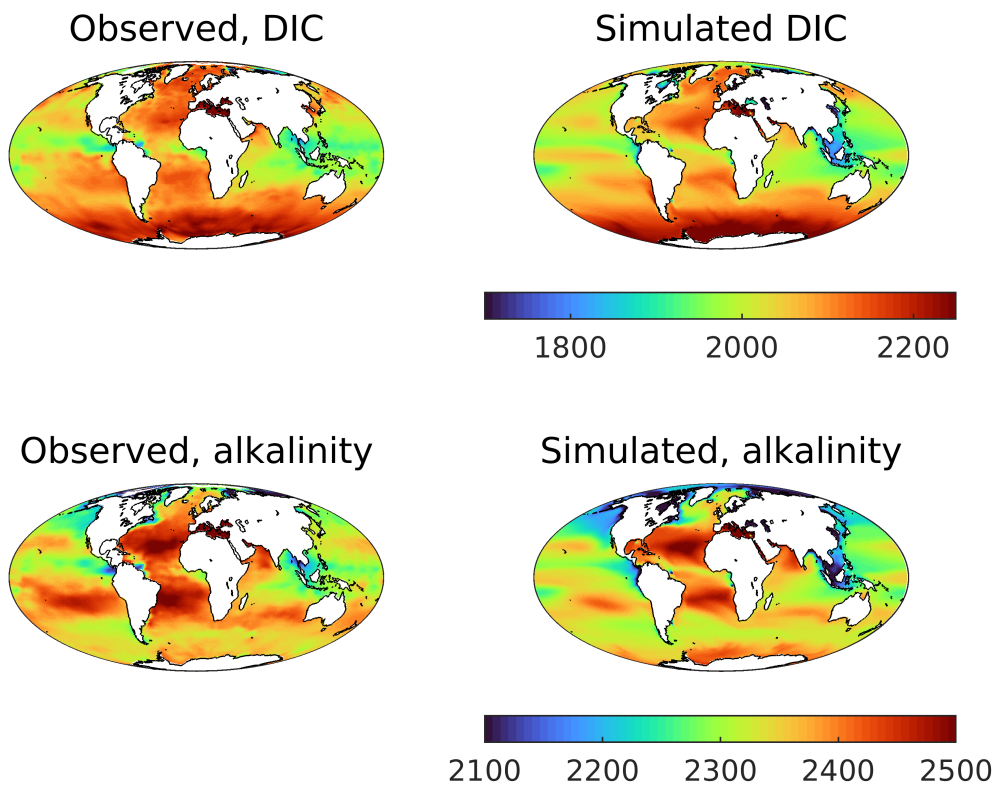


Figure 15. Observational (left; GLODAPv2) and simulated (right) annual average surface dissolved inorganic carbon (top) and total alkalinity (bottom). DIC in mmol C m^{-3} , alkalinity in meq m^{-3} .

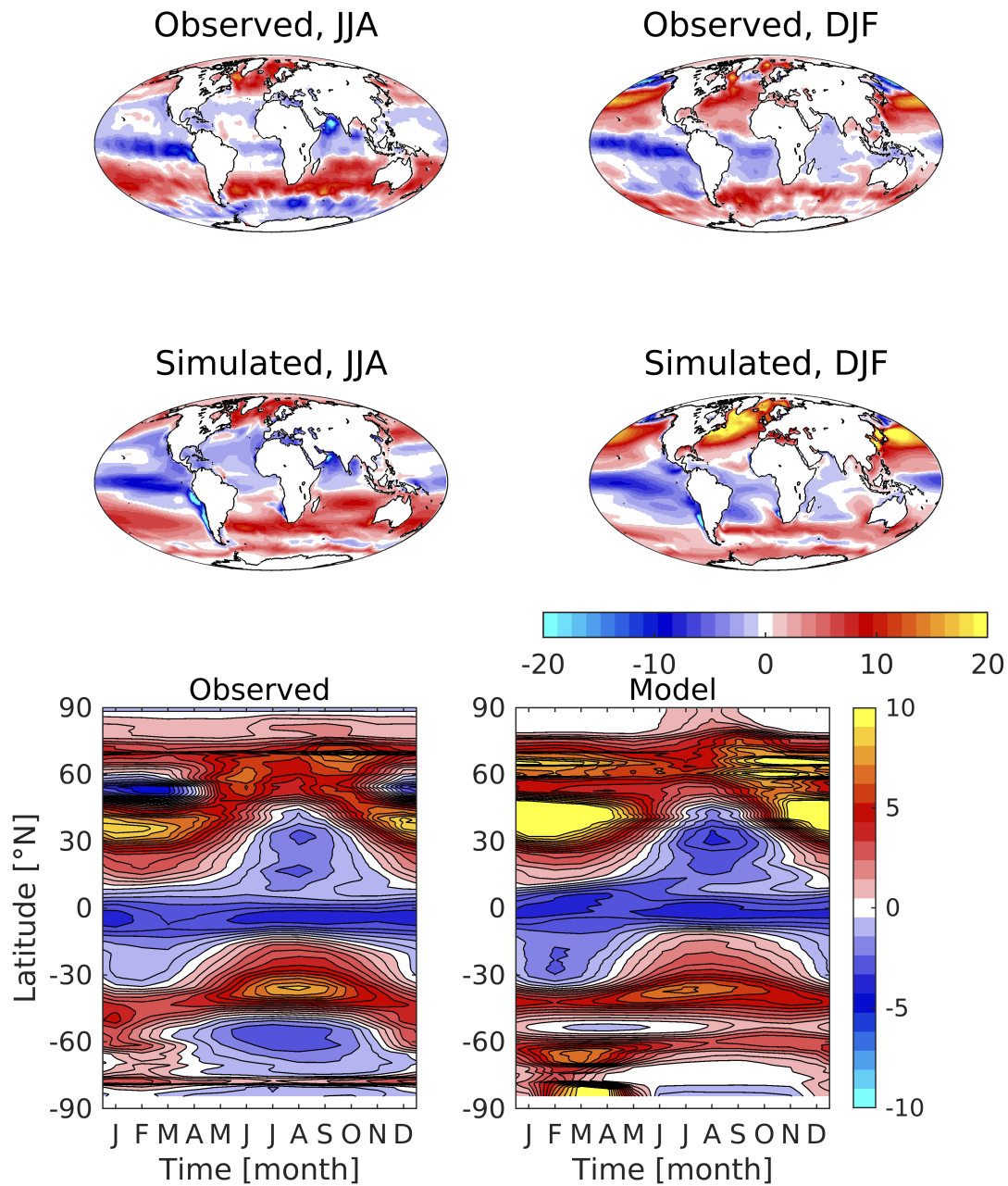


Figure 16. Observed ([lefttop](#); Rödenbeck et al. (2013)) and simulated ([rightmiddle](#)) air-sea CO₂ flux, shown geographically for northern ([topleft](#); JJA) and southern summer ([middleright](#); DJF), and as zonal Hovmöller diagrams (bottom). Red colours indicate CO₂ flux into the ocean, while blue colours denote outgassing CO₂. Flux in mmol C m⁻² d⁻¹.

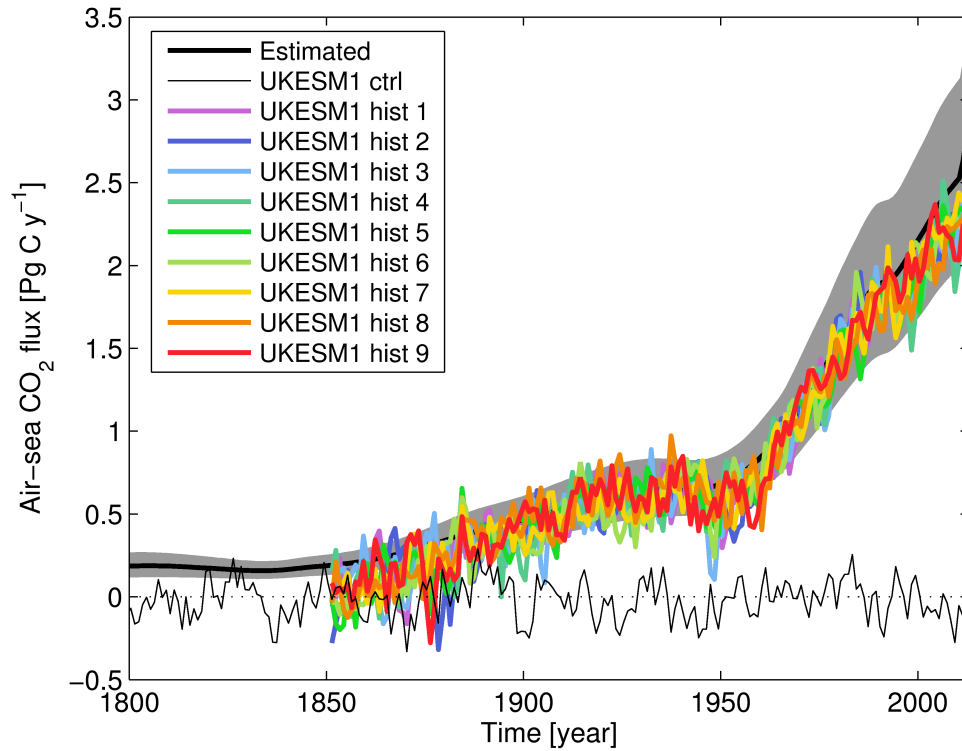


Figure 17. Time-series of globally integrated air-to-sea CO₂ flux, showing observationally-estimated mean (solid black) and range (grey shading; Khatiwala et al., 2009), and simulated ~~control~~ ~~(dashed piControl~~ ~~(thin black)~~ and ~~historical~~ ~~(solid red)~~ ~~fluxes~~ Historical ensemble ~~(solid colours)~~. Air-to-sea fluxes in Pg C y⁻¹. Note that while the Historical era of CMIP6 experiments ~~begin~~ begins in year 1850, the Industrial Revolution – and uptake of anthropogenic CO₂ by the ocean – began prior to this date.

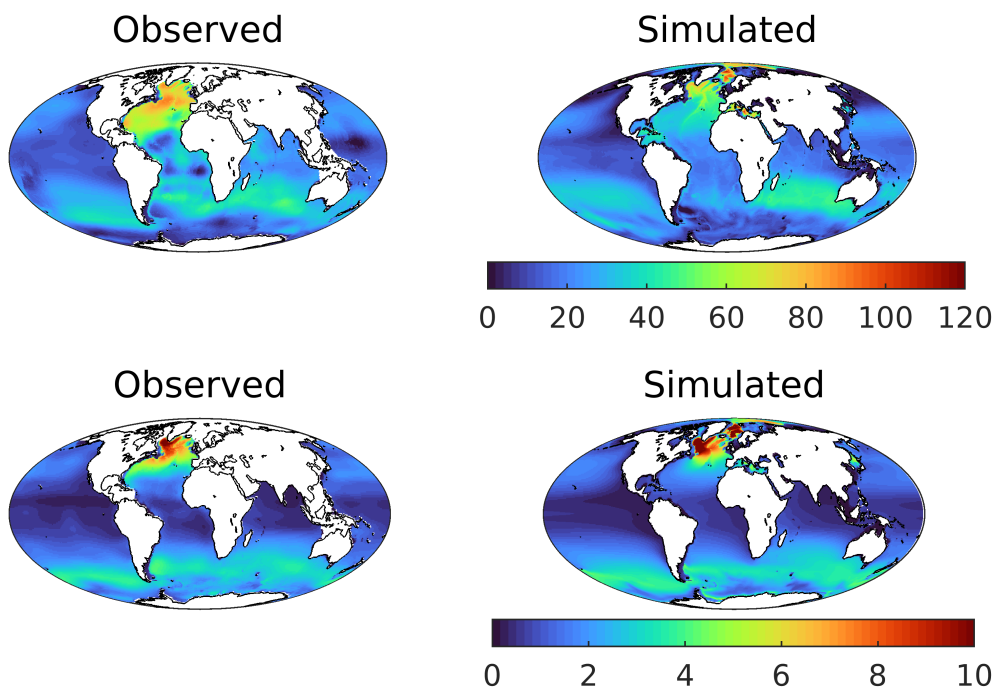


Figure 18. Observed (left; Key et al., 2004) and modelled (right) vertically-integrated anthropogenic CO₂ (top; mol C m⁻²) and CFC-11 (bottom; μmol m⁻²) in the 1990s. GLODAPv1.1 is used here as the time period used overlaps that of observational CFC-11.

1370 Observed (GLODAPv1.1; Key et al., 2004) and modelled vertical profiles of DIC (top left; mmol C m^{-3}), alkalinity (top right; meq m^{-3}), anthropogenic CO_2 (bottom left; mmol C m^{-3}) and CFC-11 (bottom right; nmol m^{-3}). GLODAPv1.1 is used here as the time period used (the 1990s) overlaps that of observational CFC-11. Model profiles from the ensemble used in this analysis are presented individually, and are geographically masked according to GLODAPv1.1.

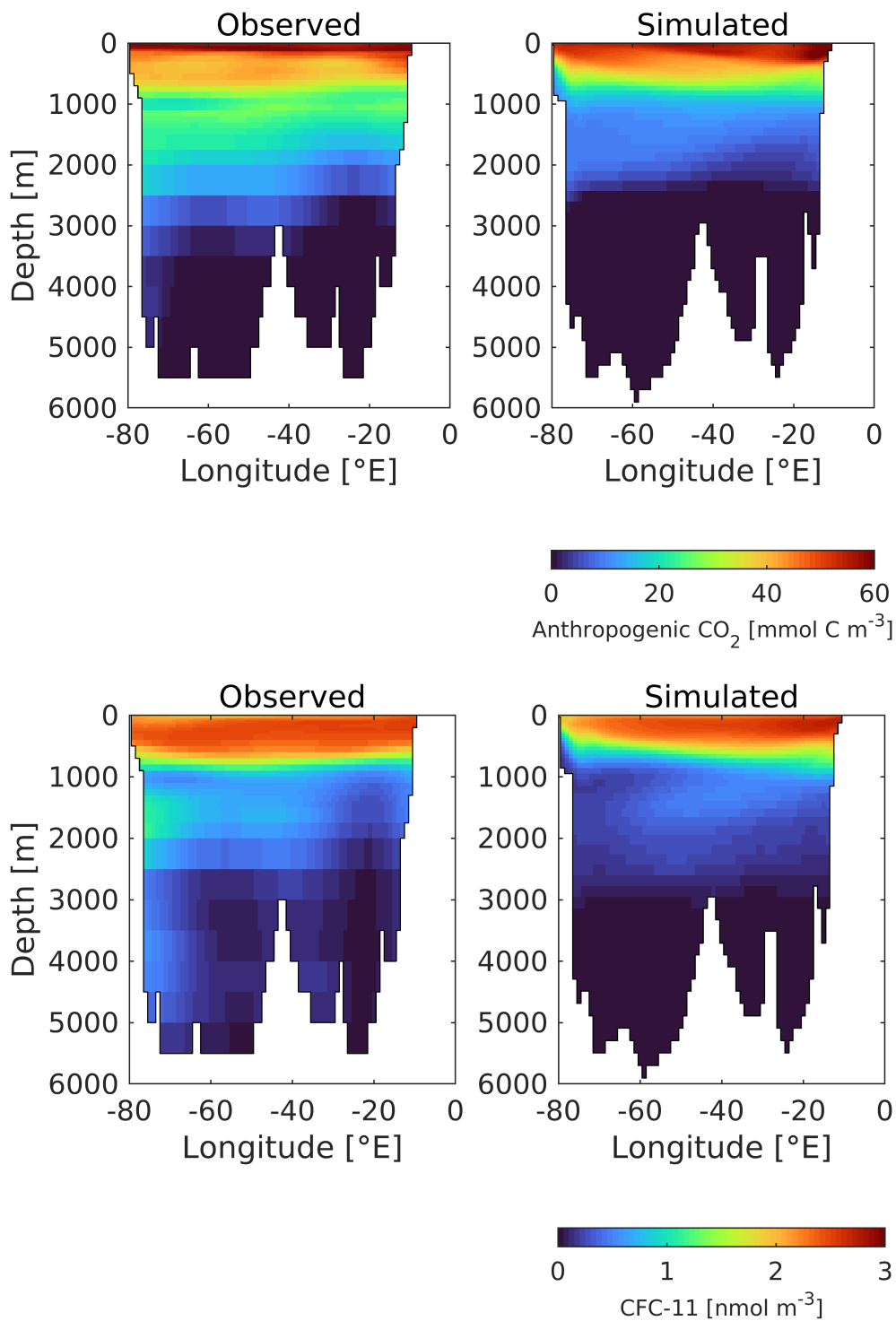


Figure 19. Observed (left; Key et al., 2004) and modelled (right) Atlantic sections (30°N) of anthropogenic CO₂ (top; mmol C m⁻³) and CFC-11 (bottom; nmol m⁻³) in the 1990s.

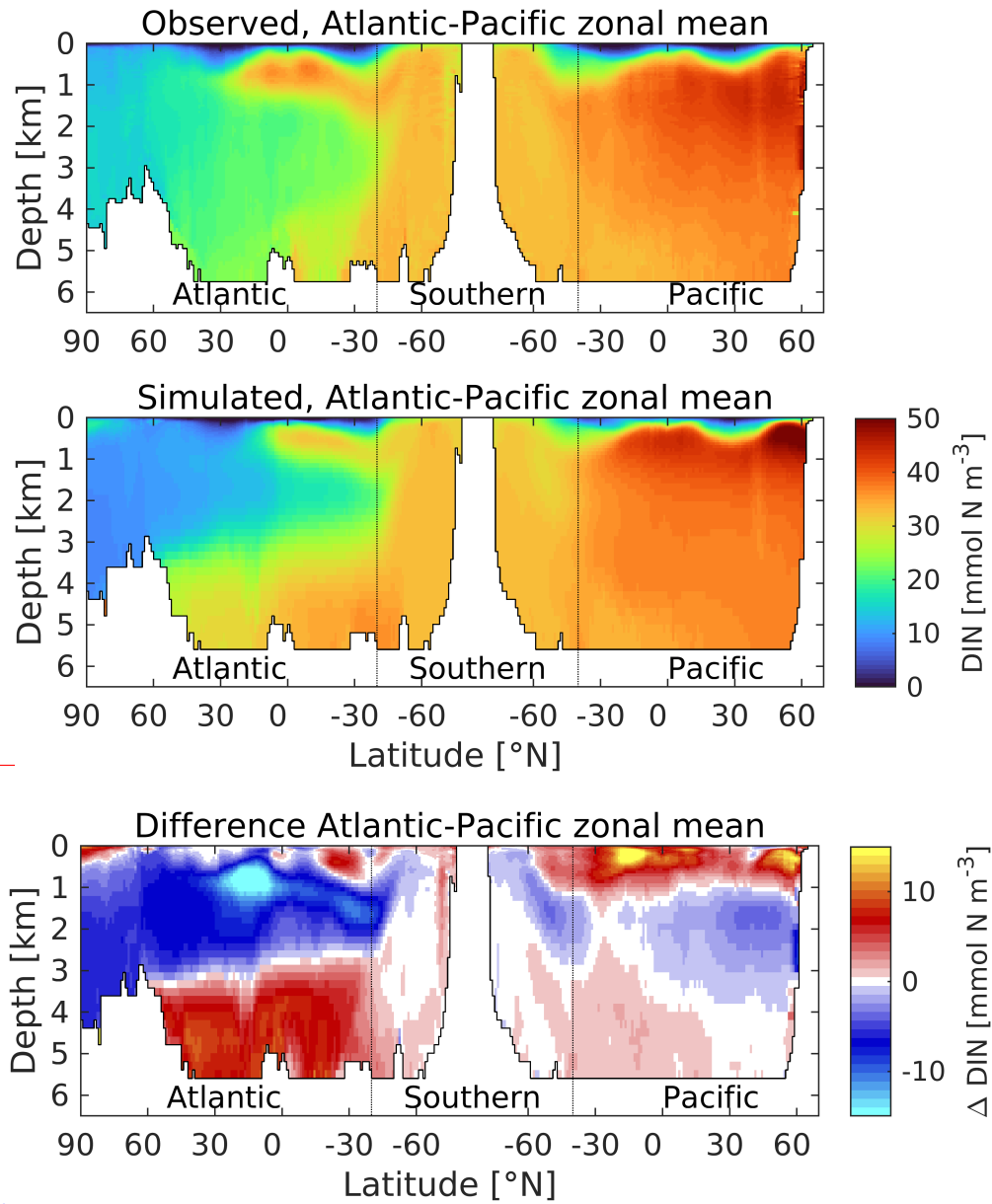


Figure 20. A “thermohaline circulation” section of observed (top) and modelled (bottom middle) zonal average dissolved inorganic nitrogen. Figure 6 explains the format of this section. Difference (simulated - observed) is shown in the bottom panel. Concentrations in mmol N m^{-3} . Figure 6 explains the format of this section.

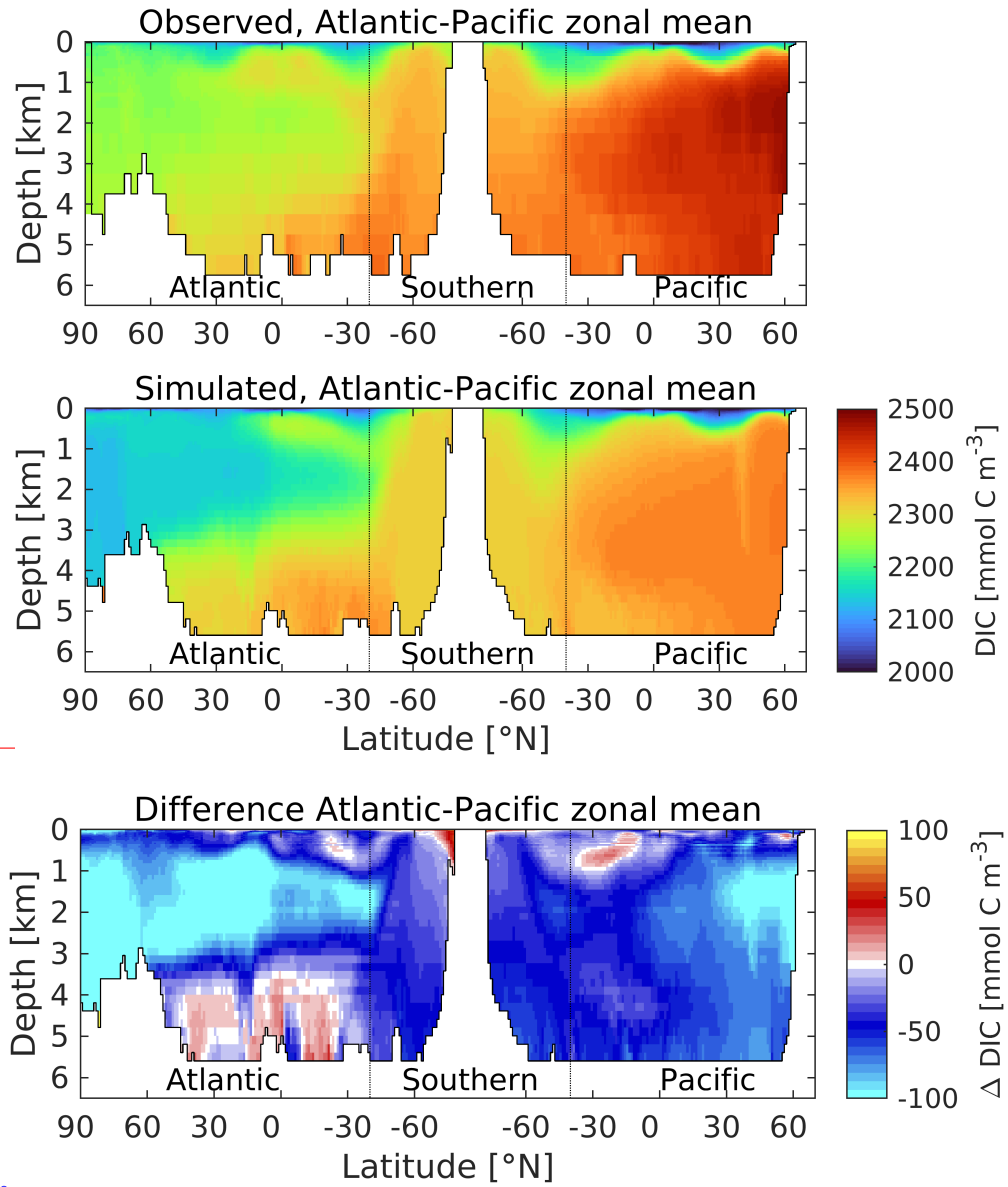


Figure 21. A “thermohaline circulation” section of observed (top) and modelled (bottom middle) zonal average dissolved inorganic carbon. Difference (simulated - observed) is shown in the bottom panel. Concentrations in mmol C m^{-3} . Figure 6 explains the format of this section. Concentrations in-

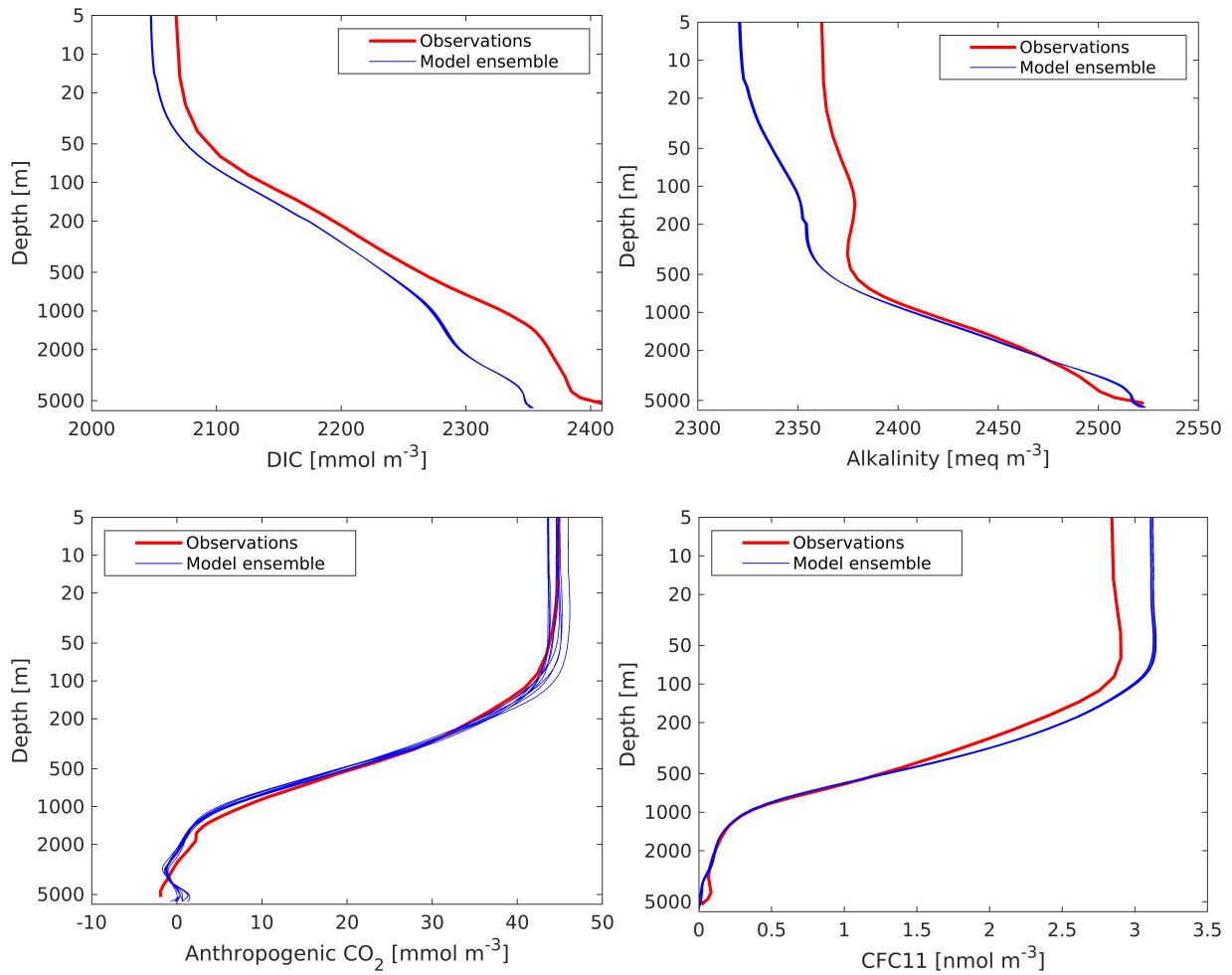


Figure 22. Observed (GLODAPv1.1; Key et al., 2004) and modelled vertical profiles of DIC (top left; mmol C m^{-3}), alkalinity (top right; meq m^{-3}), anthropogenic CO_2 (bottom left; mmol C m^{-3}) and CFC-11 (bottom right; nmol m^{-3}). GLODAPv1.1 is used here as the time period used (the 1990s) overlaps that of observational CFC-11. Model profiles from the ensemble used in this analysis are presented individually, and are geographically masked according to GLODAPv1.1.

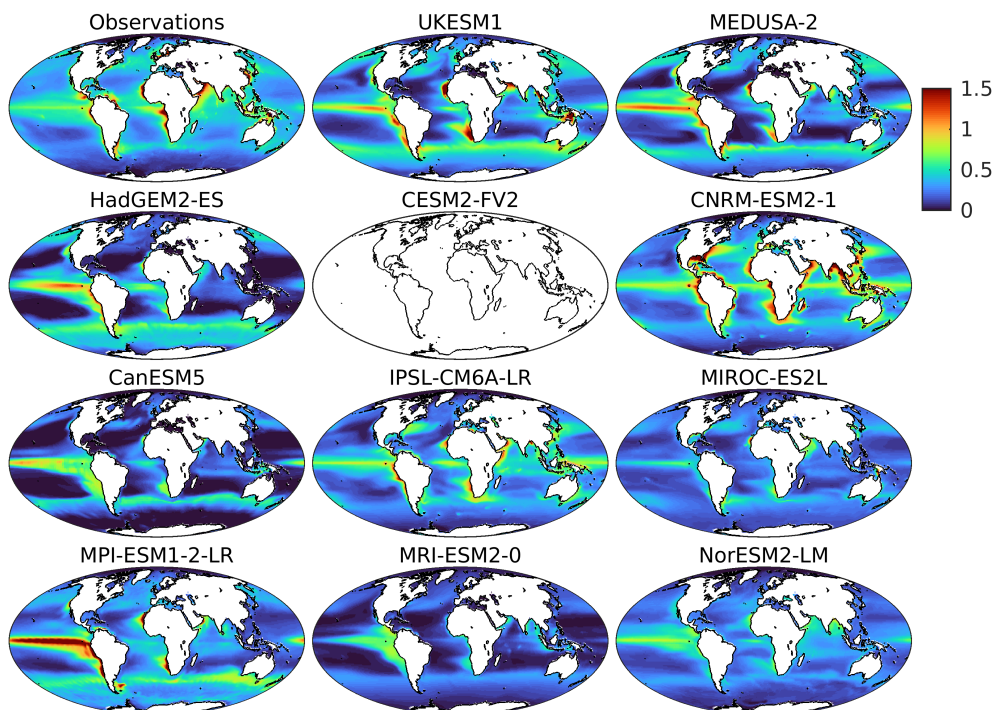


Figure 23. Intercomparison of annual mean vertically-integrated primary production between observed (top row, left), UKESM1 simulated (top row, centre) and a range of comparable CMIP6 models (rows 2–4). Results from CMIP5’s precursor to UKESM1, HadGEM2-ES (Jones et al., 2011; row 2, left) and MEDUSA–2 (Yool et al., 2013; top row, right) are shown for comparison. This field was not available for the CESM2-FV2 model and this has been left blank. Production in $\text{g C m}^{-2} \text{d}^{-1}$.

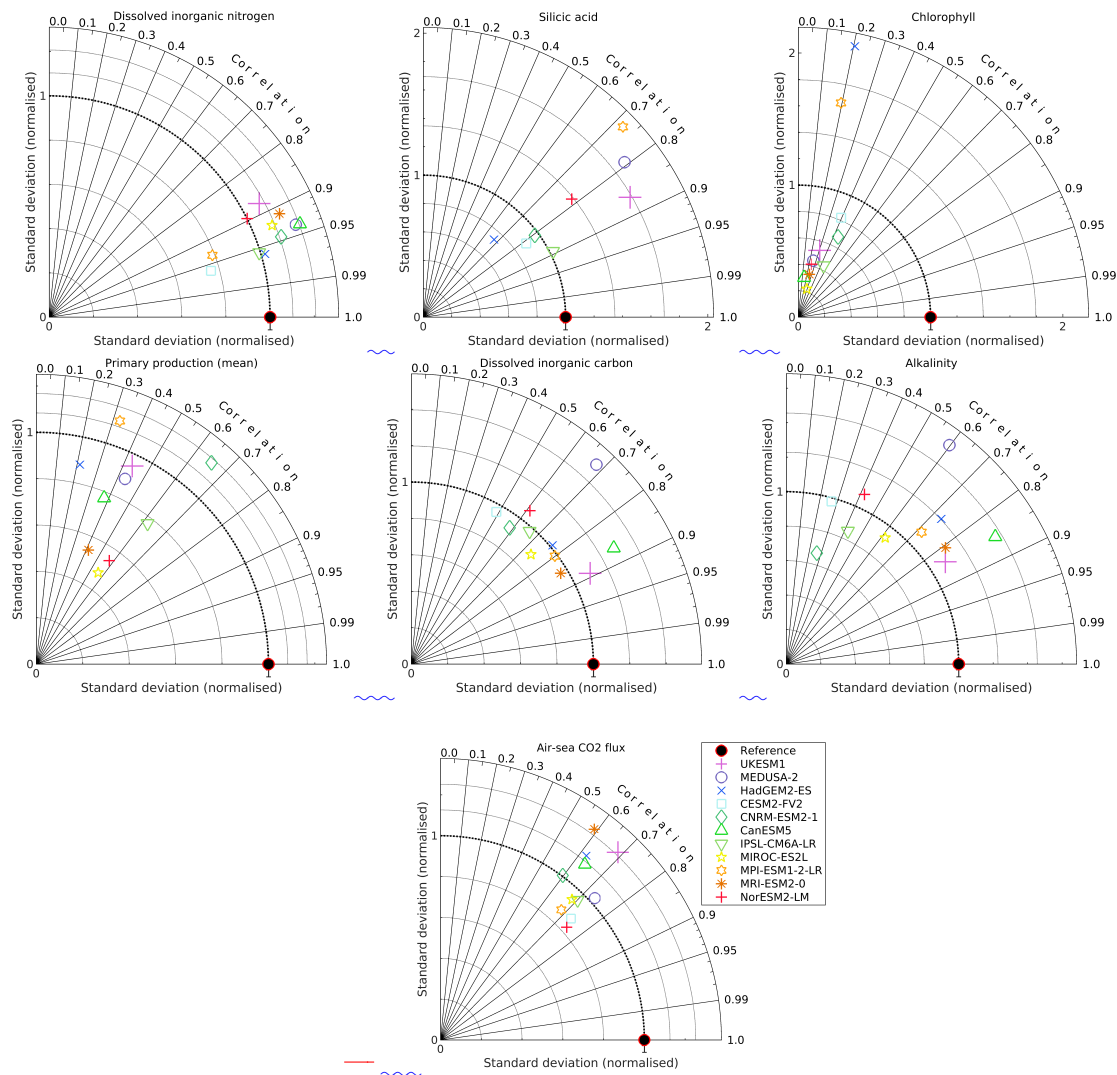


Figure 24. Taylor diagrams illustrating the skill of UKESM1 (and its precursors, HadGEM2-ES and MEDUSA-2) and a series of CMIP6 models over a set of standard surface ocean of biogeochemical properties: dissolved inorganic nitrogen (top left), silicic acid (top centre), chlorophyll (top right), primary production (middle left), dissolved inorganic carbon (middle centre), alkalinity (middle right), and air-sea CO₂ flux (bottom centre). The diagrams show model-observation comparisons based on annual average spatial fields, all regridded to the same standard grid. The diagrams share a common model key (bottom centre).

Property [units]	Mean	σ	Min.	Max.	M
AMOC	15.83	1.191	12.60	18.80	0.200
[Sv]	<i>14.77</i>	<i>0.877</i>	<i>12.61</i>	<i>16.88</i>	<i>-0.024</i>
Drake	151.52	5.908	139.75	164.18	-0.197
[Sv]	<i>154.33</i>	<i>4.017</i>	<i>144.84</i>	<i>163.37</i>	<i>0.128</i>
SST	17.76	0.175	17.45	18.42	0.015
[°C]	<i>17.67</i>	<i>0.075</i>	<i>17.45</i>	<i>17.88</i>	<i>-0.002</i>
Temperature	3.78	0.008	3.76	3.80	-0.000
[°C]	<i>3.77</i>	<i>0.007</i>	<i>3.76</i>	<i>3.79</i>	<i>-0.001</i>
SSS	34.31	0.015	34.27	34.34	-0.001
[PSU]	<i>34.31</i>	<i>0.013</i>	<i>34.28</i>	<i>34.34</i>	<i>-0.001</i>
Salinity	34.73	0.000	34.73	34.73	-0.000
[PSU]	<i>34.73</i>	<i>0.000</i>	<i>34.73</i>	<i>34.73</i>	<i>-0.000</i>
N sea-ice	12.23	0.519	10.69	13.39	-0.025
[10 ⁶ km ²]	<i>12.15</i>	<i>0.387</i>	<i>11.18</i>	<i>13.30</i>	<i>0.017</i>
S sea-ice	11.35	0.890	8.34	13.08	-0.092
[10 ⁶ km ²]	<i>11.87</i>	<i>0.554</i>	<i>10.44</i>	<i>13.27</i>	<i>-0.013</i>
MLD	50.06	0.729	48.02	52.17	0.017
[m]	<i>49.99</i>	<i>0.547</i>	<i>48.61</i>	<i>51.59</i>	<i>-0.010</i>

Table 1. Selected ocean physical properties averaged across both the Historical ensemble (upper rows) and corresponding segments of the piControl (lower *italicised* rows). For each property, the statistics refer to the full 165 y period from 1850-2015. The final statistic, M , is the linear slope of the change in the property across this full period.

Property [units]	Mean	σ	Min.	Max.	M
Surface DIN	7.52	0.146	7.18	7.90	0.008
[mmol N m ⁻³]	<i>7.49</i>	<i>0.125</i>	<i>7.20</i>	<i>7.82</i>	<i>0.003</i>
Surface silicic acid	9.30	0.312	8.69	10.10	0.038
[mmol Si m ⁻³]	<i>9.16</i>	<i>0.189</i>	<i>8.66</i>	<i>9.58</i>	<i>-0.010</i>
Surface iron	0.52	0.005	0.50	0.53	-0.000
[μ mol Fe m ⁻³]	<i>0.52</i>	<i>0.004</i>	<i>0.51</i>	<i>0.53</i>	<i>0.000</i>
Surface DIC	2020.70	16.299	2000.76	2059.53	3.267
[mmol C m ⁻³]	<i>2002.35</i>	<i>1.043</i>	<i>1999.81</i>	<i>2005.20</i>	<i>-0.017</i>
Surface alkalinity	2317.76	1.617	2314.60	2321.42	0.265
[meq m ⁻³]	<i>2316.09</i>	<i>0.817</i>	<i>2314.10</i>	<i>2317.93</i>	<i>-0.053</i>
Surface O ₂	252.05	0.735	249.28	253.33	-0.060
[mmol O ₂ m ⁻³]	<i>252.42</i>	<i>0.334</i>	<i>251.43</i>	<i>253.35</i>	<i>0.009</i>
Ocean O ₂	190.50	0.373	189.87	191.17	0.050
[mmol O ₂ m ⁻³]	<i>190.41</i>	<i>0.279</i>	<i>189.88</i>	<i>190.89</i>	<i>0.051</i>
NPP	47.96	0.728	46.04	49.83	-0.014
[Pg C y ⁻¹]	<i>48.01</i>	<i>0.682</i>	<i>46.05</i>	<i>49.74</i>	<i>-0.003</i>
Air-sea CO ₂ flux	0.81	0.674	-0.17	2.45	0.129
[Pg C y ⁻¹]	<i>-0.02</i>	<i>0.118</i>	<i>-0.33</i>	<i>0.26</i>	<i>-0.001</i>
Aeolian iron	2.41	0.228	1.89	3.07	-0.002
[Gmol Fe y ⁻¹]	<i>2.41</i>	<i>0.219</i>	<i>1.88</i>	<i>3.12</i>	<i>0.003</i>

Table 2. Selected ocean biogeochemical properties averaged across both the Historical ensemble (upper rows) and corresponding segments of the piControl (lower *italicised* rows). For each property, the statistics refer to the full 165 y period from 1850-2015. The final statistic, M , is the linear slope of the change in the property across this full period.

Field	Global	Southern	St S Atl	Eq Atl	St N Atl	Sp N Atl	St S Pac	Eq Pac	St N Pac	Sp N Pac	Arctic
<i>Surface DIN</i>											
Observed	5.227	23.473	3.904	0.366	0.436	4.444	2.515	2.472	0.456	9.942	3.298
Model	7.757	25.676	4.952	0.058	0.273	3.039	8.543	8.004	3.256	9.632	8.503
<i>Surface silicic acid</i>											
Observed	7.512	32.019	2.665	2.048	1.487	3.185	1.855	2.828	3.151	21.299	8.329
Model	10.092	62.875	5.869	0.411	0.743	2.163	1.978	0.415	0.430	4.496	4.572
<i>Surface chlorophyll</i>											
Observed	0.219	0.164	0.249	0.361	0.195	0.517	0.131	0.184	0.143	0.558	0.342
Model	0.262	0.387	0.335	0.071	0.106	0.406	0.252	0.249	0.138	0.509	0.472
<i>Primary production</i>											
Observed	0.317	0.099	0.345	0.555	0.359	0.360	0.277	0.461	0.323	0.338	0.115
Model	0.356	0.309	0.466	0.324	0.217	0.365	0.358	0.510	0.263	0.466	0.153
<i>Surface DIC</i>											
Observed	2071	2192	2113	2036	2100	2117	2074	1991	2008	2058	2053
Model	2058	2211	2096	2022	2100	2084	2055	1981	1985	2030	2021
<i>Surface Alkalinity</i>											
Observed	2355	2350	2410	2386	2446	2353	2373	2317	2324	2268	2206
Model	2327	2361	2371	2363	2438	2299	2335	2291	2286	2201	2171
<i>Air-sea CO₂ flux</i>											
Observed	1.043	-0.047	2.156	-1.493	1.331	5.136	1.665	-2.623	2.229	1.755	2.748
Model	1.350	2.122	1.862	-1.556	0.573	8.050	1.171	-3.442	1.954	5.120	3.176

Table 3. Selected biogeochemical properties averaged for specific geographical regions for annual mean fields. Observed and model values shown, with model values averaged over the Historical ensemble. Regional abbreviations are “St” for subtropical (10°S–10°N) and “Sp” for subtropical (40°–70°). In the southern hemisphere, the subtropical region falls primarily within the Southern Ocean, although as its northern margin is delineated at -50°N rather than -40°N, the southern margins of the southern subtropical Atlantic and Pacific extend to -50°N. The Indian Ocean is excluded from this analysis for simplicity. Throughout, the model domain used matches that available from observational fields.