The Sensitivity of a Coupled Climate Model to Numerical Mixing in Its Ocean Component

ALEX MEGANN,[a] DAN COPSEY,[b] AND AMBER WALSH[a]

[a] *National Oceanography Centre, Southampton, United Kingdom*
[b] *Met Office, Exeter, United Kingdom*

(Manuscript received 18 November 2025, in final form 6 August 2026, accepted 31 August 2026)

ABSTRACT: An ensemble of seven integrations of the Hadley Centre Global Environment Model, version 3, Global Coupled 5 configuration (HadGEM3-GC5) coupled model is used to investigate the sensitivity of the simulated climate system to settings in the ocean component that have been shown to affect the level of numerical mixing in forced simulations. This configuration is closely related to the U.K. contribution to CMIP7. The ensemble is integrated for 60 years with constant year 2000 greenhouse forcing. The ocean surface temperature has a strong and consistent response to numerical mixing, with increased mixing leading consistently to warming over the global ocean by up to 0.5°C, while reducing mixing cools the surface by a similar degree. The response of the surface air temperature is very similar to that of the SST but is seasonally amplified at high latitudes in the respective winter. Robust and strong sensitivities are also found for sea ice cover, rainfall, and the surface shortwave and latent heat fluxes, the latter two showing opposing changes in their global means of over 2 W m^{-2} across the ensemble. We present large-scale ocean and atmospheric metrics and discuss mechanisms for the sensitivity of surface temperatures in these simulations to numerical mixing in the ocean, in which more mixing warms the surface and vice versa. The magnitude of this sensitivity of the surface temperature is significant, since it is comparable with the changes found in CMIP simulations with historical or future greenhouse scenario forcings, and we speculate on the implications for modeling future climates.

SIGNIFICANCE STATEMENT: This study demonstrates that spurious numerical mixing in the ocean component of an IPCC-class coupled climate model can significantly affect the simulated ocean and atmospheric surface temperatures, leading to a bias in the coupled system with a magnitude comparable to that from anthropogenic greenhouse warming over the past few decades. Increasing numerical mixing leads to a warming of the surface and vice versa, with a surface temperature response of order 0.5°C in the global mean, with seasonal amplification of the surface air temperature changes at high latitudes.

KEYWORDS: Atmosphere-ocean interaction; Climate; Diapycnal mixing; Sea surface temperature; Thermohaline circulation

1. Introduction

Turbulent vertical mixing below the mixed layer base has long been recognized as a major mechanism for maintaining the thermohaline structure of the ocean. In the ocean interior, the stratification is maintained to first order by the balance between vertical advection and diffusion (Munk 1966; Huber et al. 2015). The stratification in the main thermocline is also controlled by mixing (Moum and Osborn 1986; Dewar and McWilliams 2019). It is the dominant process by which Antarctic Bottom Water (AABW) is consumed by mixing with the overlying deep waters as it propagates northward (Munk 1966; Polzin et al. 1997; Munk and Wunsch 1998; de Lavergne et al. 2016). Dense Arctic overflows undergo substantial mixing as they entrain lighter ambient waters downstream of the sills (Dickson and Brown 1994) to form North Atlantic Deep Water (NADW), and as NADW circulates southward, it mixes with both AABW and overlying Atlantic Water (Speer 1997; Talley 2013; Naveira Garabato et al. 2014).

Global diapycnal upwelling rates have been estimated from observational data, including hydrographic inverse models and tracer budgets. The total formation rate of AABW is thought to be between 17 and 19 Sv (1 Sv $\equiv 10^6$ m^3 s^{-1}) (Orsi et al. 2002; Naveira Garabato et al. 2014), and in a steady state, this must be balanced by the consumption of this water mass by mixing. Macdonald and Wunsch (1996) suggest that about 15 Sv of AABW and Circumpolar Deep Water is converted through mixing during its passage to the South Atlantic and that this mainly occurs in the Indian Ocean. Lumpkin and Speer (2007, hereafter LS2007) use an inverse model of global circulation to infer that 5.6 ± 3.0 Sv of AABW enters the Atlantic at 32°S to be transformed to NADW by diapycnal mixing at tropical and subtropical latitudes. Cessi (2019, hereafter C2019) evaluated overturning streamfunctions against potential density in the ECCO4 reanalysis to estimate water mass transformations, suggesting a lower value of about 15 Sv for the bottom water cell strength, and conclude that less than 2 Sv of AABW penetrate into the Atlantic, with most of the bottom water traveling into the Indian and Pacific Oceans. In the latter ocean basins, the absence of strong convective

Denotes content that is immediately available upon publication as open access.

Corresponding author: Alex Megann, apm@noc.ac.uk

DOI: 10.1175/JCLI-D-25-0661.1
© 2026 Author(s). This published article is licensed under the terms of a Creative Commons Attribution 4.0 International (CC BY 4.0) License.

Brought to you by NATIONAL OCEANOGRAPHY CENTRE | Unauthenticated | Downloaded 11/11/26 12:54 PM UTC

sinking at high northern latitudes means that bottom and deep waters that do not leave the basins as deep boundary currents are mixed diabatically with lighter waters and form Indian Deep Water and Pacific Deep Water, respectively: Wijffels et al. (1996) estimate that 8 Sv of bottom water crosses 10°N in the Pacific, and LS2007 estimate that 5 Sv reaches 24°N, with the implication that this net flow is balanced by diapycnal mixing to lighter water masses.

In the Atlantic, diffusive upwelling occurs between the southward-flowing NADW and the overlying intermediate waters: Scaling arguments (Robinson and Stommel 1959) suggest that the Atlantic overturning strength should depend, in the absence of other influences, on the vertical diffusivity to the 2/3 power. Although the lifting of NADW in the Southern Ocean is now believed to be dominated by wind-driven near-adiabatic upwelling (Toggweiler and Samuels 1995, 1998; Marshall and Speer 2012), the rate of mixing of NADW with overlying intermediate waters in the tropics and subtropics on the mean overturning strength has a significant effect on the overturning strength (see Cimoli et al. 2023 for a review). The streamfunction derived from inverse analysis shown as Fig. 3a of LS2007 suggests that about 4 Sv of southward-flowing NADW upwells diabatically between 24°N and 32°S. The dependence of the Atlantic meridional overturning circulation (AMOC) strength on diffusivity has been investigated extensively in numerical models. Bryan (1987) confirmed the 2/3 power in an idealized model, and Jayne (2009) used a wide range of diffusivity parameterizations in the CCSM3 climate model to show a strong positive correlation between the strength of the upper AMOC cell and the rate of dissipation. Nikurashin and Vallis (2011) demonstrated with an idealized model that, in the absence of diapycnal mixing, the AMOC strength depended linearly on Southern Ocean wind stress forcing but that with realistic wind stresses and diffusivities the mixing-driven component of the overturning was of order half the strength of the wind-driven part.

Numerical ocean models have been known for several years to exhibit spurious numerical mixing, which causes diffusion of heat and salt over and above that contributed by the model's mixing scheme. This numerical mixing arises chiefly from truncations in the tracer advection scheme (Griffies et al. 2000) and manifests in several ways, including mixing by internal waves and tides (Megann 2018, 2024), excessive mixing of overflow waters at sills (Bruciaferri et al. 2018, 2024), and by the interaction of poorly resolved mesoscale velocities with the tracer fields (Megann and Storkey 2021). The latter process results in spurious horizontal diffusion of tracers but is also associated with vertical velocities on the grid scale that are coherent in the vertical on scales of a thousand meters or more (Megann 2018). Several studies have quantified numerical mixing in forced ocean configurations through density transformation rates (Lee et al. 2002; Urakawa and Hasumi 2014; Megann 2018), through changes in global available potential energy (Ilıcak et al. 2012), or through local heat budgets (Holmes et al. 2021).

Density-coordinate, or isopycnic, models such as Miami Isopycnic Coordinate Ocean Model (MICOM) (Bleck and Smith 1990) and Generalized Ocean Layer Dynamics (GOLD) (Hallberg and Adcroft 2009) are known to provide substantial advantages in reducing numerical mixing because in this framework vertical advection of tracers is zero by construction. Nevertheless, notable disadvantages of pure isopycnic models include a poorly defined entrainment process into the mixed layer, a lack of vertical resolution in weakly stratified regions such as the Arctic, and the potential for the density used as the vertical coordinate to be nonmonotonic in certain (mainly high latitude) regions. Hybrid vertical coordinates, or arbitrary Lagrangian–Eulerian (ALE) coordinates, which combine the advantages of geopotential and isopycnic coordinates, include the Hybrid Coordinate Ocean Model (HYCOM; Bleck 2002) and version 6 of the Modular Ocean Model (MOM6; Adcroft et al. 2019) and offer at least a partial remedy for numerical mixing in both forced and coupled ocean models. The $z\sim$ filtered arbitrary Lagrangian–Eulerian vertical coordinate (Leclair and Madec 2011), which allows the coordinate interfaces to displace vertically with the vertical velocity on time scales of a few days, allows a nominally fixed-coordinate model to benefit from some of the advantages of an ALE coordinate by allowing internal waves and tides to be simulated in a nearly adiabatic way. Megann et al. (2022) demonstrated the effectiveness of the $z\sim$ coordinate in a $1/4°$ global forced Nucleus for European Modelling of the Ocean (NEMO) configuration, confirming that it was effective in reducing numerical mixing, with its impact increasing as the filter time-scale parameter was lengthened to include more of the internal wave spectrum, and Megann (2024) showed, using the same ocean configuration, that the $z\sim$ coordinate significantly reduced numerical mixing when tidal forcing was applied.

Several investigations have been made into the sensitivity of coupled climate models to the vertical coordinate of its ocean component. Megann et al. (2010) compared two climate models with identical atmosphere and sea ice components but with one (HadCM3) using a z-coordinate ocean and the other [Coupled Hadley–Isopycnic Model Experiment (CHIME)] using an ocean component based on the HYCOM hybrid-coordinate model (Bleck 2002), in which most of the ocean interior is represented by isopycnal layers with constant potential density. The model with the hybrid isopycnal ocean had a more realistic representation of the subtropical thermocline, along with better preservation of cold bottom waters in the North Atlantic. Dunne et al. (2012) compared two versions of the GFDL Earth system model, one (ESM2M) with a fixed-depth ocean component and the other (ESM2G) with the GOLD isopycnal ocean, reporting that ESM2G had a shallower and less-ventilated thermocline, more vigorous bottom flows, more extensive sea ice, shallower mixing and mode water formation, and an absence of long-term warming drift. We note that the latest version of the GFDL Earth system model (ESM4; Dunne et al. 2020), based on the physical coupled model CM4.0 (Held et al. 2019), uses the MOM6 ALE coordinate ocean model, which is configured to reduce numerical mixing.

In this study, we use an ensemble of the Met Office coupled climate model Hadley Centre Global Environment Model, version 3, Global Coupled 5 configuration (HadGEM3-GC5)

Brought to you by NATIONAL OCEANOGRAPHY CENTRE | Unauthenticated | Downloaded 11/11/26 12:54 PM UTC

to investigate the sensitivity of the simulated climate to numerical mixing in the ocean. The main objectives are to

1) quantify the numerical mixing in the 1/4° NEMO ocean component of HadGEM3-GC5 and compare with the mixing from the explicit mixing scheme and with observational estimates of mixing,
2) assess previously identified remedies for numerical mixing in this configuration,
3) identify the sensitivities of principal ocean metrics to changes in numerical mixing,
4) quantify the sensitivity of atmospheric metrics to numerical mixing in the ocean component, and
5) make recommendations for future coupled model development.

In section 2, we briefly describe the HadGEM3-GC5 model and introduce the experimental design. Section 3 defines a metric for total diapycnal mixing and discusses the sensitivity of the main ocean metrics to the changes in numerical mixing, as well as comparison of the mixing in the simulations to estimates based on observations. Section 4 describes the sensitivity of atmospheric metrics to numerical mixing in the ocean, and section 5 is a summary and discussion.

2. Model description and experimental design

a. *Model description*

The model we use is HadGEM3-GC5 (Xavier et al. 2026); this will be the physical coupled climate model of the U.K. contribution to CMIP7. The ocean and sea ice components constitute the Global Ocean–Sea Ice configuration, version 9 (GOSI9), described in detail by Guiavarc'h et al. (2025), which was developed jointly between the Met Office and the National Oceanography Centre under the Joint Marine Modelling Programme (JMMP). GOSI9 is based on version 4.0.4 of NEMO v4.0.4 (Madec et al. 2017). The sea ice model is the Sea Ice Integrated Initiative (SI^3), which is fully integrated into the NEMO model code from version 4.0 onward. The extended ORCA grid at 0.25-degree resolution (eORCA025) ocean grid, which is a tripolar C grid with 1/4° horizontal resolution (cell size about 25 km at the equator) and 75 levels in the vertical ranging from 1 m at the surface to 200 m at depth, is used for both ocean and sea ice.

The default vertical coordinate in NEMO v4 is the nonlinear free surface or variable volume layer (VVL) scheme (Adcroft and Campin 2004; Levier et al. 2007), referred to in shorthand as z^*. This represents external gravity waves as changes to the layer vertical scale e_3 scaled uniformly over the water column with the surface height anomaly. A known potential disadvantage of this coordinate, common to all geopotential coordinate systems, is that vertical motions due to internal waves and eddies cause vertical advection of tracers, and the numerical advection scheme has a small irreversible, or diffusive, component (Griffies et al. 2000) which leads to unphysical numerical tracer mixing. In this study, we shall investigate the $z\sim$ scheme of Leclair and Madec (2011), in which the coordinate surfaces are allowed to flex with internal waves and tides on time scales that are short compared with a specified filter time scale, which is supplied as a parameter in the model namelist at runtime. The $z\sim$ was demonstrated by Megann et al. (2022) and Megann (2024) to reduce numerical mixing from near-inertial internal waves and internal tides, respectively, in 1/4° global NEMO simulations forced with CORE2 surface fields, and the reduction increased markedly as the filter time scale was lengthened from 5 to 30 days. This study includes simulations with z^* and with $z\sim$; $\tau_\sim$, the filter time scale for frequency division between the quasi-isopycnal regime at high frequencies and the z^* regime at low frequencies, is set to 20 days, and τ_z, the time scale for restoration of the interface depths to z^*, is set to 60 days [see appendix A of Megann (2024), for a full description of the $z\sim$ scheme as used in this configuration].

The Thermodynamic Equation Of Seawater-2010 (TEOS10) equation of state (Roquet et al. 2015; Feistel 2018) is used in GOSI9, so the state variables are Conservative Temperature Θ and Absolute Salinity S. The vertical mixing scheme for the ocean in all the experiments is a modified version of the Gaspar et al. (1990) turbulent kinetic energy (TKE) scheme (Madec et al. 2017), and the parameterization of Simmons et al. (2004) for tidal mixing is applied. Tracer advection is performed using the total variance diminishing (TVD), also known as flux-corrected transport (FCT), scheme of Zalesak (1979): The default for GOSI9 is for fourth-order tracer advection in both horizontal and vertical directions. In most of the experiments, bi-Laplacian viscosity is used, with the viscosity coefficient reducing poleward from -1.5×10^{-11} m^4 s^{-1} at the equator with the cube of the grid scale to avoid instabilities, but in two experiments, bi-Laplacian Smagorinsky viscosity (Smagorinsky 1963; Griffies and Hallberg 2000) is used. Lateral mixing is along local neutral surfaces and uses the scheme of Redi (1982) and Cox (1987) with a Laplacian diffusivity set equal to a diffusion velocity *rn_ud* multiplied by the local grid scale: In all the simulations, *rn_ud* is 0.011 m s^{-1}, equivalent to a diffusivity at the equator of 300 m^2 s^{-1}.

GOSI9 and HadGEM3-GC5 include the Southern Ocean package (Guiavarc'h et al. 2025; Storkey et al. 2025), which consists of a weak parameterization of eddy-induced advection (Gent and McWilliams 1990; Tréguier et al. 1997) with the coefficient scaled with the cell size and with an upper limit capped at 75 m^2 s^{-1}, a change in the tracer advection from second to fourth order, and partial-slip lateral boundary conditions along the Antarctic coastline. This was found to improve the realism of the simulation of the Southern Ocean in the HadGEM3-GC3.1 coupled configuration: Specifically, the strength of the transport through the Drake Passage and the representation of water masses around the Antarctic coast were substantially more realistic than with the default configuration.

HadGEM3-GC5 is described in Xavier et al. (2026) and is composed of the GOSI9 ocean and sea ice model, coupled to the third Met Office Unified Model–JULES Regional Atmosphere and Land configuration (RAL3; Bush et al. 2025). In this ensemble, we use a relatively low-resolution atmosphere with N96 horizontal resolution, giving a cell size of about 135 km at midlatitudes. This is a coarser resolution than that of HadGEM3-GC5 eORCA025-N216, as described by Xavier et al. (2026),

Brought to you by NATIONAL OCEANOGRAPHY CENTRE | Unauthenticated | Downloaded 11/11/26 12:54 PM UTC

TABLE 1. Model integrations in the ensemble.

Experiment	Vertical coordination	Viscosity	Horizontal tracer advection	Vertical tracer advection
zstar_control	z^*	Fixed $A_{\mathrm{Lm}} = 1.5 \times 10^{11}$	4	4
ztilde_control	$z\sim$	Fixed $A_{\mathrm{Lm}} = 1.5 \times 10^{11}$	4	4
zstar_FCT2H	z^*	Fixed $A_{\mathrm{Lm}} = 1.5 \times 10^{11}$	2	4
zstar_FCT2V	z^*	Fixed $A_{\mathrm{Lm}} = 1.5 \times 10^{11}$	4	2
zstar_Smag	z^*	Smagorinsky $C_{\mathrm{MSmag}} = 3.0$	4	4
ztilde_Smag	$z\sim$	Smagorinsky $C_{\mathrm{MSmag}} = 3.0$	4	4
zstar_visc_x1.8	z^*	Fixed $A_{\mathrm{Lm}} = 2.7 \times 10^{11}$	4	4

although the ocean resolution is identical: This was chosen to allow efficient throughput and to place fewer demands on storage, while still allowing fully interactive surface interactions between ocean and atmosphere. The atmospheric forcing settings are the same as those used by Xavier et al. (2026): namely, constant 2000 greenhouse gas components.

b. Experimental design

The ensemble consists of seven members (see Table 1 for details), with identical atmosphere and sea ice configurations and using the same initial state for all components. One member, *zstar_control*, is the control for the ensemble, using the standard GC5 settings for each of the ocean, sea ice, and atmosphere, while the other seven have a set of changes applied to the NEMO namelist, either singly or two at once, each of which has previously been found to affect numerical mixing. The *ztilde_control* is identical to *zstar_control* but with the vertical coordinate changed from z^* to $z\sim$. In two simulations, *zstar_FCT2H* and *zstar_FCT2V*, the tracer advection is degraded from fourth order (as used in all other ensemble members) to second order independently in the horizontal and the vertical directions, respectively. It has been shown (Ilıcak et al. 2012; Megann and Storkey 2021) that poorly resolved mesoscale features are a significant mechanism for numerical mixing in forced ocean models and that increasing the viscosity is a practical remedy for this; following Megann and Storkey (2021), we modify the viscosity by either changing to the Smagorinsky-deformation-dependent formulation, which locally enhances the viscosity applied to regimes with large shear, or by scaling up the fixed viscosity of the default scheme by a constant ratio. Two simulations, *zstar_Smag* and *ztilde_Smag*, use a bi-Laplacian Smagorinsky deformation-dependent viscosity with the coefficient $C_{\mathrm{MS}} = 3$ and use the z^* and $z\sim$ vertical coordinate, respectively. We note that the Smagorinsky scheme in NEMO v4.0.4 imposes minimum and maximum values for the viscosity that are applied according to the cell Reynolds number; the scale factors for minimum and maximum viscosities are both set to unity (see Madec et al. 2017, for further details). Finally, in *zstar_visc_x1.8*, the fixed viscosity velocity parameter is increased by a factor of 1.8; further increase of this parameter in the present configuration proved to be unstable, even at the reduced time step of 900 s used for this integration.

The ocean is started from rest, with temperature and salinity taken from a mean of years 1995–2014 of the EN4 objective analysis dataset (Good et al. 2013), as described by Guiavarc'h et al. (2025), and the atmosphere is started from a restart file generated by an equivalent atmosphere-only spinup climate simulation. The experiments are integrated for 60 years. The default ocean time step in GOSI9 and HadGEM3-GC5 at $1/4°$ ocean resolution is 1800 s, which is 50% longer than that used in the previous U.K. Global Ocean configuration version 6 (GO6) (Storkey et al. 2018). This is possible because of the improvement in stability resulting from the use of an adaptive-implicit vertical advection scheme (Shchepetkin 2015). However, some of the changes investigated here, specifically the $z\sim$ coordinate and the two altered viscosity settings, proved not to be stable with the longer time step, so to avoid adding uncertainty related to possible sensitivity to the choice of time step, all but one of the simulations described here were integrated with a 1200-s time step. The single exception is the simulation combining $z\sim$ and Smagorinsky viscosity, for which a further reduction in the time step to 900 s was found to be required for stability.

The simulations were integrated on the MONSooN2 collaborative system at the Met Office and used 396 processors for the ocean and sea ice and 288 processors for the atmosphere. For a given configuration, the computational performance, expressed in hours per model year, is determined both by the speed of integration for the given selection of numerical settings and by the choice of time step made on the basis of stability. An added complication is that the atmosphere component is more demanding of processing resources than the ocean, so a modest slowdown in the ocean does not necessarily lead to a commensurate speed reduction in the coupled model when integrated on the same number of processors. Table 2 shows the run speed for a selection of parameter and time step combinations. Shortening the ocean time step by 33% from the default 1800 to 1200 s only slowed the integration by about 6%; however, changing from z^* to $z\sim$ at a 1200-s time step increased run time by 20%, consistent with previous studies using a forced $1/4°$ NEMO ocean (e.g., Megann et al. 2022). It was found that increasing the order of the tracer advection scheme at a given choice of time step had a negligible effect on run speed. Although direct comparisons were not made between integrations of the present configuration with fixed viscosity and with the Smagorinsky scheme at the same time step, Megann and Storkey (2021) found that this change had little effect on the run time.

For convenience, and for consistency with the forced GOSI9 (Guiavarc'h et al. 2025), which was started from the same initial state, we define the integration period as starting

Brought to you by NATIONAL OCEANOGRAPHY CENTRE | Unauthenticated | Downloaded 11/11/26 12:54 PM UTC

TABLE 2. Model run speeds.

Vertical coordination	Ocean time step (s)	Viscosity	Horizontal tracer advection	Vertical tracer advection	Hours per model year
z^*	1800	Fixed	Fourth	Fourth	9.85
z^*	1200	Fixed	Fourth	Fourth	10.5
$z\sim$	1200	Fixed	Fourth	Fourth	12.7
$z\sim$	900	Smagorinsky	Fourth	Fourth	16.5

from 1976. We shall present time series of annual means over the 60 years of each simulation; for longer time averages, we treat the first 20 years as a spinup, and averages will be evaluated over the 40 years following year 20 of each integration. The exception is the mixing analysis presented in the next section, which is based on 5-day means over years 21–30; this is the same period as used in previous work using this approach (Megann 2018; Megann and Storkey 2021; Megann et al. 2022; Megann 2024) and was originally settled on as a compromise between sufficient length of time for averaging and the requirement for large disk space.

3. Ocean sensitivities

a. *Mixing diagnostics: Definitions*

We shall use the word "mixing" in this paper to refer exclusively to the diffusive mixing of tracers, either explicit or numerical, rather than to the viscous dissipation of momentum. We base our analysis on the density transformation framework of Walin (1982). Following Lee et al. (2002), we define a transformation streamfunction $G(\phi, \rho)$ as follows:

$$G(\phi, \rho) = \Psi(\phi, \rho) + \frac{\partial}{\partial t} V(\phi, \rho), \tag{1}$$

where $\Psi(\phi, \rho)$ is the overturning streamfunction at latitude ϕ and potential density ρ and $V(\phi)$ is the volume below the isopycnal surface ρ and south of ϕ. The zonally integrated diapycnal velocity $g(\phi, \rho)$ through the density surface ρ at this latitude is then given by the meridional gradient of $G(\phi, \rho)$:

$$\int_{x_E}^{x_W} g(\phi, \rho) dx = -\frac{\partial}{\partial y} G(\phi, \rho). \tag{2}$$

The density flux $D(\phi, \rho)$ across surface ρ is then the integral of the diapycnal velocity from the densest class $\rho_{\max}$ up to ρ:

$$D(\phi, \rho) = \int_{\rho}^{\rho_{\max}} d\rho \int_{x_E}^{x_W} g(\phi, \rho) dx. \tag{3}$$

As discussed in detail in the appendix of Megann (2024), this analysis involves several approximations. As discussed by Lee et al. (2002), it neglects other contributions to the evolution of density besides dianeutral diffusion, principally the consequences of the nonlinear equation of state (cabbeling and thermobaricity) and isoneutral diffusion. Ideally Eq. (3) should be evaluated during runtime, but the NEMO code does not permit this at present. For these integrations, the highest output frequency is 5-day means for tracer and velocity fields, so the calculation of the streamfunction $\Psi(\phi, \rho)$ omits the advective effects of processes acting on shorter time scales than this, notably internal waves and internal tides. Megann (2024) demonstrated that the consequences of this choice of time sampling manifest primarily as terms that are oscillatory in time and space and that their contributions to the diagnostic metric average to a relatively small value compared to the mean. Finally, the assumption in this analysis that density only changes as a result of diffusive processes is certainly invalid in near-surface waters that are seasonally ventilated, so the analysis is only applied to regions where the density is higher than the maximum surface density at a given latitude. As in the previously mentioned studies, we evaluate the fluxes from 5-day means over the 10-yr period 1996–2005, and we include the bolus velocity from the Gent and McWilliams parameterization.

Previous work based on this approach (starting from Lee et al. 2002) has evaluated an effective diffusivity $\kappa_{\text{eff}}(\phi, \rho)$ by dividing the meridional derivative of $G(\phi, \rho)$ by the zonal integral of the vertical density gradient. While this provides a metric that is meaningful to compare between ensemble members, the calculation of the zonal-mean diffusivity relies on the assumption that the rate of diapycnal density transformation varies in the same way with longitude as the vertical density gradient integrated across the same zonal strip. Although this assumption is likely to give errors that are more or less consistent across an ensemble such as the one used in the present study, it renders quantitative comparisons between the effective diffusivity and the explicit diffusivity κ_{exp} used in the mixing scheme unreliable and is also problematic to verify directly against observational estimates of the diffusivity. For this reason, we shall in this paper instead use area integrals of the density flux $D(\phi, \rho)$ over the near-global ocean as a measure of the total diapycnal upwelling through each density surface in the respective basin, which does not rely on this approximation.

To illustrate the relative strength of numerical and explicit mixing, we shall compare the area integrals of the total diapycnal flux $D(\phi, \rho)$ with the density flux $D_{\text{exp}}(\phi, \rho)$ caused by the explicit mixing in the model, where

$$D_{\text{exp}}(\phi, \rho) = -\kappa_{\text{exp}} \frac{\partial \rho}{\partial z}, \tag{4}$$

and κ_{exp} is the explicit diffusivity used in the model mixing scheme. The $D_{\text{exp}}(\phi, \rho)$ is evaluated as a zonal mean at each latitude and density class. In this case, we evaluate the mean of the product of the diffusivity and the density gradient during runtime over a year of integration; if the effects of the nonlinearity of the density of state can be neglected, this will be close to the diffusive density flux that occurs as a result of the

Brought to you by NATIONAL OCEANOGRAPHY CENTRE | Unauthenticated | Downloaded 11/11/26 12:54 PM UTC

explicit mixing scheme. We mask out the directly ventilated regions according to the surface density as with the mixing estimate from Eq. (3), in exactly the same way as used for the calculation of the diagnosed total density flux $D(\phi, \rho)$. The difference between the diagnosed total mixing $D(\phi, \rho)$ and the explicit mixing $D_{\text{exp}}(\phi, \rho)$ is then approximately equal to the spurious numerical contribution to tracer mixing. Because $D(\phi, \rho)$ represents the total density flux from all sources, the residual includes not only the numerical mixing but also the contributions from cabbeling and thermobaricity, as well as the errors resulting from the approximations listed above. This analysis does not allow these latter mechanisms for density transformation to be distinguished from numerical mixing, but since they are likely to be similar across the model ensemble, we shall ascribe differences in $D(\phi, \rho)$ or in the residual between model simulations as arising from changes in numerical tracer mixing.

b. *Mixing diagnostics: Results*

The top three panels (Figs. 1a–c) show the zonally integrated density fluxes D_{exp} from explicit mixing, as defined in Eq. (4), in the Atlantic and Indo-Pacific basins; the second row (Figs. 1d,e, and f) shows the diagnosed fluxes $D(\phi, \rho)$ as defined in Eq. (3) in the same basins; and third row (Figs. 1g,h, and i) shows the difference between the diagnosed and explicit density fluxes. The units here are kilograms per meter per second, and positive values indicate upward (downgradient) fluxes of density. Contours of the overturning streamfunction in the respective basin are overlain in black. The bold black dashed lines illustrate the maximum surface density in that latitude class in the respective basin, evaluated from monthly means over the 10-yr analysis period; the directly ventilated water masses with densities lighter than this, in which Eqs. (2) and (3) are certainly invalid, are masked out in white. The white areas in the interior correspond to downward (upgradient) density fluxes that are most likely due to cabbeling or thermobaricity. We use the diagnosed density flux integrated over the global basin as a measure of the total mixing, which is the sum of the explicit mixing and the spurious mixing, the latter arising predominantly from numerical mixing, along with contributions arising from the approximations discussed above. The difference between the diagnosed total mixing and the explicit mixing (Figs. 1g,h, and i) may, therefore, be interpreted as the numerical component of the mixing, providing the consequences of the approximations in the derivations of the diagnosed total mixing are small.

We note, first of all, that the mixing, both explicit and numerical, is dominated by the Indo-Pacific: In contrast to the effective diffusivity analysis used by Lee et al. (2002) and Megann (2018), the upwelling here is area integrated, so the Atlantic has a much smaller contribution to the global total. Second, the density flux from numerical mixing is substantially larger than that from the explicit mixing over most of the density range, so that the residual mixing shown in Figs. 1g–i is only slightly smaller than the total, consistent with the results of Megann (2018). The spatial patterns of the explicit and diagnosed mixing fluxes are different: The former is mainly concentrated in intermediate waters with densities σ_2 lighter than about 36.3 kg m^{-3}, while the diagnosed mixing is substantially larger and has significant magnitude down to the deep waters with σ_2 greater than 36.7 kg m^{-3}. The large upward density flux with magnitude greater than 2 kg m^{-1} s^{-1} in the global means (Figs. 1d,g) south of 50°S is associated with progressive slumping of the isopycnals in the Southern Ocean: At 50°S, the 37.10 kg m^{-3} isopycnal displaces downward by about 400 m over the 10-yr analysis period, which we shall show in section 3c corresponds to a weakening of the transport through the Drake Passage by about 5 Sv over the same period.

In Fig. 1j, we show the density fluxes integrated in density classes over the global ocean, expressed in kilograms per second: The thick black line shows the integrated explicit density flux. Figure 1k shows the difference between the globally integrated density flux between each ensemble member and the explicit flux, which we interpret as the strength of the numerical mixing, and finally, Fig. 1l shows the differences from the diagnosed density flux in each ensemble member from that in *zstar_control*. The density flux from numerical mixing (Fig. 1k) is relatively small at densities lower than 35.0 kg m^{-3} and in bottom waters with $\sigma_2 > 37.0$ kg m^{-3} and peaks at a density of 36.3 kg m^{-3}, where it is larger than the explicit mixing by a factor of between 2 and 3.

Despite the approximations mentioned above in the derivation of the diapycnal upwelling rate, it is meaningful to compare the diagnosed total mixing in the simulations with observational estimates, and here, we focus specifically on *zstar_control*. For this, we shall refer to the overturning in density space represented by the streamfunction contours shown in black in Figs. 1a–c. Equation (3) implies that upwelling of streamlines in density space is directly related to the density flux in regions where this is large compared with that from the vertical drift of isopycnals, which is generally the case in the present simulations away from the Southern Ocean. We shall compare mainly with the streamfunctions constructed by LS2007 by applying an inverse analysis to hydrographic sections and with those by C2019 from the ECCO4 reanalysis. We note that the closed cell in the Indo-Pacific basin (also seen in the global streamfunction) between 40° and 10°S (Figs. 1a,c) is not seen in comparable forced simulations; the latitude and density range are consistent with a largely surface-forced density transformation of around 1 Sv along a circuit around the Australian continent. The AABW cell in Fig. 1a extends into the Northern Hemisphere, mainly in the Indo-Pacific, whereas less than 1 Sv penetrates into the Atlantic; this is consistent with the analysis of C2019, who presented a similar partition of AABW between the ocean basins, whereas in LS2007, about 5 Sv of AABW enters the Atlantic. In the present model, the AABW cell has a maximum of 12 Sv between 34° and 20°S and is consumed by mixing by about 20°N, again quantitatively similar to the corresponding cell described by C2019. The pinch off of the AABW cell at about 70°S, resulting in the separation of the downwelling cell in the Southern Ocean and the cell north of the ACC, is likely to be due to cabbeling and thermobaricity, which have been shown to be associated with upward diapycnal velocities (corresponding to apparent upgradient density fluxes) in this region of over 10 mm day^{-1} (Klocker and McDougall 2010). Figure 1j suggests that in the density range σ_2 = 36.8–37.0 kg m^{-3}, characterizing the deep waters

Brought to you by NATIONAL OCEANOGRAPHY CENTRE | Unauthenticated | Downloaded 11/11/26 12:54 PM UTC

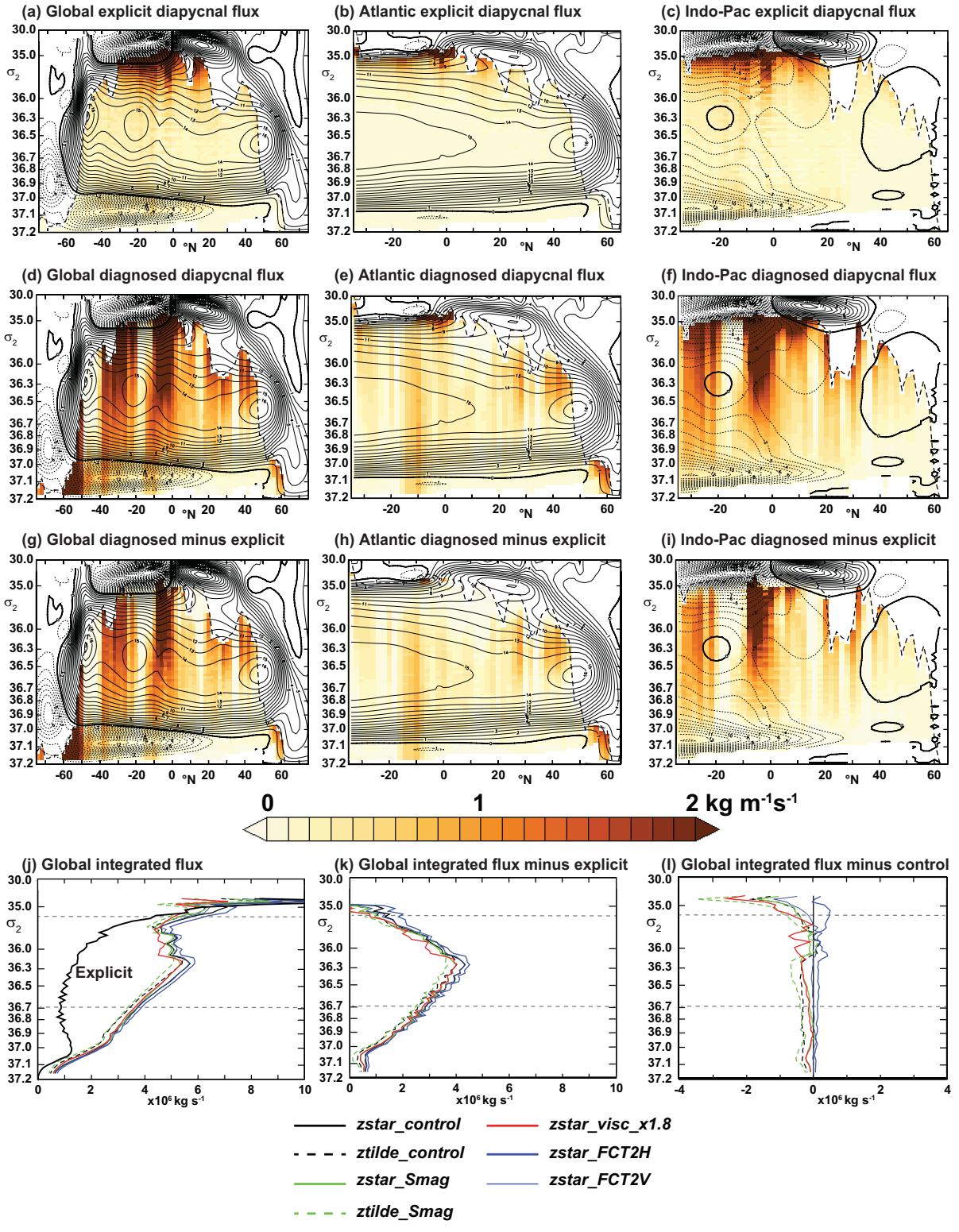


FIG. 1. Zonally integrated density fluxes (kg m^{-1} s^{-1}) from explicit mixing in density classes in *zstar_control* in (a) the global domain, (b) the Atlantic, and (c) the Indo-Pacific; zonally integrated diagnosed density fluxes in *zstar_control* in (d) the global domain, (e) the Atlantic, and (f) the Indo-Pacific; differences of the diagnosed and explicit density fluxes in *zstar_control* in (g) the global domain, (h) the Atlantic, and (i) the Indo-Pacific. (j) The diagnosed density flux in density classes, integrated over the global ocean (kg s^{-1}) in all the simulations, with that from explicit mixing shown as the bold black line; (k) the difference between explicit and diagnosed density fluxes integrated over the global ocean in all the ensemble members; and (l) the difference between the globally integrated diagnosed density flux in each ensemble member and that in *zstar_control*. The vertical axis in all panels is potential density σ_2 (kg m^{-3}). The dashed gray lines in (j)–(l) show the limits of the region in density space used to define the scalar metric D_{IDW}.

Brought to you by NATIONAL OCEANOGRAPHY CENTRE | Unauthenticated | Downloaded 11/11/26 12:54 PM UTC

TABLE 3. Integrated diapycnal fluxes, averaged over the density range 35.2–36.7 kg m^{-3}, and their ratio to the value for *zstar_control*.

	zstar_control	*ztilde_control*	*zstar_FCT2H*	*zstar_FCT2V*	*zstar_Smag*	*ztilde_Smag*	*zstar_visc_x1.8*
Flux ($\times 10^6$ kg s^{-1})	4.703	4.488	4.891	4.732	4.519	4.396	4.542
Ratio to control	0%	−5%	+4%	+0.6%	−4%	−6.5%	−3.4%

overlying AABW, there is 3–4 Sv of density transformation, which is consistent with the rising contours in Fig. 1a and also with the approximately 4 Sv of diabatic NADW upwelling shown in both Fig. 3a of LS2007 and in Fig. 2a of C2019. The bulk of the southward-flowing NADW in the model, however, is at a lower density than in the observational analysis: In *zstar_control*, it lies between 36.9 and 37.1 kg m^{-3}, in contrast to between 37.20 and 37.50 kg m^{-3} in C2019, consistent with enhanced mixing with the overlying waters in the model. The overturning streamfunction in the Indo-Pacific in *zstar_control* (Fig. 1c) may be compared with the corresponding streamfunctions presented in Fig. 3c of LS2007 and in Fig. 2b of C2019: The abyssal upwelling cell has similar northward extent in all three analyses, although its southward return flow as North Pacific Deep Water in the model is at a lower density than in C2019's analysis, being centered at 36.3 kg m^{-3} in *zstar_control* and at 36.8 m^{-3} in ECCO4. This is again consistent with the model having an excess of mixing of the deep water with the overlying intermediate waters.

Figure 1l shows the differences between the global integral of the density flux $D(\phi, \rho)$ in each of the simulations from that in *zstar_control* and confirms that this proxy for the total mixing is a useful discriminator between the experiments and that the hierarchy of profiles is consistent with expectations based on previously published work using forced ¼° global NEMO configurations. The experiment *zstar_FCT2H* with second-order horizontal tracer advection (thick blue line) has significantly higher mixing in almost all density classes than the control, while *zstar_FCT2V* with second-order vertical tracer advection (thin blue line) is not reliably distinguishable from the control. The experiments with modified viscosity all cluster to the left of zero, indicating reduced numerical tracer mixing: The experiment with the combination of Smagorinsky viscosity and $z\sim$ (dashed green line) shows the largest reduction.

We shall define, with reference to Fig. 1j, a metric D_{IDW} for total mixing, derived from the global density flux, as the area integral of $D(\phi, \rho)$ between 40°S and 40°N, averaged over densities corresponding approximately to intermediate and deep waters, with potential densities in the range $35.5 < \sigma_2 < 36.8$: These limits are demarcated by the dashed gray lines in Figs. 1j–l. Table 3 lists the values of D_{IDW} for each simulation, weighted to exclude ventilated density classes, and the second row lists the percentage change with respect to that in *zstar_control*. This confirms quantitatively that *ztilde_Smag* has levels of numerical mixing substantially (6%–8%) weaker than the control; *ztilde_control* and *zstar_Smag* and *zstar_visc_x1.8* have small but consistent (3%–5%) reductions in mixing; *zstar_FCT2V* has numerical mixing close to that in the control; and *zstar_FCT2H* has about 4% stronger mixing than in the control.

c. Ocean and sea ice biases and sensitivities

The analysis period we shall use here is the 40 years 1996–2035, which allows for a 20-yr spinup adjustment. Although a longer spinup period might be preferred, we shall show in this section that the time evolution of global-mean metrics after the first decade is generally dominated by secular drifts, rather than spinup transients. In any case, this is consistent with the approach used in the published results for the GO5.0 (Megann et al. 2014), GO6 (Storkey et al. 2018), and GOSI9 forced configurations, in which the third decade of integration is used for analysis after a 20-yr spinup period; Xavier et al. (2026) also use a 20-yr spinup for the GC5 reference integration, followed by a 40-yr analysis period. Since the model used is initialized from a 1995 to 2014 mean climatology and uses year 2000 control atmospheric composition, it is meaningful to define biases with reference to present-day climatological datasets. We note as a caveat that the ocean component in the control configuration has undergone lengthy tuning in both forced and coupled modes to reduce biases, in particular those of surface temperature, so it is highly likely that there are compensating errors in the model; in addition, the positive surface heat flux imbalance implied by year-2000 atmospheric forcing will result in warming with respect to present-day climatology. For these reasons, value judgments of individual simulations in the present ensemble on the basis of their surface biases are of limited usefulness. In this and the following section, we present an overview of the sensitivities of the major ocean and sea ice metrics to mixing; in section 3d, we shall present a more quantitative analysis, based on correlations between the time-averaged metrics and the global mixing metric D_{IDW} defined in section 3a.

In Fig. 2a, we show the sea surface temperature (SST) bias in *zstar_control*, evaluated over the 40-yr analysis period, with respect to the ESA Climate Change Initiative (CCI) climatology (Merchant et al. 2014), averaged over 1991–2010. The control can be seen to have a generally warm surface bias poleward of 45°N and about 45°S, as well as along the eastern boundaries of the Pacific and South Atlantic, which exceeds 2 K in places, and a generally cool bias between 45° and 45°S of a similar magnitude. It is interesting to compare the biases in the GC5 coupled control simulation with those in the CORE2-forced GOSI9 configuration with the same ocean component (see Fig. 8c of Guiavarc'h et al. 2025): While both simulations share the warm bias in the Southern Ocean and the North Pacific, the tropics and subtropics in the coupled control are up to 1 K cooler than the climatology, in contrast to the forced simulation where the biases are overall neutral in this latitude range. The other panels in Fig. 2 show the difference in surface temperature, evaluated over the same period, with respect to that in *zstar_control*; the control bias panel and the difference panels

Brought to you by NATIONAL OCEANOGRAPHY CENTRE | Unauthenticated | Downloaded 11/11/26 12:54 PM UTC

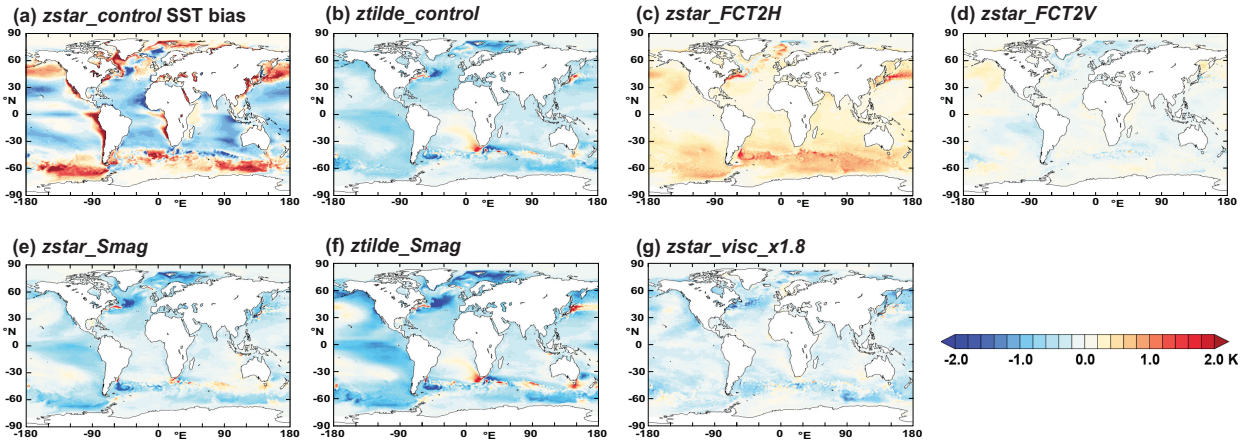


FIG. 2. (a) The SST bias (K) in *zstar_control* with reference to the ESA CCI climatology, averaged over 1991–2010, and (b)–(g) global differences in SST (K), averaged over 1996–2035, between each ensemble member and that in *zstar_control*.

have the same color scale. The four simulations with numerical mixing reduced relative to the control, namely, *ztilde_control*, *zstar_Smag*, *ztilde_Smag*, and *zstar_visc_x1.8*, have a surface temperature that is almost everywhere between 0.5 and 1.5 K cooler than in the control, whereas in *zstar_FCT2H*, the SST is warmer than the control by a similar amount. The largest changes of either sign occur consistently in the Gulf Stream separation region and in the South Atlantic and Indian Oceans north of the Antarctic Circumpolar Current (ACC), whereas in the simulation with second-order vertical tracer advection, *zstar_FCT2V*, the SST is almost unchanged relative to the control. Although the large-scale SST changes resulting from using $z\sim$ or from changing the viscosity are similar, we note the warming "bull's-eye" of about 0.5 K west of South Africa in the two simulations with $z\sim$ (Fig. 2b and f) that is not present in the simulations using z^*; the reason for this has not yet been investigated.

Figure 3a shows the sea surface salinity (SSS) bias in *zstar_control* with respect to the EN4 climatology (Good et al. 2013); in this case, the reference is the 1996–2014 mean field used for model initialization. The control simulation has a salty bias of over 2.0 g kg^{-1} in the Canadian and Siberian sectors of the Arctic and a comparable fresh bias in the Atlantic sector; away from the Arctic, the control is saltier in the South Atlantic than the climatology and fresher in the Pacific and the Southern Ocean. This pattern of surface salinity biases is very similar to that seen in the GOSI9 forced model (see Fig. 10 of Guiavarc'h et al. 2025). The other panels in Fig. 3 show the SSS difference with respect to that in *zstar_control*; as in Fig. 2, all the panels have the same color scale. Over much of the ocean, the SSS changes in *ztilde_control*, *zstar_Smag*, and *ztilde_Smag* with respect to *zstar_control* [panels (b), (e), and (f)] are of opposite sign to the biases in the control: More specifically, the fresh surface biases of up to 0.5 g kg^{-1} in the central Pacific and south of 30°S in the South Atlantic are opposed by salinity increases in the same regions, albeit of smaller magnitude than the biases, while the salty surface bias in the tropical Atlantic and Indian Oceans corresponds to a freshening in

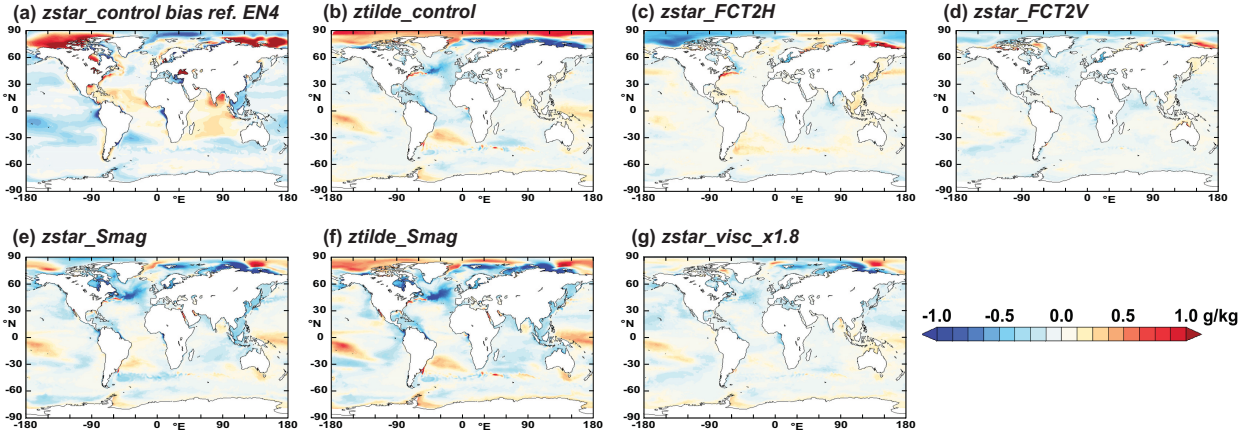


FIG. 3. (a) The SSS bias (g kg^{-1}) in *zstar_control* with reference to the EN4 climatology, averaged over 1996–2014, and (b)–(g) global differences in SSS (g kg^{-1}), averaged over 1996–2035, between each ensemble member and that in *zstar_control*.

Brought to you by NATIONAL OCEANOGRAPHY CENTRE | Unauthenticated | Downloaded 11/11/26 12:54 PM UTC

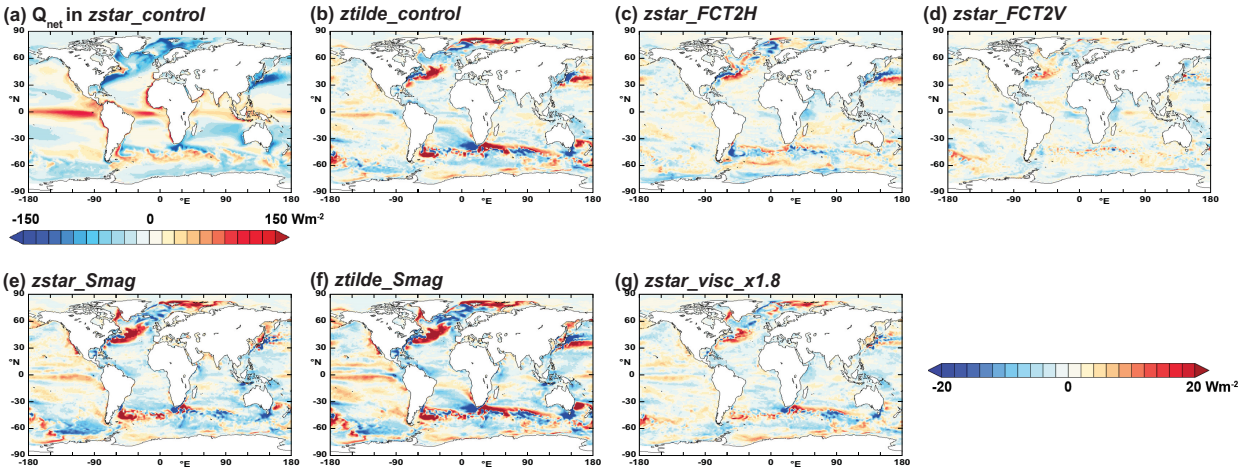


FIG. 4. (a) Net heat flux into the ocean (W m^{-2}), averaged over 1996–2035 in *zstar_control*, and (b)–(g) differences in net heat flux between each ensemble member and that in *zstar_control*. Note that (a) has a different scale to the other panels.

these simulations. In the two simulations with $z\sim$ [panels (b) and (f)], much of the Arctic is more than 0.5 g kg^{-1} saltier than the control, and this change again mainly opposes the biases in the control. The smallest overall SSS sensitivity occurs in the simulation with the vertical tracer advection degraded to second order [panel (d)].

Figure 4 shows the sensitivities of the net surface heat flux into the ocean as a difference over the 40-yr analysis period from that in the control. Figure 4a shows the net heat flux in *zstar_control*, while the other panels show the differences of the heat flux in the respective simulation from that in *zstar_control*. A consistent magnitude and spatial pattern of sensitivity is seen in *ztilde_control*, *zstar_Smag*, and *ztilde_Smag* (Figs. 4b,e, and f) in which about 20 W m^{-2} more heat enters the ocean in the Gulf Stream region, south of the Kuroshio separation, around the Atlantic sector of the ACC, and in the Barents Sea, while weaker positive flux anomalies occur in the Pacific and negative anomalies occur in the Atlantic and Indian Oceans. A similar spatial pattern is seen in *zstar_visc_x1.8* (Fig. 4f) but with about half the magnitude seen in the latter three integrations. As with the SST and SSS, the smallest sensitivity is seen in *zstar_FCTV* (Fig. 4d), while—perhaps surprisingly, given the strong and anomalous sensitivity of the surface temperature seen in Fig. 2c—the sensitivity of the surface heat flux in *zstar_FCT2H* (Fig. 2c) is rather similar in many regions to that in the simulations with reduced numerical mixing, although south of 30°S, it indeed generally has an opposite sensitivity to that in Figs. 4b,e,f. Overall, the four experiments characterized by lower numerical mixing have a consistent pattern of local changes in surface heat flux but no clear large-scale change of one sign or the other, which will be confirmed quantitatively in Table 4. We shall show in section 4 that the individual components of the heat flux, most notably the latent heat flux and the shortwave radiation, have stronger sensitivities than that of the net heat flux, and these tend to compensate in the global mean.

Figure 5 shows the temperature changes on depth levels from year 1 to the final year of each of the 60-yr integrations in the global domain and in individual ocean basins, with the horizontal (temperature) axis being the same for each basin. To enable the sensitivities in the upper ocean to be more visible, we only plot down to 2000-m depth, but the drifts below that depth are relatively small and decrease in magnitude with depth. All the panels have the same horizontal axis limits to enable the changes in different basins to be compared consistently; where the lines lie visibly close to one another, the sensitivity in the respective basin and depth is not judged to be significant.

We shall first summarize the drifts in the *zstar_control* simulation (solid black lines). The global-mean tendency (Fig. 5a) is to cool at depths shallower than 500 m, with a maximum of about −0.3 K at 100 m, and warming below 500 m, reaching a maximum of 0.2 K at about 900 m. In the Atlantic (Fig. 5b) and Indian (Fig. 5d) Oceans, the overall tendency is the same as in the global ocean but with stronger cooling in the upper ocean and stronger warming below about 300 m than in the global mean; the maximum warming occurs deeper in the Atlantic than in the Indian Ocean, occurring at about 900 m in the former and at about 600 m in the latter. The Pacific Ocean as whole (Fig. 5c) cools, with a maximum cooling of about 0.8 K at 200 m. The Arctic (Fig. 5e) and Southern Ocean (Fig. 5f) both warm at all depths, with maxima of 0.5 K at 1300 m and of 0.3 K at 800 m, respectively. The rate of globally averaged warming of *zstar_control* at 3000 m is consistent with an effective downward diffusive heat flux of about 1 W m^{-2}, while that at 1000 m corresponds to a downward flux between 1.5 and 2.0 W m^{-2}, both significantly stronger than the net surface heat flux of 0.65 W m^{-2} (see Table 4). This drift with respect to the initial state (a climatological mean from 1996 to 2014) indicates an excess of mixing in the interior.

In the upper 200 m, the simulations with the lowest levels of mixing (*ztilde_control*, *zstar_Smag*, *ztilde_Smag*, and *zstar_visc_1.8*) are the coolest of the ensemble in all basins, being between 0.1 and 0.3 K cooler than *zstar_control* in the global mean. By contrast, *zstar_FCT2H* is warmer everywhere than *zstar_control* in this depth range in every basin, again consistent

Brought to you by NATIONAL OCEANOGRAPHY CENTRE | Unauthenticated | Downloaded 11/11/26 12:54 PM UTC

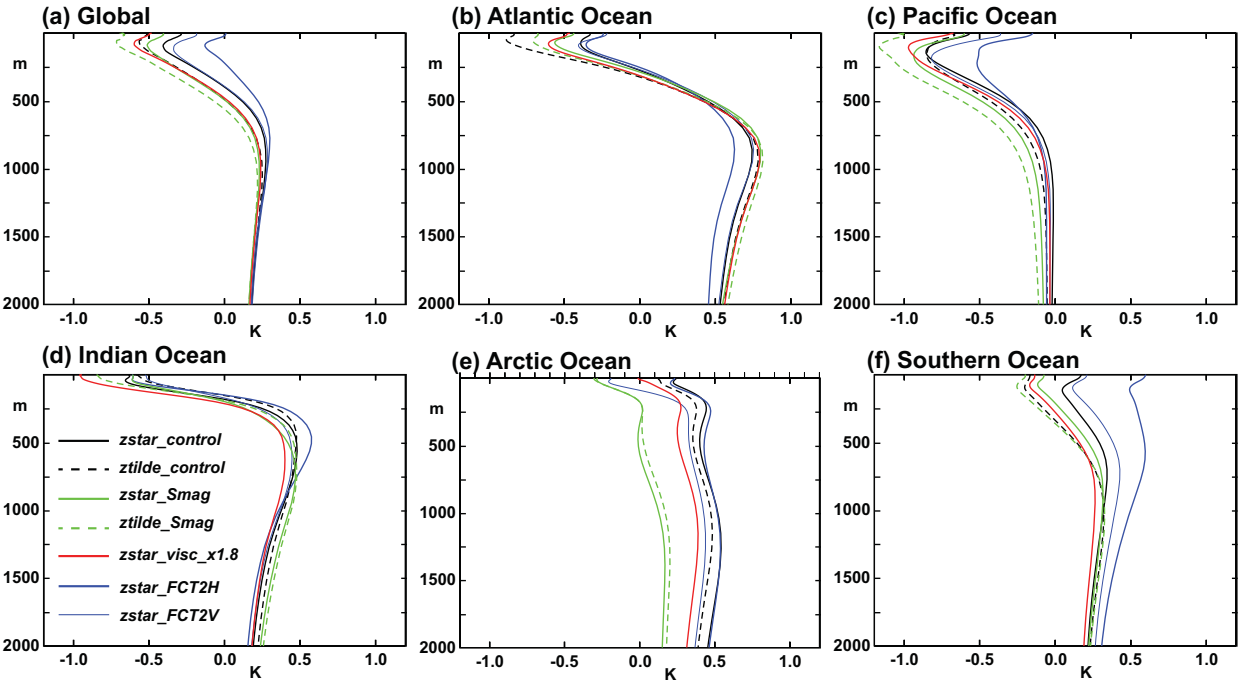


FIG. 5. Basin-mean temperature changes (K) on depth levels from year 1 to year 60 in (a) the global domain, (b) the Atlantic Ocean, (c) the Pacific Ocean, (d) the Indian Ocean, (e) the Arctic Ocean, and (f) the Southern Ocean.

with the surface temperature sensitivity shown in Fig. 2. Below this depth, however, the ordering of the profiles becomes more complex. In the Pacific (Fig. 5c), Arctic (Fig. 5e), and Southern Oceans (Fig. 5f), the simulations with lowest mixing cool most below 500 m, whereas in the Atlantic (Fig. 5b) and Indian Oceans (Fig. 5d), the simulation with strongest mixing (*zstar_FCT2H*) is the coolest below 500 m. In the Arctic (Fig. 5e) and Southern (Fig. 5f) Oceans, there is a larger spread at depths shallower than 1000 m than in the other ocean basins, with the simulations with lower mixing being consistently cooler than the control. The drifts, along with their sensitivities to mixing, partially cancel in the global mean, and as a result, the global-mean changes below 1000 m (Fig. 5a) lie within 0.1 K of that in *zstar_control*, and the sensitivity to mixing is not significant.

Apart from the sensitivity of the surface temperature to numerical mixing and the excess of diffusive downward transfer of heat into the deep ocean discussed above, there are other potential mechanisms for the spatial variation of the temperature drifts shown in Fig. 5. The weakening of the warming at 1000 m in the Atlantic by about 0.2 K as numerical mixing increases (Fig. 5b), along with the opposite tendency in the Arctic (Fig. 5e), is at least qualitatively consistent with the observation that the poleward heat transport at 45°N increases with the strength of numerical mixing by about 10% across the ensemble (see Table 4). The deeper warming in the Southern Ocean is likely to be associated with the slumping of the isopycnals across the ACC mentioned in section 3b that increases with mixing, as the warming increases with the mixing metric D_{IDW}.

Figure 6 shows time series of annual global means of the principal ocean tracer fields for the ensemble. The initial spinup phase, indicated primarily by the settling of the surface temperature (Fig. 6b) and the surface net heat flux (Fig. 6c), lasts between 10 and 20 years, after which the evolution of the volume-mean temperature (Fig. 6a) takes the form of a broadly linear trend with time, while the secular drifts in surface heat (Fig. 6c) and freshwater (Fig. 6f) fluxes are small. The long-term mean of the global-mean net surface heat flux into the ocean is between 0.5 and 0.7 W m^{-2}. The global-mean salinity (Fig. 6d) decreases approximately linearly for the first 45 years, after which it increases slightly; this is consistent with the drop in the surface freshwater flux (Fig. 6f) after that time. In all the simulations, there is a progressive surface warming (Fig. 6b) of about 0.2 K from year 20 until year 50, at which point the surface starts to cool again. Whereas the depth-integrated temperature and salinity fields are similar across the ensemble, as are the surface fluxes, the global-mean surface temperature and salinity show a consistent separation, with those simulations with the lowest numerical mixing having the coolest and freshest surfaces and vice versa. The surface salinity (Fig. 6e) freshens steadily in all runs after year 10 by about 2×10^{-3} g kg^{-1} yr^{-1}: *zstar_Smag* and *ztilde_Smag* have the largest freshening, while *zstar_FCT2H* has the smallest, with a difference of about 0.06 g kg^{-1} at the end of the integration. The surface elevation, as saved by NEMO by default, includes the dynamic height and the time-integrated freshwater flux, but not the steric contributions, so its global mean (Fig. 6g) is essentially the time integral of the total downward surface freshwater flux into the global ocean: The simulations

Brought to you by NATIONAL OCEANOGRAPHY CENTRE | Unauthenticated | Downloaded 11/11/26 12:54 PM UTC

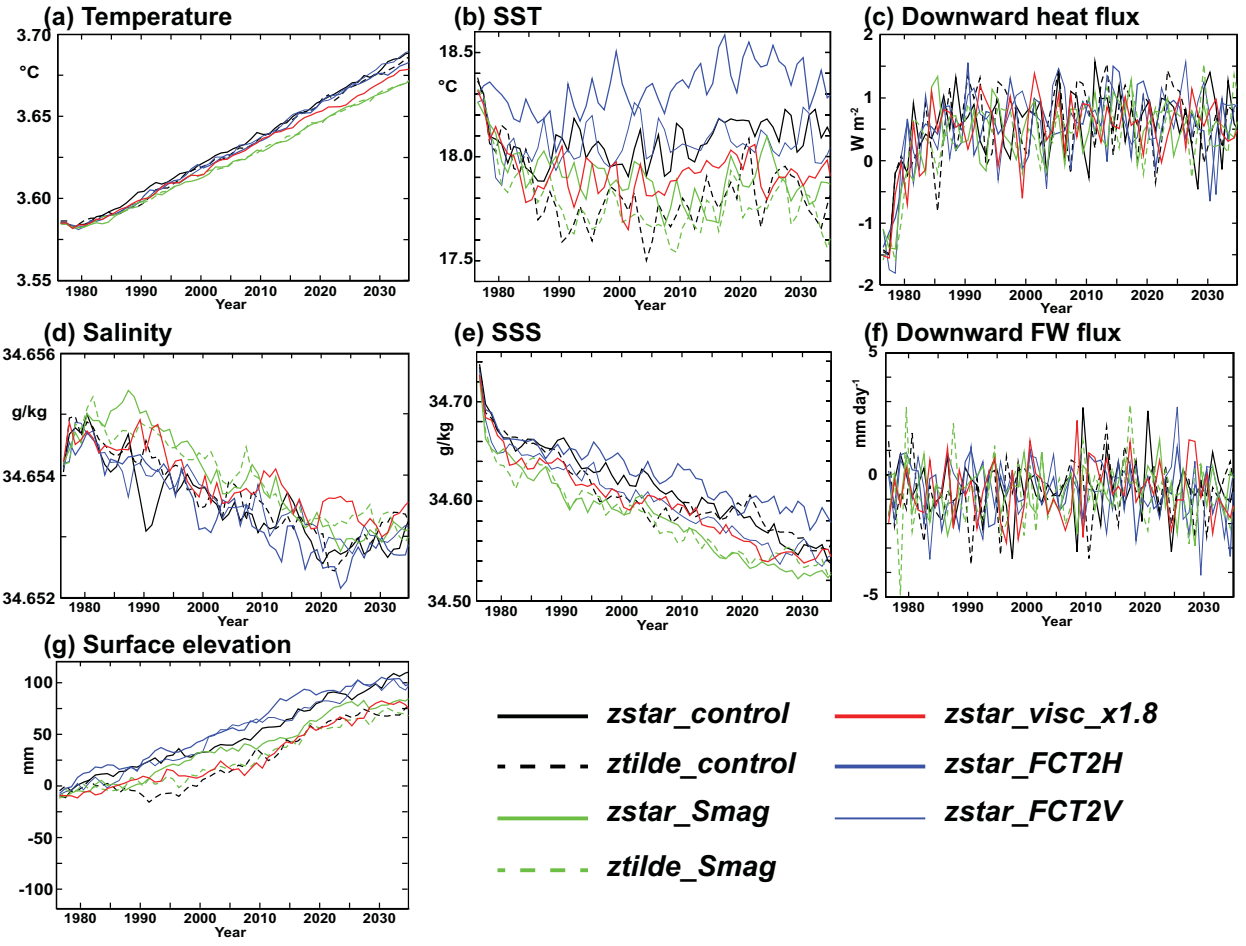


FIG. 6. Annual global-mean ocean quantities for the model integrations: (a) thickness-weighted mean temperature (°C), (b) surface temperature (°C), (c) downward surface heat flux (W m^{-2}), (d) thickness-weighted mean salinity (g kg^{-1}), (e) surface salinity (g kg^{-1}), (f) downward surface freshwater flux (mm day^{-1}), and (g) sea surface elevation (mm).

with the lowest mixing have a consistently lower rate of SSH increase in the first 20 years, but the rates of increase are more similar to one another over the remainder of the integration.

Whereas the individual curves for the volume-integrated and surface temperature and salinity are easy to distinguish from one another, with the interannual variability small relative to the spread of the ensemble, this is not true of the surface fluxes (Figs. 6c and f), and this is consistent with the relatively weak correlations between these indices across the ensemble and the mixing metric D_{IDW} as listed in Table 4.

Figures 5 and 6 demonstrate that significant secular drift occurs in the temperature and salinity fields of all the simulations, regardless of the settings that we have perturbed. This is particularly evident in the Atlantic (Fig. 5b) and the Arctic (Fig. 5e), where warming below 1000 m persists at least to year 60 of the integrations. We note that the comparatively slow net drift in the Pacific (Fig. 5c) is likely to result from local compensations. This large-scale deep warming is typical of NEMO integrations at this resolution (e.g., Megann 2018; Guiavarc'h et al. 2025) and, as we have already noted, is qualitatively consistent with the presence of numerical mixing. Weijer et al. (2020) show that long-term drifts in numerical CMIP6 models from their climatologically derived initial state are associated with changes in the vertical structure of the AMOC, with changes in the vertical structure of the overturning circulation in which the ensemble-mean depth of the AMOC maximum at around 26°N migrates upward from about 1000 m to a mean of 840 m. This corresponds in most cases to a southward-flowing NADW water mass that is too shallow and at a lower density than that in the observed ocean and, as we have discussed in section 3b, the latter is the case in the present model.

Figure 7 shows the hemispheric sea ice area in March and September in each year of the integrations. Little overall trend can be seen in either hemisphere and in either season, and there is more interannual variability in the hemispheric summer ice area (Figs. 7b and c) than in winter (Figs. 7a and d). In general, *ztilde_control* (black dashed curve) and the two simulations with Smagorinsky viscosity (green curves) have more sea ice than the control, while *zstar_FCT2H* (solid blue lines) has the least. The spread across the ensemble is larger in summer than in winter. Overall, those simulations with the

Brought to you by NATIONAL OCEANOGRAPHY CENTRE | Unauthenticated | Downloaded 11/11/26 12:54 PM UTC

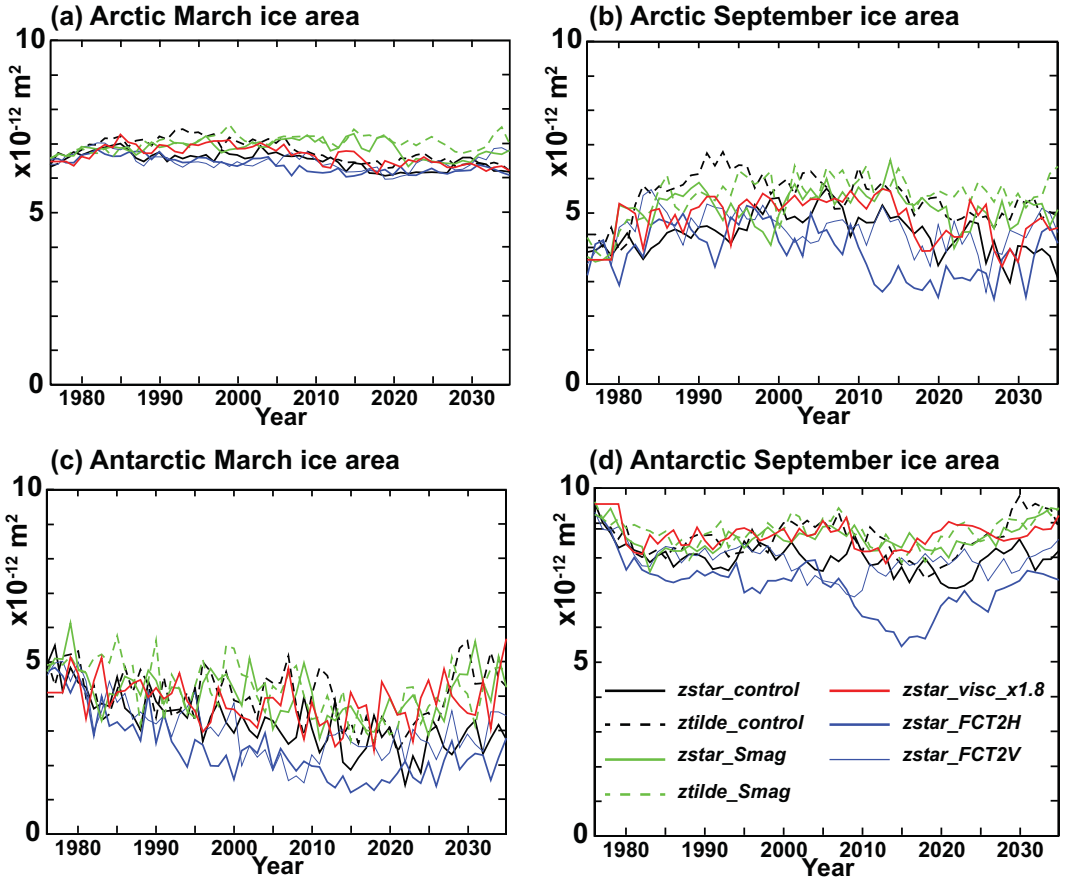


FIG. 7. Seasonal sea ice area (m^2) in the model ensemble: (a) Northern Hemisphere March, (b) Northern Hemisphere September, (c) Southern Hemisphere March, and (d) Southern Hemisphere September.

strongest mixing have the largest reduction in sea ice cover, which is consistent with the sensitivity of the sea surface temperature shown in Fig. 2.

d. Large-scale ocean transports

Figure 8 shows the strength of the AMOC at 26°N (Figs. 8a and b) and at 45°N (Figs. 8c and d), and the total volume transport in the ACC through the Drake Passage (Fig. 8e). The overturning circulation at both latitudes shows considerable interannual variability that partially obscures the dependence of the mean overturning strength on the model settings, so Figs. 8b and d show the same time series as in Figs. 8a and c, respectively, but with a nine-point Parzen filter applied. This latter shows that the mean overturning at 26°N varies by 1–2 Sv from that in *zstar_control* and reveals that there is a tendency for reduced numerical mixing (green, red, and dashed black lines in Fig. 8b) to weaken the overturning at 26°N. The converse, however, is not true: The AMOC strength in *zstar_FCT2H*, which has enhanced numerical mixing, is indistinguishable from that in the control. There is a weaker sensitivity of the overturning at 45°N to model settings. The sensitivity of the AMOC at 26°N is consistent with those observed in comparable CORE2-forced ¼° configurations: Megann and Storkey (2021), Megann et al. (2022), and Megann (2024) all showed that reduced levels of numerical mixing were well correlated with a weakened AMOC.

Storkey et al. (2025) compared the strength of the ACC in the CORE2-forced GOSI9 and in the coupled HadGEM3-GC5 (again with an eORCA025 GOSI9 ocean, but with a higher-resolution N216 atmosphere than that used in the present ensemble) and showed that enabling the Southern Ocean package (described in section 2a) in the coupled model reduced the spindown of the ACC transport, strengthening it from about 90 Sv at the end of the 30-yr integration to 105 Sv. In *zstar_control*, the ACC strength (the solid black line in Fig. 8e) shows a decline to about 110 Sv after 60 years. In the present ensemble, there is a strong and robust sensitivity to numerical mixing: In the last 30 years of integration, the spread across the ensemble is about 25 Sv. The ACC transport in *zstar_FCT2H* and *zstar_FCT2V* is barely distinguishable from that in *zstar_control*, but in *ztilde_control*, along with all three of the simulations with modified viscosity, the ACC strength robustly strengthens between 10 and 20 Sv with respect to that in the control. We conclude that reducing numerical mixing reliably addresses this recognized weakness of HadGEM3-GC5 with a ¼° ocean component. These results are again

Brought to you by NATIONAL OCEANOGRAPHY CENTRE | Unauthenticated | Downloaded 11/11/26 12:54 PM UTC

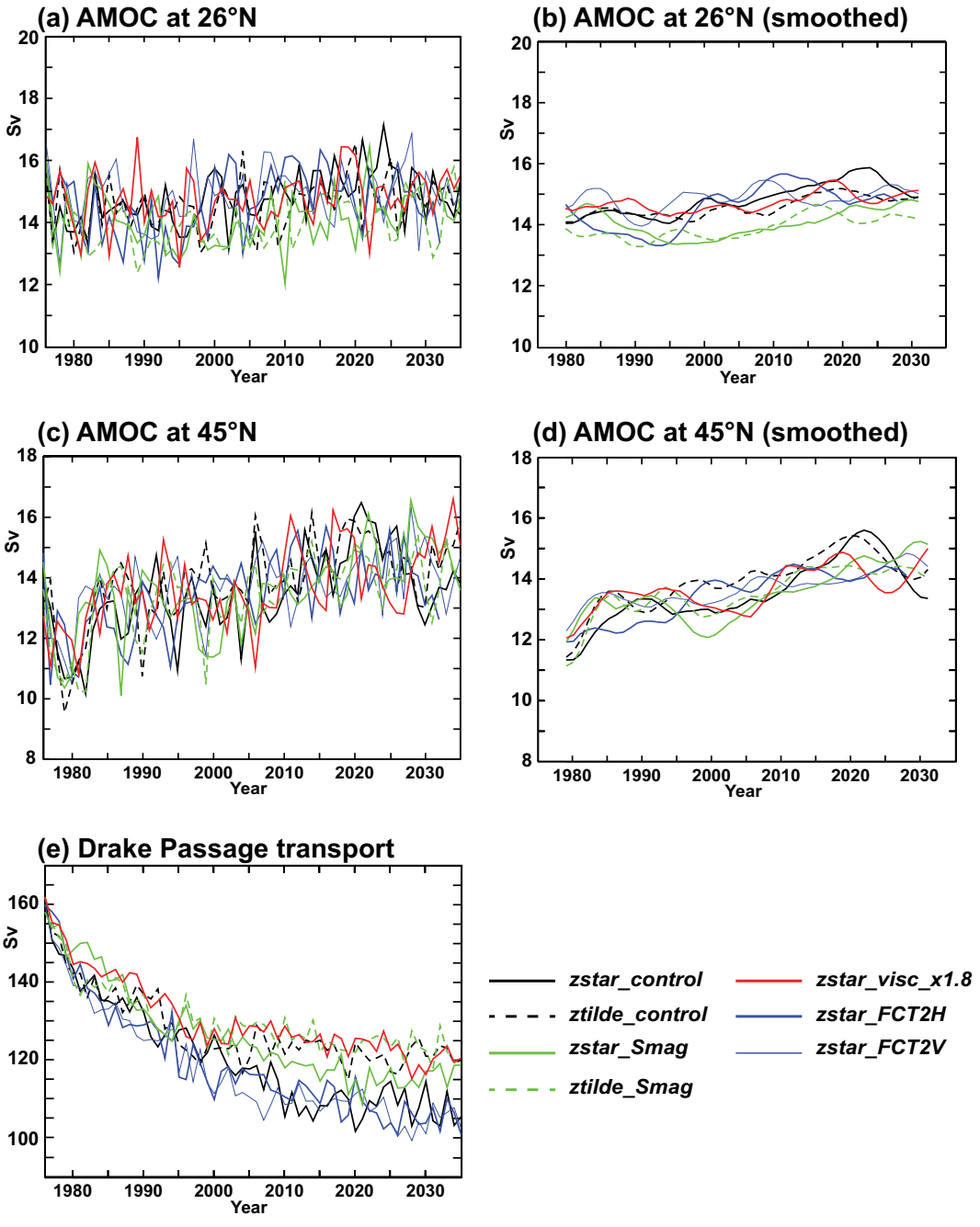


FIG. 8. Annual ocean volume transports (Sv) in the ensemble: (a) overturning strength at 26°N, (b) overturning strength at 26°N with a 9-yr smoother applied, (c) overturning strength at 45°N, (d) overturning strength at 45°N with a 9-yr smoother applied, and (e) transport through the Drake Passage.

consistent with those already published for comparable forced $^1/_4°$ configurations: Megann and Storkey (2021) found the ACC strengthened consistently with increased viscosity; Megann et al. (2022) demonstrated that $z\sim$ strengthened the ACC relative to that with z^*; Megann (2024) showed that applying tidal forcing with z^* weakened the ACC by about 10 Sv, but that with tides and $z\sim$, the ACC strength increased beyond that of the z^* non-tidal control.

e. *Correlations of ocean metrics with mixing*

We have demonstrated in the previous subsections that the four GC5 simulations with the $z\sim$ vertical coordinate and the changes in viscosity (*ztilde_control*, *zstar_Smag3*, *ztilde_Smag3*, and *zstar_visc_x1.8*) all show a significant reduction in the mixing metric D_{IDW} and that the sensitivity of the simulated ocean to these changes is consistent in both its sign and its spatial pattern across these integrations. Degrading the horizontal tracer

Brought to you by NATIONAL OCEANOGRAPHY CENTRE | Unauthenticated | Downloaded 11/11/26 12:54 PM UTC

advection to second order (*zstar_FCT2H*) tends to give changes that are opposite to those observed in the aforementioned runs, while reducing the accuracy of the vertical tracer advection (*zstar_FCT2V*) is mainly neutral in its effects. In this section, we present quantitatively the main ocean metrics for all of the simulations and present correlations between the metrics and the mixing metric D_{IDW}.

Table 4 lists the means over years 21–60 of selected ocean and sea ice quantities for each simulation. These include the surface temperature and downward surface heat flux Q_{net}; the AMOC at 26° and 45°N; the global meridional ocean heat transport (MOHT) at 30°S, 26° and 45°N; and the hemispheric mean seasonal mixed layer depths and sea ice areas. The three right-hand columns list quantities, evaluated for the last 40 years of each integration, that relate to the significance of the correlations between each metric and the mixing D_{IDW}. The first two are the standard error of the metric in each simulation, which quantifies the confidence in the respective mean value, and the standard error of the mean values over the ensemble, which describes the spread across the simulations: The ratio of these is a measure of the significance of the sensitivity of each metric to the mixing strength. The final column gives the correlation between the 40-yr time means for the simulations with the respective mixing metric D_{IDW}. We use an ad hoc critical value of $r = 0.8$ to define significance, and the coefficients which exceed this threshold are listed in bold type. We note that the metrics that have significant correlation with the density flux generally have a standard error for individual simulations that is at least 3 times smaller than the standard error across the ensemble; for example, the global-mean SST, with a correlation $r = 0.99$, has an ensemble spread of 4.2 times the standard error, whereas the surface heat flux, with a correlation of 0.64, has an ensemble spread of 1.8 times the standard error.

In Fig. 9, we present scatterplots of the main ocean metrics against the mixing metric D_{IDW}. A least squares fit and the correlation coefficient r are added to each panel. There is a remarkably strong positive correlation ($r = 0.99$) between mixing and the global-mean SST, which warms by ~0.7 K between the cases with lowest (*ztilde_Smag*) and highest (*zstar_FCT2H*) mixing. There is, by contrast, a weaker ($r = 0.64$) positive correlation between mixing and surface heat flux, and the range is only 0.1 W m^{-2}. The correlation between mixing and surface freshwater flux is lower at 0.20, and this is consistent with the large interannual variability seen in Fig. 6f.

The meridional overturning at 26°N is positively correlated with mixing, with $r = 0.70$, but that at 45°N is weakly negatively correlated: At 26°N, the range is about 1 Sv, corresponding to a standard error across the ensemble of 0.39 Sv. This is consistent with previous findings with forced models and confirms that mixing plays a significant role in maintaining the AMOC in this configuration. There is strong positive correlation between mixing and global MOHT at 26°N and at 45°N, with a range at the former latitude of about 6% of that in *zstar_control*, although the MOHT in the Southern Hemisphere is substantially less sensitive to mixing. The response of the heat transport in the Northern Hemisphere to mixing is likely to be related to that of the overturning strength in the North Atlantic, but changes in stratification attributable to mixing may also play a role. There is a strong negative correlation ($r = -0.91$) between mixing and the ACC transport: Reducing numerical mixing substantially addresses the well-known spindown in ACC transport in eORCA025 implementations of the GC3.1 coupled configuration (Storkey et al. 2025). The range in the present ensemble is over 15 Sv (the standard error is 3.3 Sv), which is comparable to that resulting from the application of the Southern Ocean package.

The mean hemispheric summer mixed layer depth is strongly negatively correlated with mixing, as is the Southern Hemisphere winter mixed layer depth (MLD), with a range of a few meters or between 4% and 7% of the mean MLD in *zstar_control*. Increasing mixing consistently reduces ice area in both hemispheres and in both seasons, which is consistent with the enhanced warming at high latitudes, and the sensitivity to mixing is both strong, with the range being between 10% and 30% of the ice area in *zstar_control* and robust, with correlation coefficients ranging between −0.92 and −0.99. This response of ice area to mixing will impact the shortwave radiative balance through the reduction in surface albedo.

f. *Development of ocean surface temperature changes*

To illustrate the time scale on which the SST sensitivities in the ensemble develop, Fig. 10 shows the evolution of the differences in annual zonal-mean SST in *zstar_FCT2H* (Fig. 10a) and *ztilde_Smag3* (Fig. 10b) from that in *zstar_control*, the latter two simulations having the highest and lowest numerical mixing, respectively. It is evident that the SST differences manifest within the first two decades of integration and are robust and stable thereafter. We note that, away from the equator, where SST variability is dominated by ENSO, the zonal-mean SST differences are almost uniform in time after the first two decades, which implies that the mean ocean circulation (which is, to first order, the same in all the integrations) rapidly equilibrates the changes around the global ocean, and the balance between vertical mixing and horizontal advection of heat is reached quickly and is stable thereafter. This uniformity of the response to model changes is also seen in the global SST and SSS fields (Figs. 2 and 3), which confirm that the spatial signature of the responses is very consistent across the ensemble, as well as in the high correlation ($r = 0.99$) of the mean surface temperature with the mixing index D_{IDW}. Figure 10 also confirms that the SST response is strongest in northern subpolar latitudes (between 35° and 60°N) and in the Southern Ocean (south of 45°S), being typically up to twice as strong as in the subtropics.

We have demonstrated that the surface temperatures in the present ensemble are warmer in the simulations with increased numerical mixing in the ocean and vice versa and that the correlation between the global-mean SST and the mixing metric D_{IDW} is remarkably strong. This is the opposite sign to the sensitivity that we might naively expect: In the absence of other influences, the first-order effect of enhanced vertical mixing would be to reduce the difference between the abyssal and surface temperatures and, hence, to result in surface cooling. We note, however, that several of the ocean responses to

Brought to you by NATIONAL OCEANOGRAPHY CENTRE | Unauthenticated | Downloaded 11/11/26 12:54 PM UTC

TABLE 4. Ocean metrics, averaged over years 21–60 of each integration. The four right-hand columns list the following quantities: the standard error of the respective time mean over the ensemble, the typical standard error of the time series, the ratio of these two, and the correlation between the time means for each simulation with the respective mixing metric D_{IW}. Values of the ratios of 3.0 or more are shown in bold, and values of the correlation coefficient r greater than or equal to 0.8 are also shown in bold.

	zstar_control	*ztilde_control*	*zstar_FCT2H*	*zstar_FCT2V*	*zstar_Smag*	*ztilde_Smag*	*zstar_visc_x1.8*	Mean std error	Ensemble std err	Ratio	Correlation with D_{IDW}
SST (°C)	18.1	17.8	18.4	18.1	17.9	17.7	17.9	±0.020	±0.084	**4.2**	**0.99**
SSS (°C)	34.59	34.58	34.61	34.58	34.56	34.56	34.57	±0.0019	±0.012	**6.3**	**0.81**
Q_{net} (W m^{-2})	0.65	0.69	0.64	0.74	0.55	0.54	0.60	±0.098	±0.18	1.8	0.64
FW flux (kg m^{-2} day^{-1})	0.0066	0.0073	0.0075	0.0063	0.0069	0.0065	0.0074	±0.0033	±0.0053	1.6	0.20
AMOC at 26°N (Sv)	15.1	14.7	14.9	15.0	14.1	14.3	14.9	±0.13	±0.39	**3.0**	0.70
AMOC at 45°N (Sv)	13.9	14.5	14.0	13.9	13.8	14.3	14.0	±0.19	±0.40	2.1	−0.52
ACC (Sv)	113	122	111	110	120	126	124	±0.43	±3.27	**7.6**	**−0.91**
MOHT 30°S (PW)	−0.485	−0.487	−0.485	−0.487	−0.540	−0.527	−0.502	±0.023	±0.062	2.7	−0.21
MOHT 26°N (PW)	1.61	1.53	1.61	1.59	1.54	1.52	1.56	±0.025	±0.069	2.8	**0.86**
MOHT 45°N (PW)	0.66	0.62	0.66	0.65	0.61	0.59	0.64	±0.010	±0.031	**3.1**	**0.80**
NH DJF MLD (m)	41.7	41.4	41.3	41.1	40.6	40.7	41.0	±0.28	±0.81	2.9	0.58
NH JJA MLD (m)	14.8	15.0	14.6	14.8	15.0	15.1	14.9	±0.07	±0.21	**3.0**	**−0.99**
SH DJF MLD (m)	26.5	27.3	26.2	26.5	26.9	27.8	26.9	±0.12	±0.41	**3.4**	**−0.92**
SH JJA MLD (m)	77.5	79.8	77.1	77.5	78.5	79.3	78.5	±0.29	±0.91	**3.1**	**−0.91**
NH March ice area (10^{12} m^2)	12.9	13.4	13.7	13.2	12.6	12.8	12.7	±0.055	±0.29	**5.3**	**−0.92**
NH September ice area (10^{12} m^2)	4.51	5.43	5.18	4.89	3.72	4.37	4.90	±0.075	±0.35	**4.7**	**−0.99**
SH March ice area (10^{12} m^2)	2.84	3.95	3.96	3.96	1.97	2.56	2.52	±0.10	±0.40	**4.0**	**−0.98**
SH September ice area (10^{12} m^2)	15.8	17.2	17.3	17.3	13.8	15.4	15.6	±0.14	±0.64	**4.7**	**−0.97**

Brought to you by NATIONAL OCEANOGRAPHY CENTRE | Unauthenticated | Downloaded 11/11/26 12:54 PM UTC

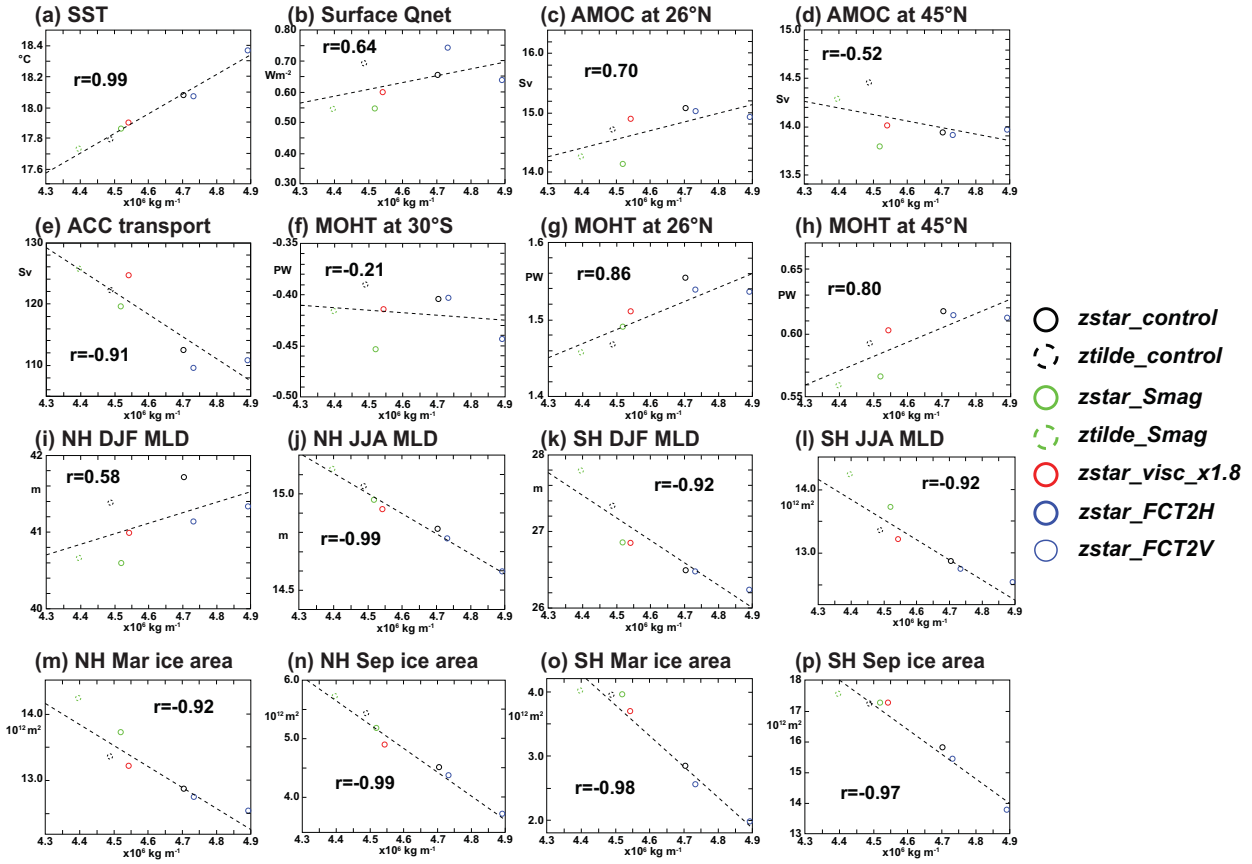


FIG. 9. Scatterplots of the principal ocean metrics against the mixing metric D_{IDW} defined in section 3a, with linear regression fits and correlation coefficients: (a) SST, (b) downward surface heat flux, (c) AMOC at 26°N, (d) AMOC at 45°N, (e) ACC transport, (f) MOHT at 30°S, (g) MOHT at 26°N, (h) MOHT at 45°N, (i) NH DJF MLD, (j) NH JJA MLD, (k) SH DJF MLD, (l) SH JJA MLD, (m) NH March sea ice area, (n) NH September sea ice area, (o) SH March sea ice area, and (p) SH September sea ice area.

numerical mixing that we have described provide plausible mechanisms contributing to the sign of the observed SST sensitivity. First of all, the strong positive correlation ($r > 0.8$) between northward ocean heat transports in the Northern Hemisphere and the mixing metric means that stronger poleward heat transport will tend to warm the upper ocean and vice versa, although the correlation in the Southern Hemisphere is much weaker. The MOHT at 26°N increases by about 0.1 PW across the ensemble (Fig. 9g) and by 0.05 PW at 45°N (Fig. 9h), corresponding in both cases to about 7% of the ensemble mean. The enhanced SST response in the northern subpolar latitude band seen in Figs. 2 and 10 is qualitatively consistent with this increase in poleward heat transport, although a full heat budget analysis would be required to confirm whether this mechanism is the dominant influence.

Second, the strong negative correlations of the summer MLD in both hemispheres (Figs. 9j,l), varying by about 3% of the mean across the ensemble in the Northern Hemisphere

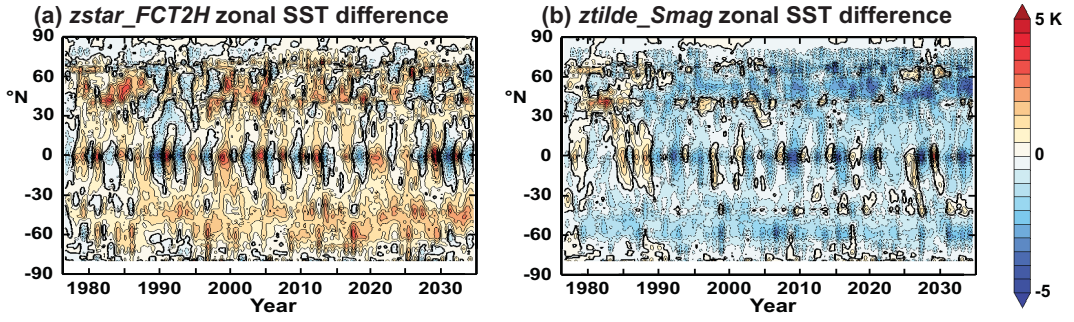


FIG. 10. Differences in annual zonal-mean SST (K) from *zstar_control* in (a) *zstar_FCT2H* and (b) *ztilde_Smag*.

Brought to you by NATIONAL OCEANOGRAPHY CENTRE | Unauthenticated | Downloaded 11/11/26 12:54 PM UTC

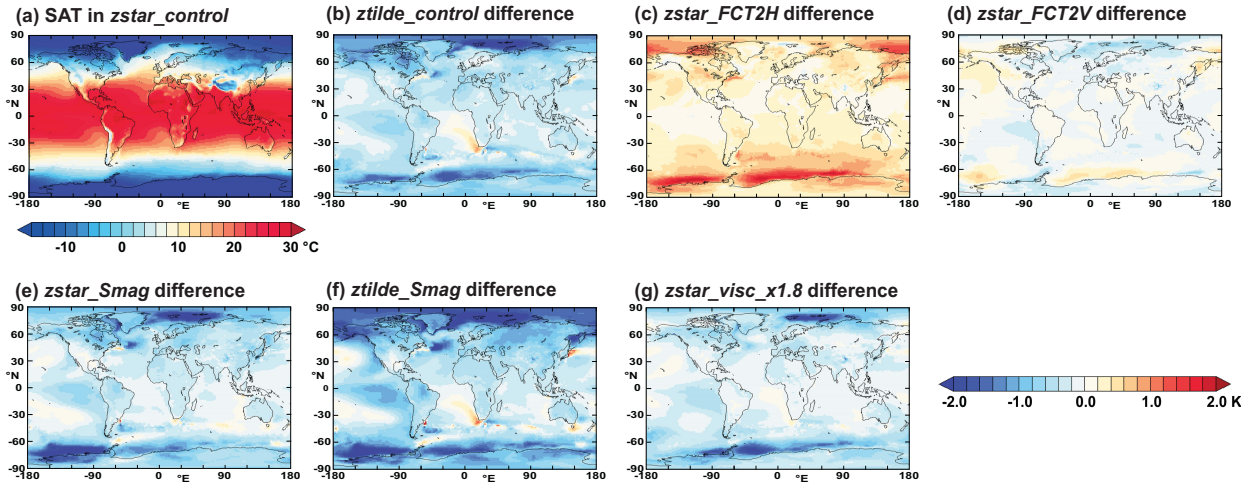


FIG. 11. (a) Annual-mean 1.5-m SAT (°C) in *zstar_control* and (b)–(g) SAT differences (K) from *zstar_control*. Note that (a) has a different scale to the other panels.

and 6% in the Southern Hemisphere, mean that, for a given heat input, the upper ocean will tend to warm more in summer in the simulations where the ML is shallower and vice versa. In polar regions, the SST sensitivity to mixing is more straightforward to explain, given the presence of warmer water beneath the shallow halocline. Finally, at high latitudes, enhanced upward heat fluxes from mixing of warm subsurface water will lead to increased melting of sea ice, which will be amplified through ice albedo feedback; Fig. 12 confirms that the response of the downward shortwave radiation to mixing is indeed enhanced south of 60°S and in the Arctic. Time scales for horizontal advection of temperature anomalies over basin scales by surface currents are generally substantially less than a decade (e.g., Jönsson and Watson 2016), so further disentangling of the effects of all these mechanisms on the decadal-scale evolution of the model is challenging and beyond the scope of this paper.

We note that the sensitivity of the sea surface temperature described in the present ensemble is substantially stronger than that seen in earlier forced ocean configurations using a 1/4° global NEMO. For example, in the viscosity experiments of Megann and Storkey (2021) and in the investigation of the $z\sim$ vertical coordinate by Megann et al. (2022), as well as other unpublished ensembles of forced 1/4° NEMO simulations, there is a tendency for stronger numerical mixing to lead to a warmer surface temperature, but the SST sensitivity is invariably 0.1–0.2 K or less, in contrast to the range of more than 1 K seen with the present ensemble. This is not surprising, since the use of bulk formulas to calculate the heat fluxes in the forced simulations, based on a fixed air temperature, will tend to damp the response of the sea surface to interior processes.

4. Atmosphere and coupled sensitivities to numerical mixing

a. *Atmosphere metrics*

In this section, we present the differences in annual means of surface atmosphere fields, evaluated over the last 40 years of the integrations, with respect to the respective field in *zstar_control*. Because the focus of this paper is on the ocean, surface heat fluxes are here defined as downward fluxes into the ocean; this is the opposite convention to that used in the reference papers published on HadGEM3 (Williams et al. 2018; Xavier et al. 2026). In the following paragraphs, we discuss the spatial dependence of the sensitivity of the surface and top-of-the-atmosphere (TOA) fields, which we shall subsequently quantify as global means (listed in Table 5) and finally present as scatterplots (Fig. 16). The sensitivities of some of the fields, notably the surface air temperature and the shortwave heat flux, have a strong seasonal dependence, and this will be referred to in the text, but for the sake of brevity, the seasonal cycle is not shown in the figures.

Figure 11a shows the mean air temperature at 1.5 m [we shall refer to this here as the surface air temperature (SAT)] in *zstar_control*, and the other panels show differences in the mean SAT from that in *zstar_control* in each simulation. To make comparisons between air and sea surface temperature changes more straightforward, this figure uses the same color scale for the difference panels as that of the figure showing sea surface temperature (Fig. 2). Between 60°S and 60°N, the sensitivity of the SAT lies between −0.5 and +0.5 K and is very close to that of the SST. At high latitudes, the SAT differences are considerably larger than those in the SST and are seasonally amplified, being up to 5 K larger in the hemispheric winter than in the respective summer, when they are closer to the SST differences. This is not unexpected, since over ice-covered oceans, the surface air can be much colder than the freezing point of water, and the response of the SAT will, therefore, respond to the changes in sea ice cover presented in Fig. 7.

Figure 12 shows the annual net downward surface shortwave radiation in *zstar_control* (Fig. 12a) and the mean differences in annual downward surface shortwave radiation from that in *zstar_control*. Reducing numerical mixing (Figs. 12b,e,f, and g) consistently reduces surface shortwave flux into the ocean by

Brought to you by NATIONAL OCEANOGRAPHY CENTRE | Unauthenticated | Downloaded 11/11/26 12:54 PM UTC

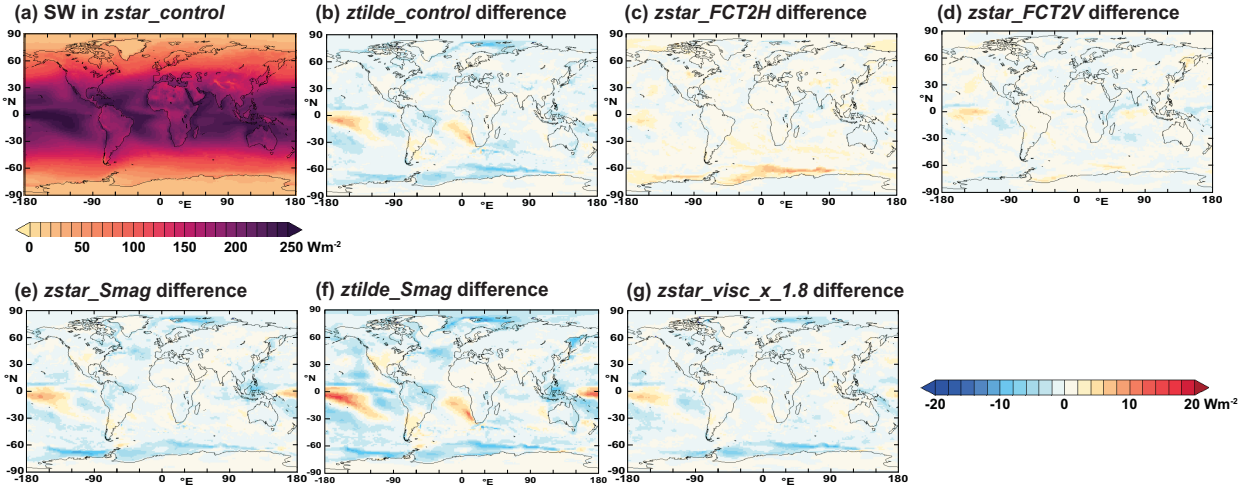


FIG. 12. (a) Annual-mean surface downward shortwave radiation (W m^{-2}) in *zstar_control* and (b)–(g) differences in downward shortwave from that in *zstar_control*. Note that (a) has a different scale to the other panels.

around 0.5 W m^{-2} and vice versa (Fig. 12c), with the exceptions of the southern tropical Pacific and the southern tropical Atlantic, where reducing mixing leads to increases in the downward shortwave heat flux. The spatial patterns of the responses to reduced mixing are consistent between the respective ensemble members: For example, while the shortwave flux is reduced almost everywhere in the simulations with altered viscosity or $z\sim$ (Figs. 12b,e,f, and g), it is consistently over 10 W m^{-2} larger in the same regions of the tropical South Atlantic and South Pacific. As with the SAT, the sensitivities are much higher in the respective hemispheric winter (not shown), consistent with the dependence of the sea ice cover on numerical mixing shown in Fig. 7. We would also expect the response of the net shortwave flux to be negatively correlated with sea ice cover, since increases in the latter increase the surface albedo: Comparison between Figs. 7 and 12 confirms that increased mixing reduces sea ice and increases the downward shortwave flux at high latitudes, and vice versa.

Figure 13 shows the differences in annual surface air–sea latent heat flux in watts per square meter, again with the sign convention that heat flux is positive into the ocean, so the latent heat flux is always negative. The latent heat flux calculated by the bulk formulas depends on the surface temperature, the surface wind speed, and the surface humidity; we note that the wind speed (not shown here) is only weakly dependent on mixing away from the Southern Ocean, while the humidity will largely be a function of surface temperatures, so we would expect surface temperature to be the dominant influence on the sensitivity of the latent heat flux. Comparison with Figs. 2 and 11 shows that the sensitivity of the latent heat flux over the ocean is indeed generally consistent with that of the surface temperatures: Increased numerical mixing warms the surface

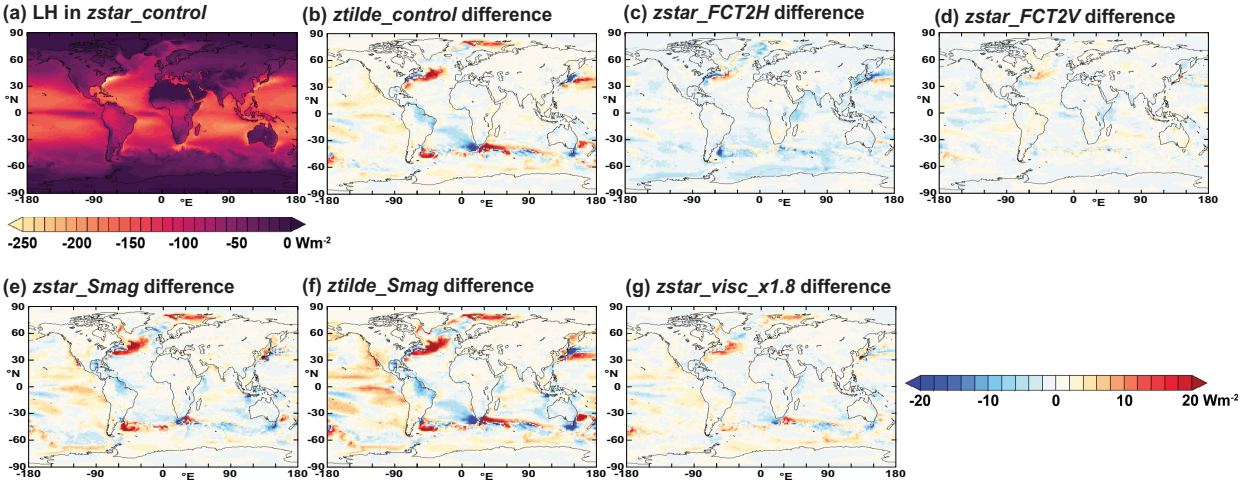


FIG. 13. (a) Annual-mean surface downward latent heat flux (W m^{-2}) in *zstar_control* and (b)–(g) differences from that in *zstar_control* on the same color scale as that in Figs. 4 and 12. Note that (a) has a different scale to the other panels.

Brought to you by NATIONAL OCEANOGRAPHY CENTRE | Unauthenticated | Downloaded 11/11/26 12:54 PM UTC

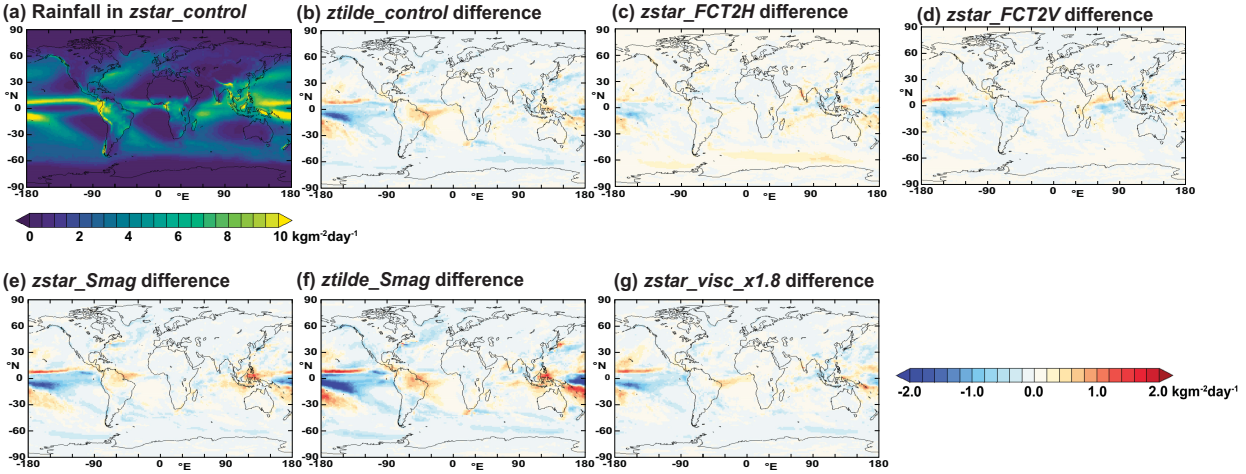


FIG. 14. (a) Annual rainfall (kg m^{-2} day^{-1}) in *zstar_control* and (b)–(g) mean differences in annual rainfall from that in *zstar_control*. Note that (a) has a different scale to the other panels.

[*zstar_FCT2H*, panel (c)], which increases the latent heat loss to the atmosphere between 2 and 5 W m^{-2}, and vice versa in the simulations with reduced mixing [panels (b), (e), (f), and (g)]. In the simulations with reduced numerical mixing, the reduction in latent heat loss from the ocean is particularly strong over the separated western boundary currents, where it exceeds 20 W m^{-2}. This again is consistent with the locations of the highest sensitivity of the surface temperature. The basin-scale sensitivity of the latent heat flux into the ocean is opposite to that of the net shortwave radiation: Warmer SST enhances latent heat loss but is also associated with a higher shortwave flux into the ocean. This is consistent with the SW adding extra heat to the ocean and the surface cooling, mainly through latent heat loss.

Figure 14 shows differences in rainfall from that in *zstar_control*, with Fig. 14a showing the rainfall in *zstar_control*. The unrealistic pattern of the double intertropical convergence zone (ITCZ) in the Pacific seen in *zstar_control* is characteristic of current climate models (Tian and Dong 2020). The sensitivities are highest in the tropics, especially in the ITCZ, chiefly because the rainfall itself is highest there. In the simulations with $z\sim$ or modified viscosity (Figs. 14b,e,f, and g), where the surface cools, rainfall is reduced over the ITCZ south of the equator and enhanced north of this region, thus increasing the realism of the simulations. A possible mechanism for this is that in the control simulation there is a negative feedback, whereby surface shortwave heats up the mixed layer, making the atmosphere more unstable, which generates more convective cloud. This would then reduce the surface shortwave, particularly in the southeast tropical Pacific. By reducing numerical mixing, we reduce the stratification and allow more heat to penetrate below the mixed layer, thereby weakening this negative feedback and reducing cloud and precipitation in the tropical southeast Pacific. This hypothesis is strengthened by the observation (Fig. 2) that sea surface temperature also cools in this location in those simulations with reduced numerical mixing. It is further corroborated by a reduction in midlevel convection and a corresponding increase in shallow convection (not shown). These changes indicate that air heated by the ocean mixed layer is not as buoyant (and, therefore, not as readily convecting to a level necessary for cloud formation) in the reduced numerical mixing simulations. Over most of the rest of the ocean, reducing numerical mixing, and cooling the SST, reduces the rainfall between 0.2 and 0.5 kg m^{-2} day^{-1}, and vice versa. Comparing Figs. 13 and 14, we see that changes in rainfall, as expected, are generally collocated and positively correlated with changes in cloud cover.

Figure 15 shows the differences in the mean cloud fraction from that in *zstar_control*. Reductions in cloud occur in regions of westerlies in the subtropical South Atlantic, North Atlantic, and South Pacific where the cloud fraction in the control is lower. More numerical mixing (*zstar_FCT2H*; Fig. 15c) reduces cloud over the ocean, while reducing numerical mixing (Figs. 15b,e,f, and g) increases cloud over the ocean, except for the south tropical Atlantic and Pacific. The sensitivity of both signs is amplified when Smagorinsky viscosity is combined with z^* (Fig. 15f). The sensitivity of the cloud fraction is in general opposite to that of the surface shortwave, as expected: more clouds mean that less shortwave radiation reaching the surface. More specifically, the similarity of the spatial patterns of the responses of rainfall and cloud cover in the tropics, particularly in the tropical Pacific, suggest that the negative cloud–SST feedback is strong in this region, and the rapid lateral SST adjustment we mention in section 3e allows this to partially suppress the subtropical positive cloud–SST feedback mechanism. This may have implications for the influence of numerical mixing in the ocean on climate warming responses.

b. *Correlation of atmospheric metrics with numerical mixing*

Table 5 lists the global means of the principal atmospheric metrics for each of the simulations, again averaged over the last 40 years of integration, and Fig. 16 shows scatterplots of the atmospheric metrics from Table 5 against the mixing

Brought to you by NATIONAL OCEANOGRAPHY CENTRE | Unauthenticated | Downloaded 11/11/26 12:54 PM UTC

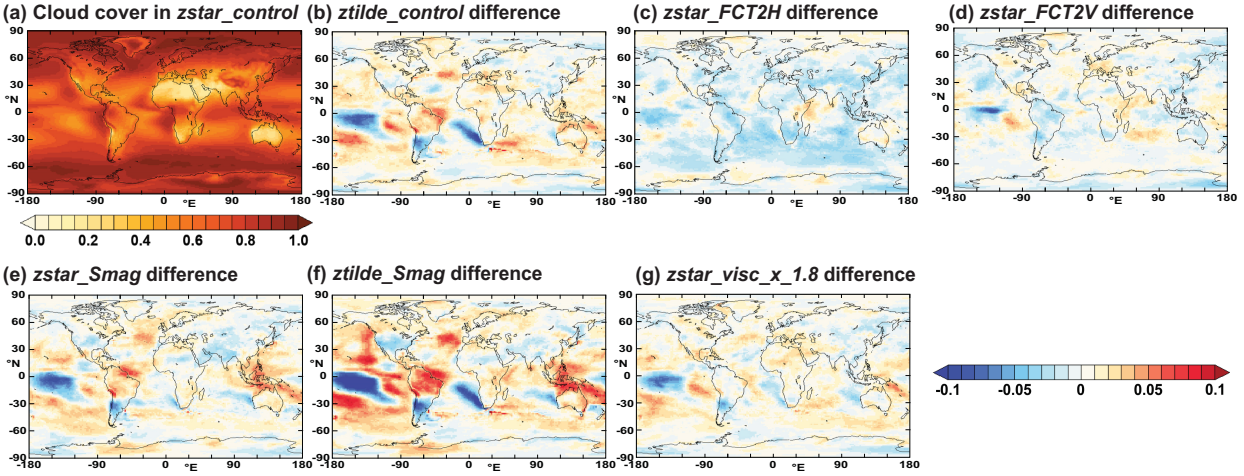


FIG. 15. (a) Annual-mean cloud fraction in *zstar_control* and (b)–(g) differences in cloud fraction from that in *zstar_control*.

metric D_{IDW}. The SAT is defined at 1.5 m and here, to aid comparison with the SST, it is again presented in degrees Celsius, rather than in the customary kelvins.

There is again a remarkably strong ($r = 0.99$) positive correlation between mixing and global SAT, consistent with that with the global SST. The total heat flux at top and bottom of the atmosphere both increase slightly with mixing (consistent with enhanced heat uptake by the ocean), but correlation is weak. The net surface shortwave radiation has very strong positive correlation with mixing; this is consistent with the negative correlations of the sea ice cover with mixing shown in Fig. 9, as well as with the cloud fraction. The net surface latent heat flux loss also increases with mixing, again consistent with the warmer surface temperatures. The main balance is between changes in the shortwave and latent heat fluxes, and these almost compensate in the global mean, but not locally. The sensitivities of the longwave and sensible heat fluxes are much weaker. Rainfall increases by up to 4% with mixing, with a robust correlation of 0.83. Regional sensitivities are much stronger: The changes described above in the ITCZ result in regional-mean rainfall changes by up to 2 kg m^{-2} day^{-1}. The albedo at the TOA is weakly sensitive to ocean mixing, reducing from 30.10% in *ztilde_Smag3* to 29.65% in *zstar_FCT2H*; this is consistent with the reduction in cloud cover observed in Fig. 15.

5. Summary and discussion

We have used an ensemble of the HadGEM3-GC5 coupled climate model with a 1/4° NEMO ocean component to investigate the sensitivity of the climate system to numerical mixing in the ocean. Within this ensemble, we have investigated the effect of changes to the ocean model settings, with respect to

TABLE 5. Atmospheric metrics, averaged over years 21–60 of each integration. The correlation coefficients which exceed the threshold of 0.8 are in bold type.

	zstar_control	*ztilde_control*	*zstar_FCT2H*	*zstar_FCT2V*	*zstar_Smag*	*ztilde_Smag*	*zstar_visc_x1.8*	Correlation with D_{IDW}
SAT (°C)	15.1	14.7	15.5	15.1	14.8	14.6	14.9	**0.99**
Albedo at TOA (%)	29.9	30.1	29.7	29.9	30.1	30.22	30.0	**−0.99**
Net TOA (W m^{-2})	0.83	0.85	0.81	0.88	0.74	0.75	0.77	0.51
Net Q_{surf} (W m^{-2})	1.43	1.48	1.39	1.48	1.38	1.39	1.40	0.09
Shortwave (W m^{-2})	165.3	164.7	165.78	165.3	164.7	164.4	164.9	**0.99**
Longwave (W m^{-2})	−56.00	−56.07	−56.00	−56.05	−56.04	−56.04	−55.98	0.40
Latent (W m^{-2})	−86.9	−86.2	−87.6	−86.9	−86.3	−86.0	−86.5	**−0.99**
Sensible (W m^{-2})	−20.7	−20.8	−20.6	−20.7	−20.8	−20.8	−20.8	**−0.96**
Rainfall (kg m^{-2} day^{-1})	2.78	2.74	2.81	2.77	2.75	2.73	2.75	**0.98**
Cloud cover	0.697	0.700	0.692	0.697	0.700	0.702	0.699	**−0.99**

Brought to you by NATIONAL OCEANOGRAPHY CENTRE | Unauthenticated | Downloaded 11/11/26 12:54 PM UTC

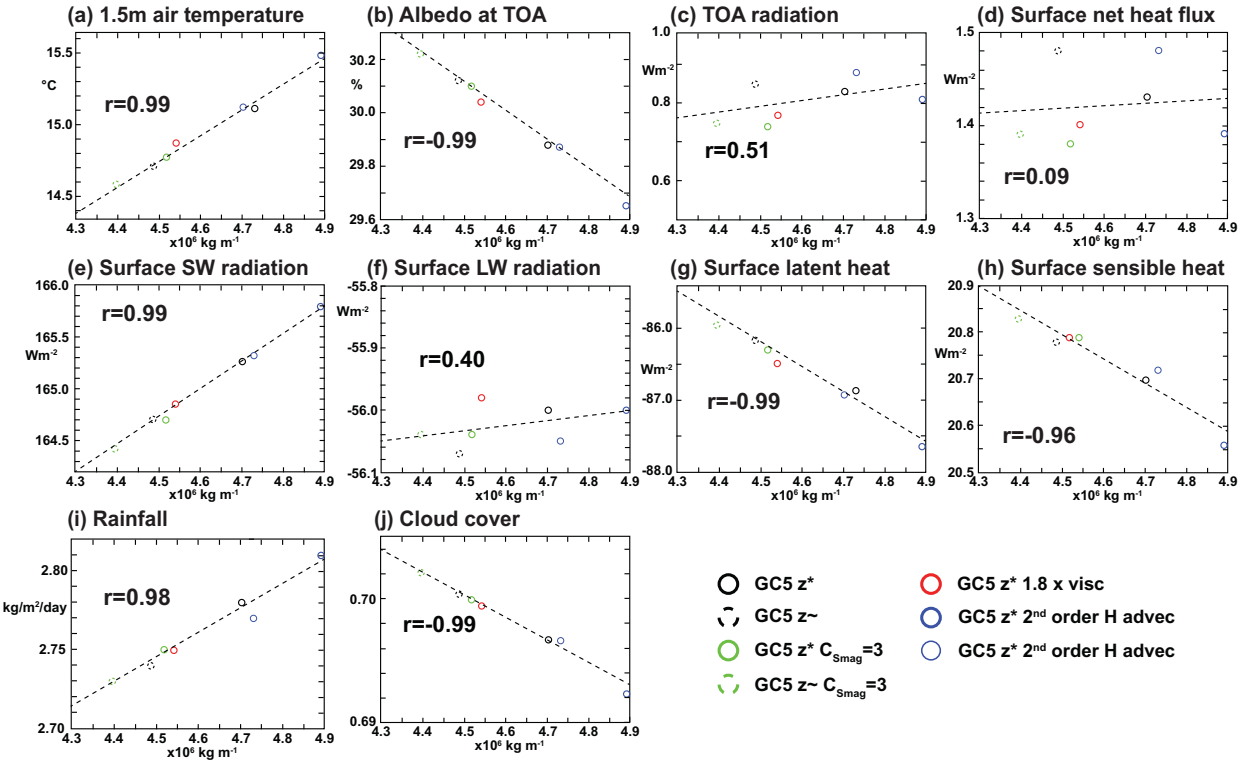


FIG. 16. Scatterplots of principal atmospheric metrics against the ocean mixing metric D_{IDW}, with linear regression fits and correlation coefficients superposed: (a) SAT, (b) albedo at TOA, (c) downwelling net radiation at TOA, (d) downwelling surface net heat flux, (e) downwelling surface shortwave, (f) downwelling surface longwave, (g) downwelling surface latent heat, (h) downwelling surface sensible heat, (i) rainfall, and (j) cloud cover.

those in the standard GOSI9 configuration, that have previously been shown to increase or decrease the numerical mixing in forced integrations of related $1/4°$ configurations. These changes include changing from the geopotential z^* vertical coordinate to the $z\sim$ filtered ALE coordinate, degrading the tracer advection in horizontal and vertical directions independently from fourth to second order, and changing the biharmonic viscosity either by switching to the Smagorinsky scheme or by scaling up the fixed viscosity. Each member is integrated for 60 years with constant year 2000 atmospheric composition and the analysis carried out on the final 40 years.

We have confirmed that a mixing metric based on the global upward density flux, evaluated from the meridional overturning streamfunction in potential density space, is a discriminating measure of total mixing across the ensemble. We have interpreted the difference between the total density flux and that inferred from the explicit diffusivity applied by the model as a measure of the numerical mixing and have demonstrated that this is close to zero for abyssal water densities; rises to a maximum of about 3 times the flux from explicit mixing at densities of about $\sigma_2 = 36.3$ kg m^{-3}, typical of intermediate waters; and reduces significantly in thermocline waters. Degrading the accuracy of the horizontal tracer advection to second order increases numerical mixing, while the $z\sim$ coordinate and Smagorinsky viscosity both consistently reduce numerical tracer mixing; the effects of the two latter changes combine additively in an integration with both applied together, leading to the lowest numerical mixing of all the simulations. Scaling the fixed viscosity by a factor of 1.8 leads to a reduction in numerical tracer mixing to a comparable degree to that obtained with the Smagorinsky viscosity, although at the cost of requiring a shorter time step for stability.

The mixing metric is strongly correlated over the whole ensemble with several large-scale ocean metrics with correlation coefficients of over $r = 0.8$ and in the case of the global-mean surface temperature as high as 0.99. Across the seven integrations, the global-mean sea surface temperature warms robustly by more than 0.8 K from the simulations with the lowest mixing to those with the highest. In the Northern Hemisphere, the poleward heat transport by the ocean increases with mixing, as also does the overturning strength at 26°N, albeit with a smaller correlation coefficient. Summer mixed layer depths in both hemispheres and winter mixed layer depths in the Southern Hemisphere reduce robustly with mixing, and hemispheric sea ice area in both winter and summer is strongly negatively correlated with mixing, consistent with the enhanced surface warming. The volume transport by the Antarctic Circumpolar Current (ACC) through the Drake Passage increases consistently with reduced mixing, with an overall strengthening of more than 15 Sv between the simulations with the highest and lowest numerical mixing. This is consistent with previous studies based on forced ocean models, which have linked the

Brought to you by NATIONAL OCEANOGRAPHY CENTRE | Unauthenticated | Downloaded 11/11/26 12:54 PM UTC

spindown of the ACC to slumping of the isopycnals across the subantarctic frontal system. We note that the strengthening by 15 Sv resulting from the combination of the Smagorinsky viscosity with the $z\sim$ coordinate is similar to the improvement reported by Storkey et al. (2025) that is gained from the application of the Southern Ocean package to the HadGEM3-GC3.1 coupled model, again with an eORCA025 ocean component.

The gross features of the overturning circulation in potential density space in the control simulation are qualitatively consistent with those estimated from observational data and reanalyses, particularly the analyses of LS2007 and C2019. In particular, the production and consumption of Antarctic Bottom Water is within the bounds of the observational estimates, as is the partition of the northward-flowing AABW between the Atlantic and Indo-Pacific basins. The overturning strength in the Atlantic is between 14 and 16 Sv, which is close to that observed by the Rapid Climate Change (RAPID) array at 26°N (Moat et al. 2020). Nevertheless, other aspects of the simulation indicate that the mixing in the model is higher than in the real ocean. In particular, the global temperature drifts of the model below the thermocline are consistent with an effective downward diffusive heat flux of between 1.5 and 2.0 W m^{-2}, which is substantially higher than the global-mean surface heat flux of 0.65 W m^{-2}. In addition, the lower potential densities of North Atlantic Deep Water and North Pacific Deep Water compared with those in the ECCO4 reanalysis suggest that there is excessive mixing with the overlying water masses.

We have also demonstrated significant, and again in some cases remarkably strong, correlations between large-scale atmospheric metrics and the mixing strength in the ocean. The surface air temperature closely follows the global-mean SST, apart from an amplification of the dependence on mixing at high latitudes in winter, related to the changes in sea ice cover, and the correlation between global-mean SAT and the mixing metric is again very strong ($r = 0.99$). Strong negative correlations with mixing are also found with the albedo at the top of atmosphere and the cloud cover, and positive correlations are found with global rainfall. The net downward shortwave and latent heat fluxes are individually strongly affected by mixing in the ocean, with the former increasing with mixing by about 1.5 W m^{-2} across the ensemble and the latter decreasing by about 2 W m^{-2}. The global net heat fluxes at the surface and the top of the atmosphere, by contrast, are not strongly sensitive to ocean mixing: These increase overall by between 0.2 and 0.3 W m^{-2}, but this change is not strongly correlated across the ensemble with ocean mixing. The changes in individual heat flux components almost cancel in the global mean, but not locally, with the shortwave radiation responding to changes in cloud cover and sea ice and the latent heat flux responding mainly to changes in the SST.

A striking observation from the results presented here is that the effective remedies applied here for numerical mixing, namely, the $z\sim$ vertical coordinate and the changes in viscosity settings, have spatially and quantitatively consistent effects on the diagnosed mixing and large-scale metrics, and the sensitivity of all the metrics to the combination of $z\sim$ and Smagorinsky viscosity is—to a good approximation—a linear sum of the responses to each of the changes made separately. Similarly, the sensitivity to degrading the accuracy of the horizontal tracer mixing is consistently opposite to that of the two former remedies for most metrics. This is evidenced quantitatively by the very high correlations between the density fluxes and many of the ocean and atmosphere metrics, as well as by the consistent spatial patterns of the responses of the surface fields. This is perhaps surprising, as these three changes are physically and numerically quite distinct from one another: $z\sim$ allows the model to represent internal waves in a more adiabatic way, increasing the viscosity reduces noise in the flow field at the grid scale, and increasing the order of the horizontal tracer advection improves the accuracy of the advection scheme in the presence of high tracer and velocity gradients (and vice versa). A possible reason for this consistency is the demonstration by Megann (2018) and Megann and Storkey (2021) that NEMO on the ¼° grid, with the default viscosity settings, has a high level of gridscale velocity noise that is concentrated near the western boundaries and intensifies poleward. Smoothing the velocity fields (and consequently the tracer fields) by increasing viscosity, or increasing the accuracy of the tracer advection, would both be expected to address this source of numerical mixing in the locations where the noise is strongest. We note that the response of the surface temperature to changes in the model settings equilibrates within a couple of decades to a near uniform pattern in both time and latitude away from the equator implies that the large-scale mean circulation, which is very similar across the ensemble, acts to homogenize the response around the global ocean within this time scale. This is consistent with the high degree of correlation of the global-mean surface temperature with the mixing metric. Finally, it is perhaps surprising that, while reducing the accuracy of the horizontal tracer advection increases numerical mixing, reducing that of the vertical tracer advection has little effect on mixing; one might expect that the latter change would have a qualitatively similar effect to that of changing from $z\sim$ to z^*, but this is not seen here. This suggests that the vertical resolution is high enough for second-order vertical advection to be sufficiently accurate overall, to deal with the vertical velocities and tracer gradients that occur in this configuration.

We have discussed potential mechanisms for the sensitivity of the ocean surface temperature to changes in numerical mixing in the ocean. These include strengthened meridional heat transport in the Northern Hemisphere with stronger mixing, shallower summer mixed layers, enhanced mixing at high latitudes of warmer waters beneath the shallow halocline, and increased insolation in regions with reduced sea ice cover, amplified through albedo feedback. These are based on responses that are shown to be robustly correlated with the mixing metric and are consistent in sign with the observed SST sensitivity, but a thorough investigation of these is outside the scope of this paper.

This study has used the HadGEM3-GC5 coupled model with fixed present-day atmospheric composition, since this was used in the model evaluation process (Xavier et al. 2026) to enable direct comparison with present-day climatology. For this reason, it is difficult to predict with confidence how the numerical mixing we have described in NEMO will affect

Brought to you by NATIONAL OCEANOGRAPHY CENTRE | Unauthenticated | Downloaded 11/11/26 12:54 PM UTC

the behavior of the simulated coupled system under plausible climate forcing scenarios. We have noted that the top-of-the-atmosphere radiation balance is relatively insensitive to the SST changes, mainly because of opposite and nearly balanced sensitivities of the surface net shortwave radiation and the latent heat flux, at least on the decadal time scales investigated here. Similarly, the rate of global ocean heat uptake is not robustly affected by the SST changes. Nevertheless, because of the nonlinearity of processes such as ice albedo feedback and also because of the different time scales of ocean, atmosphere, and cryosphere systems, the near balance of heat fluxes is unlikely to hold if the radiative balance in the atmosphere changes due to increasing greenhouse gas concentrations. We note that in Earth system models with interactive ice shelves, the rate of basal melting, and hence the rate of ice loss, is strongly dependent on the upper-ocean temperature (e.g., Gudmundsson et al. 2019; Naughten et al. 2023), so the biases reported here would be expected to have a proportionately larger effect.

As a result of this study, we make some tentative recommendations for future model development. Upgrading the tracer advection to fourth order, as part of the Southern Ocean package, was demonstrated during the development of GOSI9 (Storkey et al. 2025; Guiavarc'h et al. 2025) to substantially reduce drifts and biases in both forced and coupled simulations, and we have confirmed here that the change to fourth-order horizontal advection in particular significantly reduces numerical mixing. This has a negligible computational overhead and has no adverse effects on stability, so is recommended for this class of model in general. Increased viscosity has been shown in forced models with $1/4^\circ$ resolution (Megann and Storkey 2021) to give reliable improvements in realism; however, in GOSI9, it requires the time step to be reduced by 33%. If this is an acceptable penalty in model throughput, this change is recommended for climate models with this ocean resolution. The present study confirms, along with previous ocean-only studies (Megann et al. 2022; Megann 2024), that using the $z\sim$ coordinate significantly reduces spurious mixing, albeit at a cost in run speed, as well as requiring a shortened time step. The $z\sim$ coordinate is not implemented in standard NEMO releases from v4.2 onward, but work is ongoing to develop a more general hybrid vertical coordinate in NEMO, which will provide the option of a hybrid isopycnal coordinate similar to that in MOM6.

Acknowledgments. This work was carried out under the Joint Marine Modelling Programme (JMMP), a partnership between the Met Office and U.K. research centres: the National Oceanography Centre, the British Antarctic Survey, and the Centre for Polar Observation and Modelling. The work was supported by the Natural Environment Research Council (NERC) jointly under the Climate Linked Atlantic Sector Science (CLASS) marine research programme (NE/R015953/1) and the Atlantic Climate and Environment Strategic Science (AtlantiS) National Capability programme (NE/Y005589/1). The work was also supported under the HORIZON Europe project Explaining and Predicting the Ocean Conveyor (EPOC), HORIZON-CL6-2021-CLIMATE-01. The model was integrated on the MONSooN2 system, a collaborative facility at the Met Office provided through JMMP, and analysis was carried out on the JASMIN platform, funded by NERC and maintained by the Centre for Environmental Data Analysis (CEDA). We are grateful to Adam Blaker, Diego Bruciaferri, Jonathan Gregory, and Joel Hirschi for fruitful discussions and especially to George Nurser for his comments on the manuscript. The comments of the three anonymous reviewers have substantially clarified and strengthened the paper. We would also like to thank Monica Hanley and her team at the British Oceanographic Data Centre for archiving the model data.

Data availability statement. Monthly means of the ocean and sea ice output fields, the 5-day mean ocean fields used for the mixing analysis, and seasonal-mean atmosphere fields from each simulation are archived at the Centre for Environmental Data Analysis (CEDA). The data are available as follows: Megann and Copsey (2025a): https://www.bodc.ac.uk/data/published_data_library/catalogue/10.5285/41e78976-5752-d155-e063-7086abc0ff3c/. Data are accessible at https://data.ceda.ac.uk/bodc/soc251000. Megann and Copsey (2025b): https://www.bodc.ac.uk/data/published_data_library/catalogue/10.5285/41e78b5c-1af5-d1c4-e063-7086abc0a452/. Data are accessible at https://data.ceda.ac.uk/bodc/soc251001. Megann and Copsey (2025c): https://www.bodc.ac.uk/data/published_data_library/catalogue/10.5285/41e79307-c757-d324-e063-7086abc0b4f5/. Data are accessible at https://data.ceda.ac.uk/bodc/soc251003. Megann and Copsey (2025d): https://www.bodc.ac.uk/data/published_data_library/catalogue/10.5285/41e78dce-9b0e-d23d-e063-7086abc05c21/. Data are accessible at https://data.ceda.ac.uk/bodc/soc251002. Megann and Copsey (2025e): https://www.bodc.ac.uk/data/published_data_library/catalogue/10.5285/41e795b3-c5fa-d3b5-e063-7086abc01c43/. Data are accessible at https://data.ceda.ac.uk/bodc/soc251004. Megann and Copsey (2025f): https://www.bodc.ac.uk/data/published_data_library/catalogue/10.5285/41e79e77-1058-d599-e063-7086abc05c63/. Data are accessible at https://data.ceda.ac.uk/bodc/soc251007. Megann and Copsey (2025g): https://www.bodc.ac.uk/data/published_data_library/catalogue/10.5285/41e79a76-7299-d4be-e063-7086abc00252/. Data are accessible at https://data.ceda.ac.uk/bodc/soc251006. The ocean and sea ice namelists are archived at Zenodo: https://doi.org/10.5281/zenodo.17395842.

REFERENCES

Adcroft, A., and J.-M. Campin, 2004: Rescaled height coordinates for accurate representation of free-surface flows in ocean circulation models. *Ocean Modell.*, **7**, 269–284, https://doi.org/10.1016/j.ocemod.2003.09.003.

——, and Coauthors, 2019: The GFDL global ocean and sea ice model OM4.0: Model description and simulation features. *J. Adv. Model. Earth Syst.*, **11**, 3167–3211, https://doi.org/10.1029/2019MS001726.

Bleck, R., 2002: An oceanic general circulation model framed in hybrid isopycnic-Cartesian coordinates. *Ocean Modell.*, **4**, 55–88, https://doi.org/10.1016/S1463-5003(01)00012-9.

Brought to you by NATIONAL OCEANOGRAPHY CENTRE | Unauthenticated | Downloaded 11/11/26 12:54 PM UTC

——, and L. T. Smith, 1990: A wind-driven isopycnic coordinate model of the north and equatorial Atlantic Ocean: 1. Model development and supporting experiments. *J. Geophys. Res.*, **95**, 3273–3285, https://doi.org/10.1029/JC095iC03p03273.

Bruciaferri, D., G. I. Shapiro, and F. Wobus, 2018: A multi-envelope vertical coordinate system for numerical ocean modelling. *Ocean Dyn.*, **68**, 1239–1258, https://doi.org/10.1007/s10236-018-1189-x.

——, C. Guiavarc'h, H. T. Hewitt, J. Harle, M. Almansi, P. Mathiot, and P. Colombo, 2024: Localized general vertical coordinates for quasi-Eulerian ocean models: The Nordic overflows test-case. *J. Adv. Model. Earth Syst.*, **16**, e2023MS003893, https://doi.org/10.1029/2023MS003893.

Bryan, F., 1987: Parameter sensitivity of primitive equation ocean general circulation models. *J. Phys. Oceanogr.*, **17**, 970–985, https://doi.org/10.1175/1520-0485(1987)017<0970:PSOPEO>2.0.CO;2.

Bush, M., and Coauthors, 2025: The third Met Office Unified Model–JULES Regional Atmosphere and Land configuration, RAL3. *Geosci. Model Dev.*, **18**, 3819–3855, https://doi.org/10.5194/gmd-18-3819-2025.

Cessi, P., 2019: The global overturning circulation. *Ann. Rev. Mar. Sci.*, **11**, 249–270, https://doi.org/10.1146/annurev-marine-010318-095241.

Cimoli, L., and Coauthors, 2023: Significance of diapycnal mixing within the Atlantic Meridional Overturning Circulation. *AGU Adv.*, **4**, e2022AV000800, https://doi.org/10.1029/2022AV000800.

Cox, M. D., 1987: Isopycnal diffusion in a *z*-coordinate ocean model. *Ocean Modelling*, **74** (unpublished manuscripts), 1–5.

de Lavergne, C., G. Madec, J. Le Sommer, A. G. Nurser, and A. C. N. Garabato, 2016: On the consumption of Antarctic Bottom Water in the abyssal ocean. *J. Phys. Oceanogr.*, **46**, 635–661, https://doi.org/10.1175/JPO-D-14-0201.1.

Dewar, W. K., and J. C. McWilliams, 2019: On energy and turbulent mixing in the thermocline. *J. Adv. Model. Earth Syst.*, **11**, 578–596, https://doi.org/10.1029/2018MS001502.

Dickson, R. R., and J. Brown, 1994: The production of North Atlantic Deep Water: Sources, rates, and pathways. *J. Geophys. Res.*, **99**, 12 319–12 341, https://doi.org/10.1029/94JC00530.

Dunne, J. P., and Coauthors, 2012: GFDL's ESM2 global coupled climate–carbon Earth System Models. Part I: Physical formulation and baseline simulation characteristics. *J. Climate*, **25**, 6646–6665, https://doi.org/10.1175/JCLI-D-11-00560.1.

——, and Coauthors, 2020: The GFDL Earth System Model Version 4.1 (GFDL-ESM 4.1): Overall coupled model description and simulation characteristics. *J. Adv. Model. Earth Syst.*, **12**, e2019MS002015, https://doi.org/10.1029/2019MS002015.

Feistel, R., 2018: Thermodynamic properties of seawater, ice and humid air: TEOS-10, before and beyond. *Ocean Sci.*, **14**, 471–502, https://doi.org/10.5194/os-14-471-2018.

Gaspar, P., Y. Grégoris, and J.-M. Lefevre, 1990: A simple eddy kinetic energy model for simulations of the oceanic vertical mixing: Tests at station Papa and long-term upper ocean study site. *J. Geophys. Res.*, **95**, 16 179–16 193, https://doi.org/10.1029/JC095iC09p16179.

Gent, P. R., and J. C. McWilliams, 1990: Isopycnal mixing in ocean circulation models. *J. Phys. Oceanogr.*, **20**, 150–155, https://doi.org/10.1175/1520-0485(1990)020<0150:IMIOCM>2.0.CO;2.

Good, S. A., M. J. Martin, and N. A. Rayner, 2013: EN4: Quality controlled ocean temperature and salinity profiles and monthly objective analyses with uncertainty estimates. *J. Geophys. Res. Oceans*, **118**, 6704–6716, https://doi.org/10.1002/2013JC009067.

Griffies, S. M., and R. W. Hallberg, 2000: Biharmonic friction with a Smagorinsky-like viscosity for use in large-scale eddy-permitting ocean models. *Mon. Wea. Rev.*, **128**, 2935–2946, https://doi.org/10.1175/1520-0493(2000)128<2935:BFWASL>2.0.CO;2.

——, R. C. Pacanowski, and R. W. Hallberg, 2000: Spurious diapycnal mixing associated with advection in a *z*-coordinate ocean model. *Mon. Wea. Rev.*, **128**, 538–564, https://doi.org/10.1175/1520-0493(2000)128<0538:SDMAWA>2.0.CO;2.

Gudmundsson, G. H., F. S. Paolo, S. Adusumilli, and H. A. Fricker, 2019: Instantaneous Antarctic ice sheet mass loss driven by thinning ice shelves. *Geophys. Res. Lett.*, **46**, 13 903–13 909, https://doi.org/10.1029/2019GL085027.

Guiavarc'h, C., and Coauthors, 2025: GOSI9: UK Global Ocean and Sea Ice configurations. *Geosci. Model Dev.*, **18**, 377–403, https://doi.org/10.5194/gmd-18-377-2025.

Hallberg, R., and A. Adcroft, 2009: Reconciling estimates of the free surface height in Lagrangian vertical coordinate ocean models with mode-split time stepping. *Ocean Modell.*, **29**, 15–26, https://doi.org/10.1016/j.ocemod.2009.02.008.

Held, I. M., and Coauthors, 2019: Structure and performance of GFDL's CM4.0 climate model. *J. Adv. Model. Earth Syst.*, **11**, 3691–3727, https://doi.org/10.1029/2019MS001829.

Holmes, R. M., J. D. Zika, S. M. Griffies, A. M. Hogg, A. E. Kiss, and M. H. England, 2021: The geography of numerical mixing in a suite of global ocean models. *J. Adv. Model. Earth Syst.*, **13**, e2020MS002333, https://doi.org/10.1029/2020MS002333.

Huber, M., R. Tailleux, D. Ferreira, T. Kuhlbrodt, and J. Gregory, 2015: A traceable physical calibration of the vertical advection-diffusion equation for modeling ocean heat uptake. *Geophys. Res. Lett.*, **42**, 2333–2341, https://doi.org/10.1002/2015GL063383.

Ilıcak, M., A. J. Adcroft, S. M. Griffies, and R. W. Hallberg, 2012: Spurious dianeutral mixing and the role of momentum closure. *Ocean Modell.*, **45–46**, 37–58, https://doi.org/10.1016/j.ocemod.2011.10.003.

Jayne, S., 2009: The impact of abyssal mixing parameterizations in an ocean general circulation model. *J. Phys. Oceanogr.*, **39**, 1756–1775, https://doi.org/10.1175/2009JPO4085.1.

Jönsson, B., and J. Watson, 2016: The timescales of global surface-ocean connectivity. *Nat. Commun.*, **7**, 11239, https://doi.org/10.1038/ncomms11239.

Klocker, A., and T. J. McDougall, 2010: Influence of the nonlinear equation of state on global estimates of dianeutral advection and diffusion. *J. Phys. Oceanogr.*, **40**, 1690–1709, https://doi.org/10.1175/2010JPO4303.1.

Leclair, M., and G. Madec, 2011: $\tilde{z}$-coordinate, an Arbitrary Lagrangian–Eulerian coordinate separating high and low frequency motions. *Ocean Modell.*, **37**, 139–152, https://doi.org/10.1016/j.ocemod.2011.02.001.

Lee, M.-M., A. C. Coward, and A. G. Nurser, 2002: Spurious diapycnal mixing of the deep waters in an eddy-permitting global ocean model. *J. Phys. Oceanogr.*, **32**, 1522–1535, https://doi.org/10.1175/1520-0485(2002)032<1522:SDMOTD>2.0.CO;2.

Levier, B., A.-M. Tréguier, G. Madec, and V. Garnier, 2007: Free surface and variable volume in the NEMO code. MERSEA IP Rep. WP09-CNRS-STR-03-1A, 46 pp., https://zenodo.org/record/3244182/export/hx#.Y1JaqS8w1R4.

Brought to you by NATIONAL OCEANOGRAPHY CENTRE | Unauthenticated | Downloaded 11/11/26 12:54 PM UTC

Lumpkin, R., and K. Speer, 2007: Global ocean meridional overturning. *J. Phys. Oceanogr.*, **37**, 2550–2562, https://doi.org/10.1175/JPO3130.1.

Macdonald, A., and C. Wunsch, 1996: An estimate of global ocean circulation and heat fluxes. *Nature*, **382**, 436–439, https://doi.org/10.1038/382436a0.

Madec, G., and Coauthors, 2017: NEMO ocean engine. Note du Pôle de modélisation de l'Institut Pierre-Simon Laplace 27, 412 pp., https://doi.org/10.5281/zenodo.1464816.

Marshall, J., and K. G. Speer, 2012: Closure of the meridional overturning circulation through Southern Ocean upwelling. *Nat. Geosci.*, **5**, 171–180, https://doi.org/10.1038/ngeo1391.

Megann, A., 2018: Estimating the numerical diapycnal mixing in an eddy-permitting ocean model. *Ocean Modell.*, **121**, 19–33, https://doi.org/10.1016/j.ocemod.2017.11.001.

——, 2024: Quantifying numerical mixing in a tidally forced global eddy-permitting ocean model. *Ocean Modell.*, **188**, 102329, https://doi.org/10.1016/j.ocemod.2024.102329.

——, and D. Storkey, 2021: Exploring viscosity space in an eddy-permitting global ocean model: Is viscosity a useful control for numerical mixing? *J. Adv. Model. Earth Syst.*, **13**, e2020MS002263, https://doi.org/10.1029/2020MS002263.

——, and D. Copsey, 2025a: Model output containing ocean, atmospheric and sea ice variables from the zstar_control simulation of the HadGEM3-GC5-LM coupled climate model (1976–2035). NERC EDS British Oceanographic Data Centre NOC, accessed 30 September 2026, https://doi.org/10.5285/41e78976-5752-d155-e063-7086abc0ff3c.

——, and ——, 2025b: Model output containing ocean, atmospheric and sea ice variables from the ztilde_control simulation of the HadGEM3-GC5-LM coupled climate model (1976–2035). NERC EDS British Oceanographic Data Centre NOC, accessed 30 September 2026, https://doi.org/10.5285/41e78b5c-1af5-d1c4-e063-7086abc0a452.

——, and ——, 2025c: Model output containing ocean, atmospheric and sea ice variables from the zstar_FCT2H simulation of the HadGEM3-GC5-LM coupled climate model (1976–2035). NERC EDS British Oceanographic Data Centre NOC, accessed 30 September 2026, https://doi.org/10.5285/41e79307-c757-d324-e063-7086abc0b4f5.

——, and ——, 2025d: Model output containing ocean, atmospheric and sea ice variables from the zstar_FCT2V simulation of the HadGEM3-GC5-LM coupled climate model (1976–2035). NERC EDS British Oceanographic Data Centre NOC, accessed 30 September 2026, https://doi.org/10.5285/41e78dce-9b0e-d23d-e063-7086abc05c21.

——, and ——, 2025e: Model output containing ocean, atmospheric and sea ice variables from the zstar_Smag simulation of the HadGEM3-GC5-LM coupled climate model (1976–2035). NERC EDS British Oceanographic Data Centre NOC, accessed 30 September 2026, https://doi.org/10.5285/41e795b3-c5fa-d3b5-e063-7086abc01c43.

——, and ——, 2025f: Model output containing ocean, atmospheric and sea ice variables from the ztilde_Smag simulation of the HadGEM3-GC5-LM coupled climate model (1976–2035). NERC EDS British Oceanographic Data Centre NOC, accessed 30 September 2026, https://doi.org/10.5285/41e79e77-1058-d599-e063-7086abc05c63.

——, and ——, 2025g: Model output containing ocean, atmospheric and sea ice variables from the zstar_visc_x1.8 simulation of the HadGEM3-GC5-LM coupled climate model (1976–2035). NERC EDS British Oceanographic Data Centre NOC, accessed 30 September 2026, https://doi.org/10.5285/41e79a76-7299-d4be-e063-7086abc00252.

——, A. L. New, A. T. Blaker, and B. Sinha, 2010: The sensitivity of a coupled climate model to its ocean component. *J. Climate*, **23**, 5126–5150, https://doi.org/10.1175/2010JCLI3394.1.

——, and Coauthors, 2014: GO5.0: The joint NERC–Met Office NEMO global ocean model for use in coupled and forced applications. *Geosci. Model Dev.*, **7**, 1069–1092, https://doi.org/10.5194/gmd-7-1069-2014.

——, J. Chanut, and D. Storkey, 2022: Assessment of the $\tilde{z}$ time-filtered arbitrary Lagrangian-Eulerian coordinate in a global eddy-permitting ocean model. *J. Adv. Model. Earth Syst.*, **14**, e2022MS003056, https://doi.org/10.1029/2022MS003056.

Merchant, C. J., and Coauthors, 2014: Sea surface temperature datasets for climate applications from Phase 1 of the European Space Agency Climate Change Initiative (SST CCI). *Geosci. Data J.*, **1**, 179–191, https://doi.org/10.1002/gdj3.20.

Moat, B. I., and Coauthors, 2020: Pending recovery in the strength of the meridional overturning circulation at 26°N. *Ocean Sci.*, **16**, 863–874, https://doi.org/10.5194/os-16-863-2020.

Moum, J. N., and T. R. Osborn, 1986: Mixing in the main thermocline. *J. Phys. Oceanogr.*, **16**, 1250–1259, https://doi.org/10.1175/1520-0485(1986)016<1250:MITMT>2.0.CO;2.

Munk, W., 1966: Abyssal recipes. *Deep-Sea Res. Oceanogr. Abstr.*, **13**, 707–730, https://doi.org/10.1016/0011-7471(66)90602-4.

——, and C. Wunsch, 1998: Abyssal recipes II: Energetics of tidal and wind mixing. *Deep-Sea Res. I*, **45**, 1977–2010, https://doi.org/10.1016/S0967-0637(98)00070-3.

Naughten, K., P. R. Holland, and J. De Rydt, 2023: Unavoidable future increase in West Antarctic ice-shelf melting over the twenty-first century. *Nat. Climate Change*, **13**, 1222–1228, https://doi.org/10.1038/s41558-023-01818-x.

Naveira Garabato, A. C., A. P. Williams, and S. Bacon, 2014: The three-dimensional overturning circulation of the Southern Ocean during the WOCE era. *Prog. Oceanogr.*, **120**, 41–78, https://doi.org/10.1016/j.pocean.2013.07.018.

Nikurashin, M., and G. Vallis, 2011: A theory of deep stratification and overturning circulation in the ocean. *J. Phys. Oceanogr.*, **41**, 485–502, https://doi.org/10.1175/2010JPO4529.1.

Orsi, A., W. Smethie, and J. Bullister, 2002: On the total input of Antarctic waters to the deep ocean: A preliminary estimate from chlorofluorocarbon measurements. *J. Geophys. Res.*, **107**, 3122, https://doi.org/10.1029/2001JC000976.

Polzin, K. L., J. M. Toole, J. R. Ledwell, and R. W. Schmitt, 1997: Spatial variability of turbulent mixing in the abyssal ocean. *Science*, **276**, 93–96, https://doi.org/10.1126/science.276.5309.93.

Redi, M. H., 1982: Oceanic isopycnal mixing by coordinate rotation. *J. Phys. Oceanogr.*, **12**, 1154–1158, https://doi.org/10.1175/1520-0485(1982)012<1154:OIMBCR>2.0.CO;2.

Robinson, A., and H. Stommel, 1959: The oceanic thermocline and the associated thermohaline circulation. *Tellus*, **11** (3), 295–308, https://doi.org/10.1111/j.2153-3490.1959.tb00035.x.

Roquet, F., G. Madec, T. J. McDougall, and P. M. Barker, 2015: Accurate polynomial expressions for the density and specific volume of seawater using the TEOS-10 standard. *Ocean Modell.*, **90**, 29–43, https://doi.org/10.1016/j.ocemod.2015.04.002.

Shchepetkin, A. F., 2015: An adaptive, Courant-number-dependent implicit scheme for vertical advection in oceanic modeling. *Ocean Modell.*, **91**, 38–69, https://doi.org/10.1016/j.ocemod.2015.03.006.

Brought to you by NATIONAL OCEANOGRAPHY CENTRE | Unauthenticated | Downloaded 11/11/26 12:54 PM UTC

Simmons, H. L., S. R. Jayne, L. C. S. Laurent, and A. J. Weaver, 2004: Tidally driven mixing in a numerical model of the ocean general circulation. *Ocean Modell.*, **6**, 245–263, https://doi.org/10.1016/S1463-5003(03)00011-8.

Smagorinsky, J., 1963: General circulation experiments with the primitive equations. *Mon. Wea. Rev.*, **91**, 99–164, https://doi.org/10.1175/1520-0493(1963)091<0099:GCEWTP>2.3.CO;2.

Speer, K., 1997: A note on average cross-isopycnal mixing in the North Atlantic ocean. *Deep-Sea Res. I*, **44**, 1981–1990, https://doi.org/10.1016/S0967-0637(97)00054-X.

Storkey, D., and Coauthors, 2018: UK Global Ocean GO6 and GO7: A traceable hierarchy of model resolutions. *Geosci. Model Dev.*, **11**, 3187–3213, https://doi.org/10.5194/gmd-11-3187-2018.

——, P. Mathiot, M. J. Bell, D. Copsey, C. Guiavarc'h, H. T. Hewitt, J. Ridley, and M. J. Roberts, 2025: Resolution dependence of interlinked Southern Ocean biases in global coupled HadGEM3 models. *Geosci. Model Dev.*, **18**, 2725–2745, https://doi.org/10.5194/gmd-18-2725-2025.

Talley, L. D., 2013: Closure of the global overturning circulation through the Indian, Pacific, and Southern Oceans: Schematics and transports. *Oceanography*, **26** (1), 80–97, https://doi.org/10.5670/oceanog.2013.07.

Tian, B., and X. Dong, 2020: The double-ITCZ bias in CMIP3, CMIP5, and CMIP6 models based on annual mean precipitation. *Geophys. Res. Lett.*, **47**, e2020GL087232, https://doi.org/10.1029/2020GL087232.

Toggweiler, J. R., and B. Samuels, 1995: Effect of Drake Passage on the global thermohaline circulation. *Deep-Sea Res. I*, **42**, 477–500, https://doi.org/10.1016/0967-0637(95)00012-U.

——, and ——, 1998: On the ocean's large-scale circulation near the limit of no vertical mixing. *J. Phys. Oceanogr.*, **28**, 1832–1852, https://doi.org/10.1175/1520-0485(1998)028<1832:OTOSLS>2.0.CO;2.

Tréguier, A. M., I. M. Held, and V. D. Larichev, 1997: Parameterization of quasigeostrophic eddies in primitive equation ocean models. *J. Phys. Oceanogr.*, **27**, 567–580, https://doi.org/10.1175/1520-0485(1997)027<0567:POQEIP>2.0.CO;2.

Urakawa, S., and H. Hasumi, 2014: Effect of numerical diffusion on the water mass transformation in eddy-resolving models. *Ocean Modell.*, **74**, 22–35, https://doi.org/10.1016/j.ocemod.2013.11.003.

Walin, G., 1982: On the relation between sea-surface heat flow and thermal circulation in the ocean. *Tellus*, **34** (2), 187–195, https://doi.org/10.3402/tellusa.v34i2.10801.

Weijer, W., W. Cheng, O. A. Garuba, A. Hu, and B. T. Nadiga, 2020: CMIP6 models predict significant 21st century decline of the Atlantic Meridional Overturning Circulation. *Geophys. Res. Lett.*, **47**, e2019GL086075, https://doi.org/10.1029/2019GL086075.

Wijffels, S. E., J. M. Toole, H. L. Bryden, R. A. Fine, W. J. Jenkins, and J. L. Bullister, 1996: The water masses and circulation at 10°N in the Pacific. *Deep-Sea Res. I*, **43**, 501–544, https://doi.org/10.1016/0967-0637(96)00006-4.

Williams, K. D., and Coauthors, 2018: The Met Office global coupled model 3.0 and 3.1 (GC3.0 and GC3.1) configurations. *J. Adv. Model. Earth Syst.*, **10**, 357–380, https://doi.org/10.1002/2017MS001115.

Xavier, P., and Coauthors, 2026: Assessment of the Met Office Global Coupled model version 5 (GC5) configurations. EGUSphere, https://doi.org/10.5281/zenodo.21353929.

Zalesak, S. T., 1979: Fully multidimensional flux-corrected transport algorithms for fluids. *J. Comput. Phys.*, **31**, 335–362, https://doi.org/10.1016/0021-9991(79)90051-2.

Brought to you by NATIONAL OCEANOGRAPHY CENTRE | Unauthenticated | Downloaded 11/11/26 12:54 PM UTC