The impact of a parameterisation of submesoscale mixed layer eddies on mixed layer depths in the NEMO ocean model

Daley Calvert[a,*], George Nurser[b], Michael J. Bell[a], Baylor Fox-Kemper[c]

[a] *Met Office, Fitzroy Road, Exeter, UK*
[b] *National Oceanography Centre, Southampton, UK*
[c] *Department of Earth, Environmental and Planetary Sciences, Brown University, Providence, Rhode Island 02906, USA*

Abstract

A parameterisation scheme for restratification of the mixed layer by submesoscale mixed layer eddies is implemented in the NEMO ocean model. Its impact on the mixed layer depth (MLD) is examined in 30-year integrations of "uncoupled" ocean-ice (GO5) and "coupled" atmosphere-ocean-ice-land (GC2) 1/4° global climate configurations used by the Met Office Hadley Centre. The impact of the scheme on the MLD in GO5 is up to twice as large in subtropical and mid-latitudes when the mixed layer Rossby radius is not limited to guard against CFL-type instabilities and excessively strong volume overturning. Such a limit is not found to be necessary for stable integration of the scheme in NEMO. An alternative form of the scheme is described that approximates the mixed layer Rossby radius as a function only of latitude. This formulation is more generally robust to instability and has a comparatively larger impact on the MLD than the original formulation, but yields qualitatively similar results. The global mean impact of the scheme on the MLD is found to be almost twice as large in 1° and 2° configurations of GO5 as it is in the 1/4° configuration. This is shown to be the result of the scheme overcompensating for the decay in strength of resolved mixed layer density fronts in this model with decreasing

*Corresponding author.
Email address: daley.calvert@metoffice.gov.uk (Daley Calvert)

horizontal grid resolution. The MLD criterion defining the depth scale of the scheme is shown to affect its global mean impact on the MLD by nearly a factor of 3 in GO5 and GC2, depending on whether the criterion is chosen to capture the actively mixing layer or the well-mixed layer. The parameterisation reduces the magnitude of deep MLD biases while increasing the magnitude of shallow biases. The globally averaged winter MLD bias is reduced from 17% to 9% of climatological values in GO5 but changes from +3% to −4% in GC2. Summer mixed layers are too shallow on average in both configurations and their average magnitude is increased by the parameterisation.

Keywords: parameterisation, surface mixed layer, oceanic boundary layer, mixed layer depth, submesoscale, NEMO

1. Introduction

The vertical transports of heat and transfers of momentum within the near-surface ocean mixed layer depend on many physical processes including: convective overturning (Marshall and Schott, 1999), wind-driven inertial oscillations with occasional shear spiking (Large and Crawford, 1995), Langmuir turbulence driven by interactions between surface waves and ocean currents (Craik and Leibovich, 1976; McWilliams et al., 1997), wind-driven reductions in stability (Thomas, 2005), and instabilities of the zonal flow (Haine and Marshall, 1998) involving baroclinic (Boccaletti et al., 2007), symmetric (Stone, 1970; Thomas et al., 2013), inertial (Griffiths, 2008) and mixed Rossby-inertia wave (Sakai, 1989) instabilities.

Many of these processes are not resolved at all in ocean general circulation models used for climate simulation. Their parameterisation is clearly a complex and demanding undertaking, particularly bearing in mind that the parameterisations need to vary appropriately as the motions that are resolved by the model vary with its grid spacing; that is they should be scale-aware (Bachman et al., 2017a; Pearson et al., 2017). Parameterisations of Langmuir turbulence (Li et al., 2019), symmetric instabilities (Bachman et al., 2017b) and shear spiking

(Jochum et al., 2013) are active areas of model development.

This paper focuses on a parameterisation scheme for baroclinic sub-mesoscale mixed layer eddies (SMLEs) (Fox-Kemper et al., 2011, hereafter FK11). This SMLE parameterisation (SMLEP) is cast in terms of an overturning streamfunction that advects ocean tracers within the mixed layer, acting to slump isopycnals and restratify the upper ocean. The main impact of the SMLEP is therefore a reduction of the mixed layer depth (MLD). This reduction differs across the simulations reported by FK11 and depends on the details of the SMLEP implementation itself. Nevertheless, their simulations with the SMLEP all demonstrate a reduction in the magnitude of deep MLD biases and an increase in that of shallow MLD biases, as has been observed in other models (Weijer et al., 2012; Bentsen et al., 2013; Swapna et al., 2015). Deep winter MLD biases associated with deep convection are reduced, although its representation in the North Atlantic remains poor (Heuzé, 2017), while shallow summer MLD biases common to several models are increased (Huang et al., 2014). Other impacts noted by FK11 include a substantial increase in strength of the Atlantic meridional overturning (Farneti et al., 2015) and a reduction in its variability (Danabasoglu et al., 2012), as well as a large impact on sea ice extent and thickness.

FK11 indicate that the SMLEP is sensitive to the details of its formulation, in particular the specifications of the MLD and characteristic width of fronts in the mixed layer. However, the sensitivity to these individual parameters is obfuscated in FK11 by the use of different ocean models as well as differences in the details of the SMLEP implementation. Furthermore, the SMLEP has a scale-aware aspect related to the ability of models of different resolution to represent horizontal buoyancy gradients, such that the vertical buoyancy fluxes associated with its overturning should be approximately independent of horizontal grid resolution. As the simulations of FK11 are based on 1° ocean grids and their use of eddy-permitting configurations is limited to a short $1/10^\circ$ uncoupled simulation, it is unclear whether this resolution-independence occurs in practice in ocean models.

This paper expands on the results of FK11 by examining the impact of the SMLEP on standard global configurations of the NEMO ocean model (Madec et al., 2017), focussing mainly on its impact on the MLD. Our primary objectives are to quantify the dependence of this impact on the details of the SMLEP formulation and on the horizonal grid resolution, and to uncover the reasons for these dependencies.

Section 2 describes the formulation of the SMLEP of FK11, its implementation in NEMO and an approximate formulation that we have used. Section 3 describes the simulations performed and the diagnostics used in the paper. Sections 4, 5 and 6 explore the dependence of the impact of the SMLEP on three aspects of its formulation: section 4 examines the dependence on the specification of the mixed layer Rossby radius used by the parameterisation, section 5 explores the dependence on the horizontal resolution of the model calling the parameterisation, and section 6 investigates the dependence on the density difference used by the parameterisation to define the MLD. Section 7 examines the impact of the SMLEP on MLD biases in our standard global configurations of NEMO. Section 8 summarises the main results.

2. The sub-mesoscale mixed layer eddy parameterisation (SMLEP)

2.1. Overview of the FK11 formulation

The basis of the FK11 SMLEP is that sub-mesoscale baroclinic instabilities within the ocean mixed layer transport light water upward and toward the dense water side of mesoscale fronts, thus releasing potential energy. These $\mathcal{O}$(100m-10km) instabilities are not resolved in present climate models so their net buoyancy transports need to be parameterised. Their vertical buoyancy transports are particularly important, as they compete with vertical mixing processes also present in the upper ocean. FK11 follow Gent and McWilliams (1990) in parameterising eddy effects via an overturning streamfunction, $\underline{\Psi}$, the additional velocity field $\underline{u}^*$ being given by

$$\underline{u}^* = \nabla \times \underline{\Psi}. \tag{1}$$

FK11 provide a formulation for $\underline{\Psi}$ in OGCMs (see their eq. (5)), which is based on a combination of physical arguments and submesoscale-resolving Large Eddy Simulations (LES). It can be written in the form:

$$\underline{\Psi} = \frac{C_e S}{L_f} \frac{H^2 \nabla \overline{b}^z \times \underline{\hat{z}}}{\sqrt{f^2 + \tau^{-2}}} \mu(z/H). \tag{2}$$

Here C_e is a non-dimensional efficiency coefficient estimated from LES to be between 0.06 and 0.08; S is a function of the horizontal grid spacing; L_f is the characteristic width of a mixed layer front; H is the mixed layer depth; $b = g(\rho_0 - \rho)/\rho_0$ is the buoyancy; $\nabla \overline{b}^z$ is the buoyancy gradient averaged over the mixed layer; $\underline{\hat{z}}$ is an upward pointing unit vector; f is the Coriolis parameter; z is height; and $\mu(z/H)$ is a vertical structure function which has a value of zero at the surface and beneath the mixed layer and one in the centre of the mixed layer. The term $\sqrt{f^2 + \tau^{-2}}^{-1}$ gives a time scale that tends to the inertial timescale $|f|^{-1}$ at mid-latitudes but near the equator reduces to a frictional slumping time scale $\tau \sim$ 1–20 days. The following paragraphs describe L_f and S in more detail.

FK11 represent L_f using a modified mixed layer Rossby radius (their eq. (13)),

$$L_f^m = \max(L_f^N, L_f^A, L_f^{\min}), \tag{3}$$

where L_f^N is the mixed layer Rossby radius defined in terms of the mixed layer average of the buoyancy frequency,

$$L_f^N = \frac{\overline{N}^z H}{\sqrt{f^2 + \tau^{-2}}}, \tag{4}$$

with lower limits imposed by the expected end state after geostrophic adjustment or mixing by symmetric instabilities,

$$L_f^A = \frac{|\nabla_H \overline{b}^z| H}{f^2}, \tag{5}$$

and by a parameter $L_f^{\min}$, both of which are used to ensure numerical stability. In (4), we have replaced f with $\sqrt{f^2 + \tau^{-2}}$ following the CCSM implementation of FK11.

The function S is defined by

$$\begin{aligned} S =& \Delta s, \quad \Delta s \leq L_u \\ =& L_u, \quad \Delta s > L_u, \end{aligned} \tag{6}$$

where Δs is the local horizontal grid spacing and L_u is an upper limit which is explained shortly. Subsection 2.1.3 of FK11 justifies the linear dependence of S on Δs in (6) by considering the vertical buoyancy flux induced by the SMLEP,

$$\overline{w'b'}_{\Psi} = \underline{\Psi} \times \nabla_H \overline{b}^z = \frac{C_e H^2}{\sqrt{f^2 + \tau^{-2}}} \frac{S}{L_f} |\nabla_H \overline{b}^z|^2. \tag{7}$$

The vertical buoyancy flux is the critical outcome of the parameterisation in governing its effects on mixed layer depth through competition with vertical mixing processes. Its magnitude should be approximately independent of Δs but the magnitude of the horizontal buoyancy gradient, $|\nabla_H \overline{b}^z|$, depends on the scales resolved by the model. FK11 show that $|\nabla_H \overline{b}^z|^2$ can be assumed to scale with Δs^{-1} over a range $L_f \ll \Delta s < L_u$, where $L_u = 1^\circ \approx 111$km is determined from the MESO simulations of Hallberg and Gnanadesikan (2006). They show this to be equivalent to a k^{-2} slope in horizontal power spectra of $\overline{b}^z$, found in both observational datasets and models. The S/L_f scaling term is thus expected to make the vertical buoyancy flux (7) and the impact of the SMLEP independent of resolution.

For convenience we will define

$$R \equiv S |\nabla_H \overline{\rho}^z|^2. \tag{8}$$

The parameterisation (2) with (6) is constructed in the expectation that R is approximately independent of model resolution.

The use of (3) to represent L_f is a significant source of uncertainty in the SMLEP. Firstly, while L_f^N has been found in observations to be a useful guide to L_f (Hosegood et al., 2006) under conditions of co-evolving mesoscale and submesoscale fronts and eddies, recent studies have questioned this finding. Callies and Ferrari (2018) find that L_f^N underestimates L_f by at least an order of magnitude when calculated over scales larger than L_f, so its use in (2) would over-estimate the vertical buoyancy flux. The simplifications exploited in (2) may not hold in if a more complete theory of the arrest scale of fronts is used instead of the deformation radius approximate scale (Sullivan and McWilliams, 2018; Bodner et al., 2020). Furthermore, the determination of $\overline{N}^z$ in models is very uncertain and depends on the choice of calculation method as well as the details of parameterisations used by the model. Due to the uncertainty in its definition and calculation, FK11 consider L_f to be a tuneable parameter that can be altered to reduce model bias.

Secondly, the S/L_f scaling term can result in excessively large $\underline{u}^*$ for coarse horizontal grids (large S) and weakly stratified mixed layers (very small L_f^N), which may cause numerical instability. This is prevented by limiting L_f via the $L_f^{\min}$ parameter in L_f^m, but this can significantly reduce the impact of the parameterisation. FK11 find for their 1° simulations that $L_f^{\min} = 5$km prevents $\underline{\Psi}$ from exceeding the overturning strength of the Atlantic meridional overturning circulation in volume transport (although the meridional heat and freshwater transports are much weaker), but note that choosing $L_f^{\min} = 1$km nearly doubles the reduction in global mean MLD by the parameterisation. The choice of $L_f^{\min}$ is therefore important and should be as small as possible while avoiding excessively large $\underline{u}^*$ and instability.

2.2. Implementation in NEMO

Here we describe the implementation of the FK11 SMLEP in NEMO, which is henceforth described as the "FK11 scheme".

The parameterisation is discretised on the Arakawa C-grid by calculating the streamfunction vector $\underline{\Psi}$ at velocity points (the zonal component at v points and the meridional component at u points). The mixed layer average buoyancy, $\overline{b}^z$, and MLD, H, are initially calculated at tracer points. H is calculated using a finite-difference potential density criterion, $\Delta\sigma_\theta$, with respect to the density at 10m depth. The buoyancy frequency is then defined as

$$\overline{N}^{z\,2} = \frac{g\Delta\sigma_\theta}{\rho_0 H}. \tag{9}$$

H and $\nabla\overline{b}^z \times \underline{\hat{z}}$ vectors are then calculated at velocity points; H as the minimum of its value at adjacent tracer points and $\nabla\overline{b}^z \times \underline{\hat{z}}$ as the horizontal gradient in $\overline{b}^z$. The induced velocity $\underline{u}^*$ is then added to the model velocities. As L_f^A is not required to guard against numerical instability (described in more detail below), it is disregarded in (3) so that L_f^m is a function of L_f^N and $L_f^{\min}$ only.

Unless explicitly specified to the contrary, the "standard" parameter values used with the FK11 scheme in this paper to obtain $\underline{\Psi}$ from (2), (3), (4), (6) and (9) are $C_e = 0.06$, $L_u = 111$km, $\tau = 2$ days, $L_f^{\min} = 5$km and $\Delta\sigma_\theta = 0.03\text{kgm}^{-3}$. The above specification of H is the optimal finite-difference criterion of de Boyer Montégut et al. (2004), which captures the depth of the well-mixed layer over timescales greater than a day. Their choice of a 10m reference depth was found to largely avoid the strong diurnal cycle of the mixing layer, while $\Delta\sigma_\theta = 0.01\text{kgm}^{-3}$ tended to capture the actively mixing layer after strong diurnal forcing and $\Delta\sigma_\theta = 0.05\text{kgm}^{-3}$ yielded H within the seasonal thermocline, rather than at its top. The choice $L_f^{\min} = 5$km is the largest value recommended by FK11 for their 1° simulations and is therefore conservative with regards to preventing numerical instability and excessive overturning strength.

We have found that the FK11 scheme integrates stably in all NEMO configurations described in this paper when $L_f^{\min}$ does not limit L_f^m. The following argument suggests that this is an inherent characteristic of our method for calculating $\overline{N}^z$. Substituting (9) into (4) one obtains

$$L_f^N = \frac{\sqrt{H}}{A\sqrt{f^2 + \tau^{-2}}}, \tag{10}$$

where $A^{-1} \equiv \sqrt{g\Delta\sigma_\theta/\rho_0} \approx 0.0169\text{m}^{1/2}\text{s}^{-1}$ is a constant. With $L_f = L_f^N$, substituting (10) into (2) gives

$$\underline{\Psi} = C_e A S H^{3/2} \left(\nabla \overline{b}^z \times \hat{\underline{z}}\right) \mu(z/H), \tag{11}$$

which is more evidently robust than (2) to excessively large $\underline{u}^*$, as it has no denominator that can become vanishingly small. Another interpretation is that $\overline{N}^z$, and therefore L_f^N, cannot become vanishingly small so L_f^{min} is not required to limit L_f^m. However, (11) does not guarantee that $\underline{u}^*$ will not be detrimental to the accuracy of mixed layer currents and overturning transports. L_f^{min} may therefore still be required to limit this impact. The degree to which L_f^{min} limits L_f^m has implications for the impact of the SMLEP on the mixed layer; this is examined further in section 4.

In the following subsection we present an alternative form of L_f that, like (10), cannot become vanishingly small and for which $\underline{\Psi}$ has a similar form to (11).

2.3. Approximate formulation

The calculation of $\overline{N}^z$ in NEMO has a useful property in that it cannot become vanishingly small. This will not necessarily be the case in other models and L_f may need to be artificially limited, as in (3), to avoid excessively large $\underline{u}^*$ and numerical instability. Furthermore as already mentioned, the calculation of $\overline{N}^z$ is generally very uncertain and dependent on the details of parameterisations used by the model, which can make it difficult to determine an appropriate value for L_f^{min}.

Other implementations of (2) circumvent this issue by representing L_f as a constant (Fox-Kemper et al., 2008a) or as a function of Δs only (see the CM2Gα implementation of Fox-Kemper et al., 2011). This approach of defining

L_f without a dependence on $\overline{N}^z$ is permitted by the present lack of theoretical and observational constraints on its value.

Here we introduce a further definition of L_f with this property, that we will refer to as the "approximate scheme". L_f is represented as a function of latitude approximating L_f^N,

$$L_f^a \sqrt{(f^2 + \tau^{-2})} = L_0 f_0 \equiv \frac{1}{B}, \tag{12}$$

where L_0 and f_0 are reference values of L_f^N and f respectively, and B^{-1} is equivalent to a constant value of $\overline{N}^z H$. After substituting (12) in (2), the latter can then be written as

$$\underline{\tilde{\Psi}} = C_e B S H^2 \left(\nabla \overline{b}^z \times \underline{\hat{z}} \right) \mu(z/H). \tag{13}$$

(13) is similar to (11) in that it is more robust than (2) to excessively large $\underline{u}^*$, but has the advantage that this property does not require a specific method for calculating $\overline{N}^z$. The main difference from the FK11 scheme is that the scaling term S/L_f is constant in time and does not depend on the model state, with B replacing $L_f^{\min}$ as the relevant tuneable parameter. Additionally, (13) is more sensitive to H ($\underline{\tilde{\Psi}} \propto H^2$) than (11) ($\underline{\Psi} \propto H^{3/2}$) as L_f no longer depends on H.

All terms in (13) that appear in (11) are defined in the same way. Unless explicitly specified to the contrary, the "standard" parameter values used with the approximate scheme to obtain $\underline{\tilde{\Psi}}$ from (6), (12) and (13) are $L_0 = 5$km and $f_0 = f(20°)$ (equivalent to $B^{-1} \approx 0.249\text{ms}^{-1}$) with other parameter values the same as for the FK11 scheme ($C_e = 0.06$, $L_u = 111$km and $\Delta\sigma_\theta = 0.03\text{kgm}^{-3}$). As for the standard value of $L_f^{\min}$ in the FK11 scheme, this value of B is a conservative choice with regards to preventing excessively strong overturning by the SMLEP. The resulting L_f^a profile, shown later in section 4, has a global minimum of roughly 1.7km and lies within the $1\text{km} \leq L_f^{\min} \leq 5\text{km}$ range used by FK11 for their 1° simulations, but is generally much larger than L_f^N as calculated using (10) except in deep mixed layers at high latitudes. B might be further adjusted to give an L_f^a profile more consistent with L_f^N and/or to improve model biases.

In section 4 we compare the impact of the approximate scheme (13) and its dependence on B with that of the FK11 scheme (11) and its respective tuneable parameter, $L_f^{\min}$.

3. Configurations and diagnostics

3.1. Configurations

We present results from simulations based on three standard global configurations of the NEMO ocean model; the GO5 (Megann et al., 2014) and GO6 (Storkey et al., 2018) "uncoupled" ocean-ice configurations, and the GC2 "coupled" atmosphere-ocean-ice-land configuration (Williams et al., 2015) in which the ocean component is GO5. These references should be consulted for a full description of each configuration, but some relevant details are described below.

GO5 and GO6 use the tripolar ORCA horizontal grids (Madec and Imbard, 1996), which have been extended further south in GO6 to permit the modelling of circulation beneath the Antarctic ice shelves but are otherwise identical. Both configurations use 75 vertical levels with refinement towards the surface and a partial step representation of bottom bathymetry (Barnier et al., 2006). The vertical coordinate is discretised on z-levels in GO5 and uses the variable volume (nonlinear free surface) z^* coordinate of Adcroft and Campin (2004) in GO6.

Lateral advection of momentum is formulated using an energy and enstrophy conserving vector-invariant scheme (Arakawa and Lamb, 1981) and lateral advection of tracers uses the FCT scheme of Zalesak (1979). The irrotational momentum advection term uses the Hollingsworth et al. (1983) correction in GO6 but not in GO5. Lateral diffusion of momentum is along geopotential surfaces and uses a Laplacian operator in the 1° and 2° simulations and a biharmonic operator in the 1/4° and 1/12° simulations. Lateral diffusion of tracers is along isoneutral surfaces and uses a Laplacian operator. The lateral viscosity and diffusion coefficients scale with the horizontal grid spacing following Willebrand et al. (2001). A parameterisation of adiabatic eddy mixing

(Gent and McWilliams, 1990) is used with a spatially-varying coefficient (Held and Larichev, 1996) in the 1° and 2° simulations. A free slip lateral boundary condition on momentum is used in all configurations except the 1/4° and 1/12° GO6 configurations, which use partial slip and no-slip boundary conditions respectively around the Antarctic coastline.

Vertical mixing uses a modified version of the Gaspar et al. (1990) TKE scheme (Madec et al., 2017) with representations of surface (Craig and Banner, 1994), near-inertial (Rodgers et al., 2014) and tidal (Simmons et al., 2004; Koch-Larrouy et al., 2008) wave breaking, double-diffusive mixing (Merryfield et al., 1999) and Langmuir turbulence (Axell, 2002). In the uncoupled simulations, a weak sea surface salinity restoration flux of -33.33 mm day^{-1}psu^{-1} towards monthly mean climatological values is applied. The climatological data used are Levitus et al. (1998) merged with PHC2.1 (Steele et al., 2001) in Arctic regions for GO5 and a 1995-2014 average of the EN4 monthly objective analysis (Good et al., 2013) for GO6.

The sea ice component in all configurations is CICE (Hunke and Lipscomb, 2010); specifically the GSI6.0 configuration (Rae et al., 2015) in GO5 and GC2, and the GSI8.1 configuration (Ridley et al., 2018) in GO6.

GO5 and GO6 have each been integrated as a traceable set of simulations with nominal horizontal resolutions of 1/4°, 1° and 2°, and 1/12°, 1/4° and 1° respectively. The main grid-dependent physical parameter settings for these simulations are given in table 1. Simulations of GC2 have been performed using an N96 grid for the UM atmosphere model (approximately 135km resolution in mid-latitudes) coupled to the 1/4° configuration of GO5.

The uncoupled GO5 and GO6 simulations are forced by the CORE2 surface forcing dataset (Large and Yeager, 2009) over the period 1976-2005. All simulations have been integrated for at least 30 years starting from rest, with temperature and salinity initialised from a 2004-2008 average of the EN3 monthly objective analysis (Ingleby and Huddleston, 2007) in GO5 and GC2, and a 1995-2014 average of the EN4 monthly objective analysis (Good et al., 2013) in GO6.

Simulations with the SMLEP use the approximate scheme and standard parameters described in subsection 2.3 unless specifically stated otherwise. The impact of the SMLEP on MLD biases is presented in section 7 for the GO5 and GC2 configurations using this SMLEP configuration. The sensitivity of the SMLEP MLD impact to the details of its formulation are presented in sections 4, 5 and 6 for the GO5 configuration, using this SMLEP configuration as a reference. We note that despite the different dependence on H, the sensitivities obtained using the approximate scheme are qualitatively similar to those obtained using the FK11 scheme. In section 6, we also present results for the GC2 configuration to indicate how the sensitivity of the SMLEP may differ when a coupled atmosphere model is used.

3.2. Diagnostics

We describe restratification of the mixed layer by the SMLEP in terms of the vertical buoyancy flux, (7), re-scaled as an equivalent upward heat flux across the mixed layer (Fox-Kemper et al., 2008b),

$$Q = \frac{C_p \rho_0}{g \alpha_T} \underline{\hat{z}} \cdot (\underline{\Psi} \times \nabla_H \bar{b}^z), \tag{14}$$

and its subsequent impact on the diagnosed MLD. It is important to distinguish between the MLD diagnostic used to evaluate this impact, H_d, and the MLD appearing in the expression for the SMLEP overturning streamfunction, H, described in section 2. H and H_d are not necessarily the same quantity. H may need to be consistent with vertical length scales used by other parameterisations within the model, for example the boundary layer depth used in mixed layer models such as KPP (Large et al., 1994) or Turner and Kraus (1967). Similarly H_d may need to be consistent with other models, in a model intercomparison project for example, or with observational datasets. Here we define H_d using a finite-difference density criterion $\Delta\sigma_\theta = 0.03\text{kgm}^{-3}$ calculated with respect to the density at 10m depth. This definition captures the depth of the seasonal mixed layer and is identical to that of de Boyer Montégut et al. (2004), allowing a direct comparison with their climatology. H_d is therefore consistent in definition with H for the standard choices of parameter settings in both the FK11 and approximate schemes, though H_d is calculated at tracer points and H at velocity points.

The diagnostics in this paper are calculated for each model time step and presented as 25-year monthly mean climatologies for years 6 to 30 of the simulations, unless otherwise stated. The mean, maximum and minimum values of these monthly mean climatologies at a given grid point are used to calculate an annual average, a seasonal maximum and a seasonal minimum. For H_d and Q, the seasonal maximum and minimum are considered to be representative of the local winter and summer respectively. For a quantity x these three temporal samplings are denoted by $\overline{x}$, $\overline{x}^{\max}$ and $\overline{x}^{\min}$ respectively. The impact of the SMLEP on H_d is expressed as a relative change

$$\Delta(H_d) = 100\frac{H_d^+ - H_d^-}{H_d^-}, \tag{15}$$

where the + and − superscripts respectively denote simulations with and without the SMLEP. Results are mainly shown for the impact on the seasonal maximum, $\Delta(\overline{H_d}^{\max})$, as the parameterisation tends to have the largest impact during the local winter when MLDs are deepest.

4. Dependence on the specification of L_f

In this section we examine the impact of the SMLEP on the MLD when using the FK11 (subsections 2.1, 2.2) and approximate (subsection 2.3) schemes in the 1/4° uncoupled GO5 configuration. These schemes differ in their specification of the mixed layer frontal width L_f; the following results therefore indicate the sensitivity of the SMLEP impact to this specification.

Both schemes represent L_f using a form of the mixed layer Rossby radius, L_f^N. The FK11 scheme uses a modified form, L_f^m (3), which is the minimum of L_f^N and a parameter $L_f^{\min}$, and the approximate scheme uses an approximation, L_f^a (12), which is a function of latitude and a parameter B. Simulations have been performed using two parameter settings for each scheme; the FK11 scheme simulations with $L_f^{\min} = 5000$m and $L_f^{\min} = 200$m, and the approximate scheme simulations with B defined using $L_0 = 5000$m and $f_0 = f(20°)$, $f_0 = f(10°)$. Figure 1, panel (a) shows 25-year zonal averages of L_f calculated using L_f^m and L_f^a with these parameter settings and L_f^N calculated using (10). $L_f^{\min} = 5000$m and $f_0 = f(20°)$ are chosen as conservative upper bounds on L_f, as described in subsection 2.3. $L_f^{\min} = 5000$m (blue line) significantly limits L_f with respect to L_f^N outside of the tropics, while the L_f profile of $f_0 = f(20°)$ (green line) is generally much larger than that of L_f^N, particularly in the tropics. $L_f^{\min} = 200$m and $f_0 = f(10°)$ are chosen as lower bounds for which L_f is more consistent with L_f^N. $L_f^{\min} = 200$m (orange line) does not limit L_f with respect to L_f^N, as H has a lower limit imposed by the reference depth of the finite-difference criterion, 10m, so that $L_f^N \geq\sim 367$m. The L_f profile of $f_0 = f(10°)$ (red line) is similar to that of L_f^N and is globally a better approximation than $f_0 = f(20°)$, which is consistent with L_f^N only in deep mixed layers at high latitudes but is still much larger in the tropics.

Figure 2 shows the impact of the FK11 and approximate schemes on the MLD seasonal maximum, $\Delta(\overline{H_d}^{\max})$ as defined in subsection 3.2, while figure 1, panel (b) shows zonal averages of $\Delta(\overline{H_d}^{\max})$. There is a great deal of geographical coherence in $\Delta(\overline{H_d}^{\max})$ between the simulations and both approximate scheme simulations give qualitatively similar results to the FK11 scheme simulation using $L_f^{\min} = 200$m. In the FK11 simulations, $\Delta(\overline{H_d}^{\max})$ is up to a factor of 2 larger in subtropical and mid-latitudes and up to a factor of 4 larger in the Arctic when using $L_f^{\min} = 200$m than when using $L_f^{\min} = 5000$m. These results are consistent with those of FK11, who report for their 1° uncoupled simulations that the global mean MLD decreases by nearly the same amount between simulations using $L_f^{\min} = 1000$m and $L_f^{\min} = 5000$m, as between a simulation using $L_f^{\min} = 5000$m and one without the SMLEP. In the approximate scheme simulations, $\Delta(\overline{H_d}^{\max})$ is around a factor of 1.5 larger between the equator and mid-latitudes when using $f_0 = f(10°)$ than when using $f_0 = f(20°)$.

Most of the differences in $\Delta(\overline{H_d}^{\max})$ between the simulations can be explained in terms of their specification of L_f, as this is the only way in which they differ. This is shown by the zonal averages of L_f and $\Delta(\overline{H_d}^{\max})$ in figure 1, panels (a) and (b) respectively. A decrease in L_f implies an increase in Q, the vertical buoyancy flux re-scaled as an equivalent heat flux across the mixed layer, and therefore stronger restratification of the mixed layer by the SMLEP. Outside of the tropics, L_f is significantly larger for $L_f^{\min} = 5000$m (blue line) than for $L_f^{\min} = 200$m (orange line) and therefore reduces the impact of the FK11 scheme. Similarly, the $f_0 = f(20°)$ profile of L_f (green line) is globally larger than that of $f_0 = f(10°)$ (red line) and reduces the impact of the approximate scheme.

At high-latitudes the response of $\Delta(\overline{H_d}^{\max})$ to changes in L_f is more nonlinear. Large $\Delta(\overline{H_d}^{\max})$ are found to coincide with changes in sea ice concentration, suggesting that feedbacks between the SMLEP and sea ice are important. This is particularly evident in the Southern Ocean. South of 75°S the two approximate scheme simulations have a similar impact, while the $L_f^{\min} = 5000$m simulation has a larger impact than the $L_f^{\min} = 200$m simulation. Between 60 - 65°S the SMLEP instead acts to increase $\overline{H_d}^{\max}$, with larger L_f acting to reduce this impact. This positive $\Delta(\overline{H_d}^{\max})$ is evident in figure 2 and is attributed to a spurious polynya in the Weddell Sea in GO5 (Megann et al., 2014), which is known to be sensitive to details of the parameterised vertical mixing (Heuzé et al., 2015). While these feedback mechanisms would benefit from further study, they are not the focus of the present analysis and are not discussed further.

The average L_f profiles for the simulations using the approximate scheme with $f_0 = f(10°)$ and FK11 scheme with $L_f^{\min} = 200$m are very similar outside the tropics, but $\Delta(\overline{H_d}^{\max})$ is generally larger for the $f_0 = f(10°)$ simulation. Within the tropics, L_f is larger in the $f_0 = f(10°)$ simulation than in the $L_f^{\min} = 200$m simulation but $\Delta(\overline{H_d}^{\max})$ is similar. Differences in $\Delta(\overline{H_d}^{\max})$ between these simulations are better explained in terms of the scaling of Q with H in the two SMLEP formulations. Q scales as BH^2 in the approximate scheme (13) and scales as $AH^{3/2}$ in the FK11 scheme (11). Figure 3 shows that Q is larger for the approximate scheme than for the FK11 scheme (orange line) for

$H \geq\sim 55$m and $H \geq\sim 215$m when B is defined using $f_0 = f(10°)$ (red line) and $f_0 = f(20°)$ (green line) respectively. Therefore except in the tropics where H is small, mixed layer restratification by the SMLEP is expected to be stronger in the approximate scheme with $f_0 = f(10°)$ than in the FK11 scheme with $L_f^{\min} = 200$m.

The Q scalings indicate that shallow MLDs limit the sensitivity of the SMLEP to changes in L_f. Panel (a) of figure 1 shows that L_f is larger than L_f^N in the tropics for both profiles used by the approximate scheme, by a factor of 1.5 for the $f_0 = f(10°)$ profile and by a factor of 3 for the $f_0 = f(20°)$ profile. The corresponding profiles of $\Delta(\overline{H_d}^{\max})$ in panel (b) are only slightly smaller than those for the FK11 scheme simulations. This suggests that further adjustment of these profiles to better approximate L_f^N in the tropics would have only a small impact on the MLDs here.

In the following two sections we explore the sensitivity of the SMLEP MLD impact to parameters other than L_f, using the approximate scheme with the standard parameters $L_0 = 5000$m and $f_0 = f(20°)$. The results presented in this section might suggest that these sensitivities should differ from those obtained using the FK11 scheme, but additional experiments with the FK11 scheme are found to yield qualitatively similar sensitivities.

5. Dependence on horizontal resolution

In this section we examine the impact of the SMLEP on the MLD when using the approximate scheme (13) in the 1/4°, 1° and 2° uncoupled GO5 configurations. The following results therefore indicate the sensitivity of the SMLEP as a function of the horizontal grid scale Δs.

Table 2 shows the global mean impact of the SMLEP on the MLD in the GO5 simulations. The annual mean MLD impact, $\Delta(\overline{H_d})$, is almost twice as large in the 1° and 2° simulations as it is in the 1/4° simulation. This is dominated by differences in the winter ($\Delta(\overline{H_d}^{\max})$); differences in the summer ($\Delta(\overline{H_d}^{\min})$) are much smaller and are negligible between the 1° and 2° simulations. Panel (a) of figure 4 shows that within $\pm 45°$ of the equator, $\Delta(\overline{H_d}^{\max})$ is much larger in the $1°$ simulation than in the $1/4°$ simulation but of a similar magnitude in the $2°$ simulation. Panels (a) and (b) of figure 5 show that the large-scale spatial characteristics of $\Delta(\overline{H_d}^{\max})$ in this region are very similar between the simulations (not shown for the $2°$ simulation) but are often substantially different poleward of $\pm 45°$, most notably in the Southern Ocean and North Atlantic.

Restratification of the mixed layer by the SMLEP vertical buoyancy flux, $\overline{w'b'}_\Psi$, is shown via (7), (8) and (13) to depend on R, the mixed layer horizontal density gradient multiplied by Δs, and H^2, the square of the MLD ($H^{3/2}$ when using the FK11 scheme (11) instead of the approximate scheme (13)). It is nontrivial to directly relate differences in these quantities between simulations to those in the MLD impact of the SMLEP shown in figures 4 and 5. Instead we examine the more straightforward relationship between R, H^2 and restratification by the SMLEP expressed as an upward heat flux, Q (14).

Figures 4 and 5 show the seasonal maximum of Q, $\overline{Q}^{\max}$. These results are consistent with the observational estimates of Johnson et al. (2016) for both the $1/4°$ and $1°$ simulations, although the modelled fluxes are much larger in some regions e.g. the south Indian Ocean. $\overline{Q}^{\max}$ generally corresponds well with that of the MLD impact of the SMLEP, $\Delta(\overline{H_d}^{\max})$, and within $\pm 45°$ of the equator is similarly larger in magnitude in the $1°$ simulation than in the $1/4°$ simulation. In some regions, $\overline{Q}^{\max}$ does not correspond well with $\Delta(\overline{H_d}^{\max})$. Around Antarctica, $\overline{Q}^{\max}$ is small but $\Delta(\overline{H_d}^{\max})$ is large. This was identified in section 4 as a region where feedbacks between the SMLEP and sea ice are important in determining $\Delta(\overline{H_d}^{\max})$; H_d is therefore indirectly affected by the SMLEP via its impact on the sea ice. In close proximity to the Amazon outflow, $\overline{Q}^{\max}$ is large but $\Delta(\overline{H_d}^{\max})$ is small. In this case a large horizontal density gradient coincides with a shallow, strongly stratified mixed layer; H_d is already small and further restratification by the SMLEP has little impact.

GO5 simulations using the FK11 scheme with $L_f^{\min} = 200$m produce qualitatively similar results to those in figures 4, 5 and table 2. The FK11 scheme produces larger $\overline{Q}^{\max}$ and $\Delta(\overline{H_d}^{\max})$ than the approximate scheme, as shown in section 4, but their increase in magnitude between the 1/4° and 1° simulations is robust to the choice of SMLEP formulation. In particular, the global averages of $\Delta(\overline{H_d})$ and $\Delta(\overline{H_d}^{\max})$ in table 2 remain almost twice as large in the 1° simulation as in the 1/4° simulation when using the FK11 scheme.

As described in subsection 2.1, the FK11 scheme assumes that R is approximately independent of Δs. Figure 6 shows instantaneous zonal averages of R from GO5 and GO6 simulations that do not use the SMLEP, calculated for the south-east Pacific region (100 - 160°W) used by FK11 in a similar analysis (their figure 2). It is evident that in this region R is not independent of Δs in either set of simulations. Between 35 - 50°S, R is on average larger by factors of around 5 and 2 in the 1° and 1/4° GO6 simulations compared to the 1/12° GO6 simulation, and by a factor of 3 in the 1° GO5 simulation compared to the 1/4° GO5 simulation. The magnitude of R is more similar between the 1° and 2° GO5 simulations, suggesting that it is correct to apply an upper limit of $L_u = 1°$ to S in (6).

Figure 7 shows maps of instantaneous R from the 1/12°, 1/4° and 1° GO6 simulations. Over large scales, the spatial characteristics of R are similar in all simulations but the magnitude clearly increases with Δs, consistent with figure 6. Similar results are found for equivalent data from the GO5 simulations, where R is again found to be qualitatively similar in the 1° and 2° simulations. Similar results to those presented in figures 6 and 7 are also obtained from GO5 simulations using the SMLEP.

The dependence of R on Δs implies that the strength of resolved mixed layer density fronts, $|\nabla_H \overline{\rho}^z|^2$, does not scale as Δs^{-1} and that in turn, horizontal power spectra of $\overline{\rho}^z$ do not scale with the wavenumber as k^{-2} (Fox-Kemper et al., 2011). Power spectra for the 1/12° and 1/4° simulations (figure 8) using the same data and region as in figure 6 show that $\overline{\rho}^z$ transitions with increasing k from a k^{-2} to k^{-4} scaling at around $k \equiv 2°$. This result, also found for other months and years of the simulations, implies that the resolved mesoscale fronts shown in figure 7 are weaker than expected. The S/L_f scaling therefore overcompensates for the decrease in resolved frontal strength when these fronts are unresolved by the grid; Q is generally larger in the $1/4°$ and $1°$ simulations than it is in the $1/12°$ simulation.

In order to quantitatively establish the relative contributions of differences in R and H^2 to differences in Q between the simulations, we use (13) in (14) to infer for each of the simulations that

$$Q = Q_0 R H^2 \tag{16}$$

where Q_0 is a constant. If a quantity x is represented in two different simulations as x_a and x_b and its difference as $\delta x = x_a - x_b$, simple algebra shows that

$$\delta Q = \delta Q_H + \delta Q_R, \quad \delta Q_H \equiv Q_0 R_b \delta(H^2), \quad \delta Q_R \equiv Q_0 H_a^2 \delta R. \tag{17}$$

In the following analysis δQ_H and δQ_R have been calculated offline from monthly mean diagnostics.

Figure 9 shows maps of 25-year averages of the terms in (17), $\overline{\delta Q}$, $\overline{\delta Q_H}$ and $\overline{\delta Q_R}$, calculated for the $1°$ GO5 simulation with respect to the $1/4°$ simulation. Except at high latitudes, $\overline{\delta Q_R}$ corresponds well with $\overline{\delta Q}$ and generally has the opposite sign to $\overline{\delta Q_H}$; $\overline{\delta Q_R}$ therefore generally more than accounts for $\overline{\delta Q}$ in these regions. In the North Atlantic and around Antarctica, $\overline{\delta Q_H}$ has the same sign as $\overline{\delta Q}$ and the opposite sign to $\overline{\delta Q_R}$; here $\overline{\delta Q}$ tends to be dominated by $\overline{\delta Q_H}$. This negative correlation between $\overline{\delta Q_H}$ and $\overline{\delta Q_R}$ is consistent with the results of Johnson et al. (2016). Qualitatively similar results are obtained when this analysis is repeated using the $1/4°$ and $2°$ simulations.

Given that $\overline{Q}^{\max}$ and $\Delta(\overline{H_d}^{\max})$ are generally well correlated except at high latitudes (figure 5), the differences in $\Delta(\overline{H_d}^{\max})$ between the $1/4°$ and $1°$ simulations can be attributed to changes in R in these regions. Hence, except at high latitudes, the larger magnitude of R in the coarser resolution GO5 simulations results in stronger restratification of the mixed layer by the SMLEP than in the $1/4°$ simulation.

6. Dependence on specification of the MLD criterion

In this section we examine the impact of the SMLEP on the MLD when using the approximate scheme (13) in the $1/4°$ uncoupled GO5 and coupled GC2 configurations with two different specifications for the MLD parameter, H. The following results therefore indicate the sensitivity of the SMLEP to the specification of H.

In order to quantify the sensitivity in a straightforward way, we use the finite-difference criterion described in subsection 2.2 for both specifications of H and alter only the $\Delta\sigma_\theta$ value. Simulations have been performed with $\Delta\sigma_\theta = 0.01\mathrm{kgm}^{-3}$ in addition to those with the standard value of $\Delta\sigma_\theta = 0.03\mathrm{kgm}^{-3}$. We will denote these two criteria by $\Delta\sigma_\theta = 0.01$ and $\Delta\sigma_\theta = 0.03$ respectively. The former criterion tends to capture the depth of the actively mixing layer in strongly stratified profiles (de Boyer Montégut et al., 2004), which could be considered a lower bound on the definition of H. Figure 10 shows that this criterion produces MLDs that are shallower than the $\Delta\sigma_\theta = 0.03$ criterion by more than 100m in the Labrador and Greenland-Iceland-Norwegian (GIN) Seas, and in the Ross and Weddell Sea sectors.

As described in subsection 3.2 we distinguish H, which appears in the expression for the SMLEP overturning streamfunction (13), from the MLD diagnostic used to evaluate the impact of the SMLEP, H_d. $\Delta_{\sigma_0} H_d$ is used to denote differences in H_d between the simulation with the SMLEP using $\Delta\sigma_\theta = \sigma_0$ and the control simulation without the SMLEP.

Panels (a) and (c) of figure 11 show the impact of the SMLEP on the MLD in the GO5 simulations using $\Delta\sigma_\theta = 0.01$ and $\Delta\sigma_\theta = 0.03$, i.e. $\Delta_{0.01}(\overline{H_d}^{\max})$ and $\Delta_{0.03}(\overline{H_d}^{\max})$, while panel (e) shows their difference and illustrates the sensitivity of $\Delta(\overline{H_d}^{\max})$ to the criterion used to determine H. Panels (b), (d) and (f) respectively show the same quantities for the GC2 simulations.

$\Delta_{0.01}(\overline{H_d}^{\max})$ is generally small and negative in the GO5 simulation and larger with both positive and negative values in the GC2 simulation. These differences between the GO5 and GC2 simulations indicate the effect of coupled atmosphere-ocean feedbacks on the impact of the SMLEP, which was also noted by FK11. Around Antarctica $\Delta(\overline{H_d}^{\max})$ is much smaller in the GC2 simulation than in the GO5 simulation, as the extent of Antarctic sea ice is much smaller in GC2 and the effects of ocean-sea ice feedbacks, such as the spurious Weddell Sea polynya, are less prevalent than in GO5. In the GO5 simulations, $\Delta_{0.03}(\overline{H_d}^{\max})$ has qualitatively similar spatial characteristics to $\Delta_{0.01}(\overline{H_d}^{\max})$ but is much larger in magnitude. This increase in magnitude (panel (e) of figure 11) is larger than the magnitude of $\Delta_{0.01}(\overline{H_d}^{\max})$ (panel (a) of figure 11). In the GC2 simulations, $\Delta_{0.03}(\overline{H_d}^{\max})$ is more uniformly negative than $\Delta_{0.01}(\overline{H_d}^{\max})$, suggesting that increased restratification by the SMLEP competes more favorably with the impact of coupled feedbacks on the mixed layer.

Table 3 lists the global mean values of the impact of the SMLEP on H_d for both $\Delta\sigma_\theta$ criteria in the GO5 and GC2 simulations. The annual mean impact, $\Delta(\overline{H_d})$, is nearly a factor of 3 larger when using the $\Delta\sigma_\theta = 0.03$ criterion than when using the $\Delta\sigma_\theta = 0.01$ criterion. This is also the case for the local winter ($\Delta(\overline{H_d}^{\max})$) and summer ($\Delta(\overline{H_d}^{\min})$) impacts, except that the latter differs by less than a factor of 2 in the GC2 simulations. The similar global mean sensitivity to $\Delta\sigma_\theta$ in the GO5 and GC2 simulations again suggests that in the GC2 simulation using $\Delta\sigma_\theta = 0.03$, increased restratification by the SMLEP competes more favourably with the impact of coupled atmosphere-ocean feedbacks on the mixed layer.

Figure 12 shows fields of $\overline{Q}^{\max}$ calculated using only the last 5 years of the simulations. $\overline{Q}^{\max}$ is qualitatively similar in the GO5 and GC2 simulations using $\Delta\sigma_\theta = 0.01$ (panels (a) and (b)) and is only slightly larger in the respective simulations using $\Delta\sigma_\theta = 0.03$ (panels (c) and (d)). These relatively small changes in $\overline{Q}^{\max}$ contrast with the relatively large changes in $\Delta(\overline{H_d}^{\max})$ shown in panels (a), (b), (e) and (f) of figure 11 and with the factor of 3 increase in the global average of $\Delta(\overline{H_d}^{\max})$ shown in table 3. Q, as defined by (14), is the vertical buoyancy flux of the SMLEP expressed as an equivalent upward heat flux across the mixed layer and is proportional to H^2. As changes in H are relatively large (figure 10, panel (b)) but changes in Q are relatively small (figure 12, panels (c) and (d)), the vertical distribution of the buoyancy fluxes, determined by $\mu(z/H)$ in (2), is evidently more important than the amount of buoyancy transported by the fluxes across the mixed layer, represented by Q. The sensitivity of $\Delta(\overline{H_d}^{\max})$ to $\Delta\sigma_\theta$ is therefore driven by changes in the depth of overturning, rather than changes in the rate of mixed layer restratification by the overturning.

1° GO5 simulations using the FK11 scheme with $L_f^{\min} = 200$m produce qualitatively similar results to those in figures 11, 12 and table 3. $\overline{Q}^{\max}$ and $\Delta(\overline{H_d}^{\max})$ are generally larger in the FK11 scheme than in the approximate scheme, as shown in section 4, and are generally larger in the 1° simulations than in the 1/4° simulations, as shown in section 5. $\overline{Q}^{\max}$ decreases in magnitude with $\Delta\sigma_\theta$ in these simulations with the FK11 scheme, in contrast with the relatively small increase in magnitude in the simulations with the approximate scheme, shown in figure 12, panel (c). However, the global averages of $\Delta(\overline{H_d})$ and $\Delta(\overline{H_d}^{\max})$ in table 3 remain nearly a factor of 3 larger in the simulation using $\Delta\sigma_\theta = 0.03$ than in the simulation using $\Delta\sigma_\theta = 0.01$. This further supports the sensitivity of $\Delta(\overline{H_d}^{\max})$ to $\Delta\sigma_\theta$ being driven mainly by changes in the overturning depth.

As for the other parameter sensitivity results presented in this paper, the impact of the SMLEP on the MLD will depend on the criterion used to define H_d. This choice is particularly important when interpreting the present results, given that the impact on the MLD depends mainly on the depth of overturning by the SMLEP. When H is defined using a $\Delta\sigma_\theta = 0.01$ criterion, the depth of overturning is shallower than the MLD as defined by H_d using a $\Delta\sigma_\theta = 0.03$ criterion. In this case only part of the diagnosed mixed layer is restratified and the impact of the SMLEP on the MLD is reduced, increasing the apparent sensitivity of the impact to H. When H_d is instead calculated using a $\Delta\sigma_\theta = 0.01$ criterion, both H criteria overturn the full diagnosed mixed layer and the factor of 3 difference in the global averages of $\Delta(\overline{H_d})$ and $\Delta(\overline{H_d}^{\max})$ in table 3 reduces to slightly less than a factor of 2.

The qualitative similarity of $\overline{Q}^{\max}$ and its lack of response to changes in H

in the GO5 and GC2 simulations in figure 12 contrasts with the differing spatial characteristics of $\overline{H_d}^{\max}$ in figure 11. This suggests that if the other parameter sensitivities examined in this paper were also determined for the coupled configuration, the vertical buoyancy fluxes induced by the SMLEP might be similarly affected, but would likely yield different MLD impact sensitivities due to the indirect response of the coupled system to their changes.

7. Impact of the SMLEP on MLD biases

In this section we examine the impact of the SMLEP on MLD biases in the 1/4° uncoupled GO5 and coupled GC2 configurations, using the approximate scheme (13) with the standard parameters described in subsection 2.3. The MLD biases are calculated using the climatological dataset of de Boyer Montégut et al. (2004), updated to include ARGO data to September 2008 and using a criterion consistent with that of H_d described in subsection 3.2.

Panels (a), (c), (e) and (g) of figure 13 show maps of winter (seasonal maximum) and summer (seasonal minimum) MLD biases for the GO5 and GC2 simulations without the SMLEP. Global averages and standard deviations of these biases are shown in table 4.

Winter MLDs ($\overline{H_d}^{\max}$) in GO5 and GC2 (panels (a) and (c) of figure 13 respectively) are too deep in regions of deep water formation in the North Atlantic, most notably in the Labrador and GIN Seas. In GO5 winter MLDs in the Ross and Weddell Sea sectors are too deep, but in GC2 they are too shallow throughout the Southern Ocean. Table 5 shows that the magnitude of these regional biases is large when compared to the global averages in table 4. Globally, winter MLDs are too deep in GO5 (22m) and too shallow in GC2 (-6m). The average magnitude of biases in the North Atlantic is generally much larger in GO5 than in GC2: 518m for GO5 and 255m for GC2 in the Labrador Sea, 162m for GO5 and -29m for GC2 in the Greenland Sea. In the Ross and Weddell Sea sectors, the deep MLD biases in GO5 are between 200m and 450m while the shallow MLD biases in GC2 are between -35m and -70m. This difference is particularly remarkable in the Weddell Sea, where the average winter MLD is more than 3 times the climatological value in GO5 but less than 70% of the climatological value in GC2. This very deep MLD bias in GO5 corresponds to the spurious polynya described in section 4.

Summer MLDs ($\overline{H_d}^{\mathrm{min}}$) in GO5 and GC2 (panels (e) and (g) of figure 13 respectively) are generally too shallow, particularly in the Southern Ocean. The average magnitude of the summer MLD bias is larger in GC2 (-4m) than in GO5 (-2m).

Panels (b), (d), (f) and (h) of figure 13 show maps of the change in magnitude of the MLD bias for the winter ($\Delta(\overline{H_d}^{\mathrm{max}})$) and summer ($\Delta(\overline{H_d}^{\mathrm{min}})$) when the SMLEP is used in the GO5 and GC2 simulations. When compared with the corresponding maps of MLD bias in panels (a), (c), (e) and (g) it is evident that the SMLEP generally reduces the magnitude of deep biases and increases the magnitude of shallow biases, and that this change is proportional to the magnitude of the bias. This is expected given that the SMLEP acts to systematically reduce the MLD and has a larger impact on deep mixed layers.

Table 4 shows that the SMLEP tends to reduce the magnitude of the global average MLD bias when it is positive (too deep) and increase the magnitude when it is negative (too shallow). The SMLEP more than halves the global average winter bias in GO5, from 17% to 9% of climatological values, but in GC2 this changes from +3% to -4%. The standard deviation of winter biases is reduced in both simulations with the SMLEP, from 193m to 180m in GO5 and from 117m to 98m in GC2. The impact on the global average annual mean ($\overline{H_d}$) bias is smaller but qualitatively similar to that on the winter bias, decreasing from 10% to 3% of climatological values in GO5 but increasing from -2% to -7% in GC2. The global average summer bias increases in both simulations with the SMLEP, from -3% to -4% of climatological values in GO5 and from -9% to -11% in GC2. The standard deviation of summer biases slightly increases in GC2, from 9m to 10m, but changes by very little in GO5.

Table 5 shows that the impact on the regional winter biases is similarly dependent on model configuration. In the Labrador Sea the deep MLD bias is more than halved in GC2, from 62% to 15% of climatological values, but is reduced by less in GO5, from 185% to 157%, and remains substantial. In the Greenland Sea the deep MLD bias in GO5 is reduced, from 54% to 30% of climatological values, but the shallow MLD bias in GC2 is increased, from 2% to −25%. In the Ross Sea sector the deep MLD bias in GO5 is reduced by nearly 70%, from 161% to 50% of climatological values, but the increase in magnitude of the shallow MLD bias in GC2 is less than 4% of climatological values. In the Weddell Sea sector the deep MLD bias in GO5 and the shallow MLD bias in GC2 both increase in magnitude, but the increase in GO5 is much larger, from 362% to 421% of climatological values, while that in GC2 is close to 2% of climatological values. This large increase in GO5 is related to the impact of the SMLEP on the Weddell Sea polynya, described in section 4.

The SMLEP was developed with the aim of representing one of the physical processes active in the restratification of deep winter mixed layers. This physical basis, the overall reduction of deep climatological MLD biases and the reduction in the global standard deviation of the biases show that the SMLEP improves NEMO's representation of the near-surface ocean and is an important parameterisation to include in the model.

8. Concluding summary and discussion

The impact of the sub-mesoscale mixed layer eddy parameterisation (SMLEP) of Fox-Kemper et al. (2011) (hereafter FK11) on the mixed layer depth (MLD) is examined in global "uncoupled" ocean-ice (GO5) and "coupled" atmosphere-ocean-ice-land (GC2) configurations that use the NEMO ocean model. Specifically, we explore the sensitivity of this impact to three aspects of the parameterisation.

The first aspect is the specification of the characteristic width of a mixed layer front, L_f. Two methods for specifying L_f are examined using $1/4°$ simulations of GO5. The first method is a time-varying specification referred to as the "FK11 scheme", where L_f is calculated following FK11 as the minimum of the mixed layer Rossby radius L_f^N and a parameter $L_f^{\min}$. The second method is a time-fixed specification referred to as the "approximate scheme" where L_f^N is approximated as a function of latitude and two constants, L_0 and f_0. This latter specification is used for all other simulations with the SMLEP discussed in this paper.

The impact of the SMLEP on the MLD is found to be sensitive to the details of both L_f specifications. In subtropical and mid-latitudes the impact of the FK11 scheme is reduced by up to a factor of 2 when the limit on L_f is $L_f^{\min} = 5000$m, instead of $L_f^{\min} = 200$m which does not limit L_f^N. For the approximate scheme two profiles are used with the parameters $L_0 = 5000$m and $f_0 = f(20°)$ or $f_0 = f(10°)$. The impact of this scheme is around a factor of 1.5 less between the equator and mid-latitudes when the $f(20°)$ profile is used instead of the $f(10°)$ profile. However, the impact of both profiles is qualitatively similar to that of the FK11 scheme using $L_f^{\min} = 200$m.

The $f(10°)$ L_f profile of the approximate scheme closely approximates the 25-year zonal average of L_f^N, while the $L_f^{\min} = 200$m L_f profile of the FK11 scheme is identical to L_f^N. For these respective profiles the approximate scheme has a generally larger impact on the MLD than the FK11 scheme outside of the tropics. This is shown to be because the approximate scheme has a stronger dependence on the MLD than the FK11 scheme.

In the FK11 scheme it is desirable for $L_f^{\min}$ to be as small as possible while avoiding excessively strong overturning by the SMLEP, which can result in numerical instability. The 1/4°, 1° and 2° simulations of GO5 using the FK11 scheme are found to be numerically stable when L_f is not limited by $L_f^{\min}$ (and therefore equal to L_f^N). Subsection 2.2 argues that this is a property of the method used to calculate the mixed layer buoyancy frequency in NEMO and is not necessarily transferrable to other models. The approximate scheme introduced in this paper is generally robust to instability due to its use of a fixed profile for L_f. Although neither formulation is numerically unstable in our simulations, it is still possible for the SMLEP overturning to be excessively strong and detrimental to the accuracy of mixed layer currents and overturning transports. While we have not examined the impact of the SMLEP overturning on mixed layer currents in detail, the impact on annual average currents was found to be very small.

Both the FK11 and approximate schemes estimate L_f as the mixed layer Rossby radius L_f^N, but in practice L_f is not well-constrained and recent studies have questioned this assumption (Callies and Ferrari, 2018). The profile of L_f used by the approximate scheme might therefore be further adjusted to reduce model bias as suggested by FK11. Our sensitivity results suggest that except in the tropics, the MLD impact of the SMLEP is sensitive to the specification of L_0 and f_0. However, determining an optimal set of values that minimise MLD biases is beyond the scope of this paper.

The second aspect of the parameterisation explored here is the dependence on the local horizontal grid spacing of the model calling the parameterisation, Δs. The SMLEP is constructed so that the vertical buoyancy flux induced by the parameterisation is to be approximately independent of Δs over large horizontal scales. A scaling term S/L_f (see (2) and (6)) is implemented to achieve this, where $S \propto \Delta s$ and $R \equiv S|\nabla_H \overline{\rho}^z|^2$ is approximately independent of Δs over large scales. This assumes that $|\nabla_H \overline{\rho}^z|^2$ will scale as Δs^{-1}, which is consistent with horizontal power spectra of $\overline{\rho}^z$ scaling with the wavenumber as k^{-2} as found in observations and models with adequate representation of the mesoscale. In their analysis of the regional MESO simulations of Hallberg and Gnanadesikan (2006) FK11 find this assumption to be valid for $\Delta s \leq 1^\circ$. In our uncoupled NEMO simulations spanning $1/12^\circ \leq \Delta s \leq 2^\circ$ we find that R is proportional to Δs, although limiting Δs to a maximum of 1° as imposed by FK11 is sufficient to make R independent of Δs in the 1° and 2° simulations. Horizontal power spectra of $\overline{\rho}^z$ for these simulations scale as k^{-4} for k larger than approximately $k \equiv 2^\circ$, suggesting that the smallest resolved mesoscale density fronts are weaker than expected. The assumptions underlying the S/L_f scaling term are therefore violated in this model and the term does not correctly rescale $|\nabla_H \overline{\rho}^z|^2$ to account for grid spacing, such that R is proportional to Δs. This drives an increase in mixed layer restratification by the SMLEP on coarser model grids, so that the global mean impact on the MLD in the 1° simulation is nearly twice that in the 1/4° simulation.

Given that k^{-2} scalings for $\overline{\rho}^z$ have been found in a number of other modelling studies (Hallberg and Gnanadesikan, 2006; Capet et al., 2008) we suggest that this property depends on certain details of the model configuration. Hallberg and Gnanadesikan (2006) (see their appendix A) employ a spectral nudging procedure to ensure that their surface buoyancy fluxes do not suppress eddy variability. The use of a 2° surface bulk forcing dataset (CORE2; Large and Yeager, 2009) in our uncoupled simulations may similarly act to suppress variability at finer scales, which could explain the k^{-4} spectral slope over this range of k. However, power spectra for coupled GC3 simulations (the 1950 piControl simulations of HighResMIP; Haarsma et al., 2016) with N96, N216 and N512 atmospheric resolution (approximately equivalent at mid-latitudes to grid spacings of 135km, 60km and 25km respectively) were also found to have a k^{-4} spectral slope. Furthermore, the horizontal buoyancy gradient in an ocean model is not only a function of forcing and resolution but also of numerical scheme and subgrid dissipation and viscosity. These differ between NEMO, which uses the FCT upwinding for diffusivity and the scalings of Willebrand et al. (2001), and the MESO simulations of Hallberg and Gnanadesikan (2006) used to study R in FK11, which use the Griffies and Hallberg (2000) closure. Such choices can have a strong impact on the energy of resolved mesoscale dynamics (Bachman et al., 2017a; Pearson et al., 2017); we note that the largest change in R occurs between the eddy-permitting 1/4° and non-eddying 1° simulations which represent mesoscale eddies in entirely different ways. Further work might attempt to identify the key processes affecting the scaling of $\overline{\rho}^z$ and seek to improve the specification of the S/L_f scaling.

The third aspect of the parameterisation explored here is the specification of the density difference, $\Delta\sigma_\theta$ with units kgm^{-3}, used to determine the depth scale, H, within (2). FK11 state that H should be consistent with the MLD where possible, the definition of which may vary considerably between models. The global mean impact of the SMLEP on the MLD is reduced by nearly a factor of

3 when H is defined to be close to the mixing layer depth ($\Delta\sigma_\theta = 0.01$) rather than the seasonal mixed layer ($\Delta\sigma_\theta = 0.03$). This sensitivity is observed in both the $1/4°$ GO5 and GC2 simulations, although the spatial characteristics of the impact differ significantly due to the response of coupled atmosphere and sea ice feedbacks. We suggest that the factor of 3 sensitivity is partly due to the SMLEP overturning only part of the mixed layer when using a $\Delta\sigma_\theta = 0.01$ criterion for H. This factor is reduced to slightly less than 2 when the MLD diagnostic, H_d, is defined using a $\Delta\sigma_\theta = 0.01\text{kgm}^{-3}$ criterion instead of the $\Delta\sigma_\theta = 0.03\text{kgm}^{-3}$ criterion used throughout the paper. The sensitivity of the SMLEP MLD impact to other aspects of its formulation is expected to be similarly dependent on the choice of criterion for H_d and may also differ in configurations with a coupled atmosphere model.

The impact of the SMLEP on MLD biases in the $1/4°$ uncoupled GO5 and coupled GC2 simulations is investigated using the approximate scheme with its standard parameter settings. The SMLEP systematically reduces the MLD in both GO5 and GC2 and tends to reduce the magnitude of deep MLD biases while increasing that of shallow MLD biases. Summer mixed layers are generally too shallow in both GO5 and GC2. Their global mean bias increases by around 2% of climatological values when the SMLEP is introduced, but their standard deviation is only slightly affected. Winter mixed layers are too deep in certain regions in both GO5 and GC2, but these deep MLD biases are larger in magnitude and more prevalent in GO5. This is reflected by the global mean bias, which reduces from 17% to 9% of climatological values in GO5 but changes from +3% to −4% in GC2 when the SMLEP is introduced. The global standard deviation of the winter MLD biases is reduced in both configurations, from 193m to 180m in GO5 and from 117m to 98m in GC2. The SMLEP has a much larger impact on the deep regional winter MLD biases; in the Labrador Sea the average bias is reduced from 185% to 157% in GO5 and from 62% to 15% in GC2. For the other regions studied, winter mixed layers are on average too deep in GO5 and too shallow in GC2, and the SMLEP respectively tends to decrease and increase the magnitude of these MLD biases.

The SMLEP was developed with the aim of representing one of the physical processes active in the restratification of deep winter mixed layers (Mahadevan et al., 2012; Swart et al., 2015; Thompson et al., 2016). This physical basis, the overall reduction of deep climatological MLD biases and the reduction in the global standard deviation of the biases demonstrate that the SMLEP is an important parameterisation to include in NEMO. FK11 note that while the reduction of MLD biases by the SMLEP is desirable, it is not a robust indicator of accurately parameterised MLE physics. The spatial and temporal variations in the impact of the SMLEP on MLD biases in NEMO, in the simulations of FK11 and in other ocean models (Weijer et al., 2012; Bentsen et al., 2013; Swapna et al., 2015) demonstrate that the representation of other physical processes also requires improvement. There remain significant differences in MLD biases between ocean models (Huang et al., 2014; Heuzé, 2017), due in part to uncertainty in the representation of vertical mixing processes (Li et al., 2019). As a result, FK11 suggest that other sub-grid scale parameterisations may need to be retuned to account for the SMLEP. We note that a particularly sensitive parameterisation of inertial wave breaking (Rodgers et al., 2014) was adjusted in the GO5 (Megann et al., 2014) and GO6 (Storkey et al., 2018) configurations to improve summer MLDs in the Southern Ocean. While outside the scope of this study, our results suggest that this parameterisation could be further adjusted to counteract the detrimental impact of the SMLEP on summer MLD biases, just as Langmuir mixing and SMLEP have been coordinated in other models to see improvements in summer MLD biases (Li and Fox-Kemper, 2017; Li et al., 2019).

The sensitivities discussed in this paper have important implications for the impact of the SMLEP on ocean model biases. The MLD, H, and mixed layer frontal width, L_f, have a clear physical definition, but their specification is poorly constrained and their calculation may vary significantly between models. We have used only a limited number of specifications for these parameters, in order to quantify the sensitivity of the SMLEP in a straightforward way. We therefore do not make any general recommendations regarding these parameters other than to reiterate those of FK11; that their specification be reported as part of the model configuration. A more complete investigation of the dependence of the SMLEP on H might consider other details of the MLD calculation, such as the choice of reference depth in the finite-difference criterion used here or the use of other criteria such as gradient, integral and regression methods (Thomson and Fine, 2003).

We have shown that the S/L_f scaling term does not necessarily ensure that the SMLEP is independent of horizontal resolution, but at present it is unclear whether it should be modified accordingly and how this could be done. As a first step towards this, we suggest that further work identify the key details of ocean model configurations affecting the strength of resolved mixed layer density fronts.

We have shown that ocean-sea ice feedbacks can result in large impacts on the MLD by the SMLEP. This is consistent with the results of FK11, who find that the SMLEP has a large impact on sea ice extent and thickness but do not examine how this in turn impacts the MLD. In particular, there is a strong interaction between the SMLEP and the Weddell Sea polynya that develops in GO5 (Megann et al., 2014; Heuzé et al., 2015). The Weddell Sea polynya is considered to be spurious and detrimental to several properties of the modelled ocean. Winter mixed layers in GO5 are too deep in this region and the SMLEP further increases the magnitude of this bias. Following FK11 we suggest that further work more closely examine the interaction between the SMLEP and sea ice, and how this in turn affects ocean properties other than the MLD.

Acknowledgements

Baylor Fox-Kemper was supported by NSF 1350795 and ONR N00014-17-1-2393. Daley Calvert and Mike Bell were supported by the Met Office Hadley Centre Climate Programme funded by BEIS and Defra. George Nurser was supported by the UK Natural Environment Research Council through the OSMOSIS project, NE/N018095/1. The authors are grateful to Stephen Belcher, John Siddorn and Adrian Hines for their guidance during the early stages of this work. We are grateful to two anonymous reviewers, whose constructive and insightful suggestions improved this manuscript.

Declaration of interest

None.

References

Adcroft, A., Campin, J.-M., 2004. Rescaled height coordinates for accurate representation of free-surface flows in ocean circulation models. Ocean Modelling 7 (3-4), 269–284.

Arakawa, A., Lamb, V. R., 1981. A potential enstrophy and energy conserving scheme for the shallow water equations. Monthly Weather Review 109 (1), 18–36.

Axell, L. B., 2002. Wind-driven internal waves and Langmuir circulations in a numerical ocean model of the southern Baltic Sea. Journal of Geophysical Research: Oceans 107 (C11).

Bachman, S. D., Fox-Kemper, B., Pearson, B., 2017a. A scale-aware subgrid model for quasi-geostrophic turbulence. Journal of Geophysical Research: Oceans 122 (2), 1529–1554.

Bachman, S. D., Fox-Kemper, B., Taylor, J. R., Thomas, L., 2017b. Parameterization of frontal symmetric instabilities. I: Theory for resolved fronts. Ocean Modelling 109, 72–95.

Barnier, B., Madec, G., Penduff, T., Molines, J.-M., Treguier, A.-M., Le Sommer, J., Beckmann, A., Biastoch, A., Böning, C., Dengg, J., Derval, C., Durand, E., Gulev, S., Remy, E., Talandier, C., Theetten, S., Maltrud, M., McClean, J., De Cuevas, B., 2006. Impact of partial steps and momentum advection schemes in a global ocean circulation model at eddy-permitting resolution. Ocean Dynamics 56 (5), 543–567.

Bentsen, M., Bethke, I., Debernard, J. B., Iversen, T., Kirkevåg, A., Seland, Ø., Drange, H., Roelandt, C., Seierstad, I. A., Hoose, C., Kristjánsson, J. E., 2013. The Norwegian Earth System Model, NorESM1-M – Part 1: Description and basic evaluation of the physical climate. Geosci. Model Dev 6 (3), 687–720.

Boccaletti, G., Ferrari, R., Fox-Kemper, B., 2007. Mixed layer instabilities and restratification. J. Phys. Oceanogr. 37, 2228–2250.

Bodner, A. S., Fox-Kemper, B., Van Roekel, L. P., McWilliams, J. C., Sullivan, P. P., 2020. A perturbation approach to understanding the effects of turbulence on frontogenesis. Journal of Fluid Mechanics 883, A25.

Callies, J., Ferrari, R., 2018. Note on the rate of restratification in the baroclinic spindown of fronts. Journal of Physical Oceanography 48 (7), 1543–1553.

Capet, X., McWilliams, J. C., Molemaker, M. J., Shchepetkin, A., 2008. Mesoscale to submesoscale transition in the California Current System. Part I: Flow structure, eddy flux, and observational tests. Journal of physical oceanography 38 (1), 29–43.

Cavalieri, D., Parkinson, C., 2008. Antarctic sea ice variability and trends, 1979–2006. Journal of Geophysical Research: Oceans 113 (C7).

Craig, P. D., Banner, M. L., 1994. Modeling wave-enhanced turbulence in the ocean surface layer. Journal of Physical Oceanography 24 (12), 2546–2559.

Craik, A., Leibovich, S., 1976. A rational model for Langmuir circulation. J. Fluid Mech. 73, 401–426.

Danabasoglu, G., Yeager, S. G., Kwon, Y.-O., Tribbia, J. J., Phillips, A. S., Hurrell, J. W., 2012. Variability of the Atlantic meridional overturning circulation in CCSM4. Journal of Climate 25 (15), 5153–5172.

de Boyer Montégut, C., Madec, G., Fischer, A. S., Lazar, A., Iudicone, D., 2004. Mixed layer depth over the global ocean: An examination of profile data and a profile-based climatology. Journal of Geophysical Research: Oceans (1978–2012) 109 (C12).

Farneti, R., Downes, S. M., Griffies, S. M., Marsland, S. J., Behrens, E., Bentsen, M., Bi, D., Biastoch, A., Böning, C., Bozec, A., Canuto, V. M., Chassignet, E., Danabasoglu, G., Danilov, S., Diansky, N., Drange, H., Fogli, P. G., Gusev, A., Hallberg, R. W., Howard, A., Ilicak, M., Jung, T., Kelley, M., Large, W. G., Leboissetier, A., Long, M., Lu, J., Masina, S., Mishra, A., Navarra, A., Nurser, A. G., Patara, L., Samuels, B. L., Sidorenko, D., Tsujino, H., Uotila, P., Wang, Q., Yeager, S. G., 2015. An assessment of Antarctic Circumpolar Current and Southern Ocean meridional overturning circulation during 1958–2007 in a suite of interannual CORE-II simulations. Ocean Modelling 93, 84–120.

Fox-Kemper, B., Danabasoglu, G., Ferrari, R., Griffies, S., Hallberg, R., Holland, M., Maltrud, M., Peacock, S., Samuels, B., 2011. Parameterization of mixed layer eddies. III: Implementation and impact in global ocean climate simulations. Ocean Modelling 39 (1), 61–78.

Fox-Kemper, B., Danabasoglu, G., Ferrari, R., Hallberg, R., 2008a. Parameterizing submesoscale physics in global climate models. CLIVAR Exchanges 13 (1), 3–5.

Fox-Kemper, B., Ferrari, R., Hallberg, R., 2008b. Parameterization of mixed layer eddies. Part II: Prognosis and impact. Journal of Physical Oceanography 38 (6), 1166–1179.

Gaspar, P., Grégoris, Y., Lefevre, J.-M., 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. Journal of Geophysical Research: Oceans 95 (C9), 16179–16193.

Gent, P. R., McWilliams, J. C., 1990. Isopycnal mixing in ocean circulation models. Journal of Physical Oceanography 20 (1), 150–155.

Good, S. A., Martin, M. J., Rayner, N. A., 2013. EN4: Quality controlled ocean temperature and salinity profiles and monthly objective analyses with uncertainty estimates. Journal of Geophysical Research: Oceans 118 (12), 6704–6716.

Griffies, S., Hallberg, R., 2000. Biharmonic friction with a Smagorinsky-like viscosity for use in large-scale eddy-permitting ocean models. Monthly Weather Review 128 (8), 2935–2946.

Griffiths, S. D., 2008. The limiting form of inertial instability in geophysical flows. J Fluid Mech 605, 115–143.

Haarsma, R. J., Roberts, M. J., Vidale, P. L., Senior, C. A., Bellucci, A., Bao, Q., Chang, P., Corti, S., Fučkar, N. S., Guemas, V., von Hardenberg, J., Hazeleger, W., Kodama, C., Koenigk, T., Leung, L. R., Lu, J., Luo, J.-J., Mao, J., Mizielinski, M. S., Mizuta, R., Nobre, P., Satoh, M., Scoccimarro, E., Semmler, T., Small, J., von Storch, J.-S., 2016. High Resolution Model Intercomparison Project (HighResMIP v1.0) for CMIP6. Geoscientific Model Development 9 (11), 4185–4208.
URL https://www.geosci-model-dev.net/9/4185/2016/

Haine, T. W. N., Marshall, J., 1998. Gravitational, symmetric, and baroclinic instability of the ocean mixed layer. J. Phys. Oceanogr. 28, 634–658.

Hallberg, R., Gnanadesikan, A., 2006. The role of eddies in determining the structure and response of the wind-driven Southern Hemisphere overturning: Results from the Modeling Eddies in the Southern Ocean (MESO) project. Journal of Physical Oceanography 36 (12), 2232–2252.

Held, I. M., Larichev, V. D., 1996. A scaling theory for horizontally homogeneous, baroclinically unstable flow on a beta plane. Journal of the atmospheric sciences 53 (7), 946–952.

Heuzé, C., 2017. North Atlantic deep water formation and AMOC in CMIP5 models. Ocean Science 13 (4), 609.

Heuzé, C., Ridley, J., Calvert, D., Stevens, D., Heywood, K., 2015. Increasing vertical mixing to reduce Southern Ocean deep convection in NEMO. Geoscientific Model Development Discussions 8 (3), 2949–2972.

Hollingsworth, A., Kållberg, P., Renner, V., Burridge, D., 1983. An internal symmetric computational instability. Quarterly Journal of the Royal Meteorological Society 109 (460), 417–428.

Hosegood, P., Gregg, M. C., Alford, M. H., 2006. Sub-mesoscale lateral density structure in the oceanic surface mixed layer. Geophysical research letters 33 (22).

Huang, C. J., Qiao, F., Dai, D., 2014. Evaluating CMIP5 simulations of mixed layer depth during summer. Journal of Geophysical Research: Oceans 119 (4), 2568–2582.

Hunke, E. C., Lipscomb, W. H., 2010. The Los Alamos sea ice model, documentation and software user's manual, version 4.1. Los Alamos National Laboratory Tech. Rep. LA-CC-06-012.

Ingleby, B., Huddleston, M., 2007. Quality control of ocean temperature and salinity profiles- Historical and real-time data. Journal of Marine Systems 65 (1), 158–175.

Jochum, M., Briegleb, G., Danabasoglu, G., Large, W. G., Norton, N. J., Jayne, S. R., Alford, M. H., Bryan, F. O., 2013. The impact of oceanic near-inertial waves on climate. J. Climate 26 (9), 2833–2844.

Johnson, L., Lee, C. M., D'Asaro, E. A., 2016. Global estimates of lateral springtime restratification. Journal of Physical Oceanography 46 (5), 1555–1573.

Koch-Larrouy, A., Madec, G., Blanke, B., Molcard, R., 2008. Water mass transformation along the Indonesian throughflow in an OGCM. Ocean Dynamics 58 (3-4), 289–309.

Large, W., McWilliams, J., Doney, S., 1994. Oceanic vertical mixing: A review and a model with nonlocal boundary layer parameterization. Rev. Geophys. 32, 363 – 403.

Large, W., Yeager, S., 2009. The global climatology of an interannually varying air–sea flux data set. Climate dynamics 33 (2-3), 341–364.

Large, W. G., Crawford, G. B., 1995. Observations and simulations of upper-ocean response to wind events during the Ocean Storms Experiment. J. Phys. Oceanogr. 25, 2831 – 2852.

Levitus, S., Boyer, T., Conkright, M., O'Brien, T., Antonov, J., Stephens, C., Stathoplos, L., Johnson, D., Gelfeld, R., 1998. NOAA Atlas NESDIS 18 and World Ocean Database 1998, vol. 1. U.S. Gov. Printing Office, Washington DC.

Li, Q., Fox-Kemper, B., 2017. Assessing the effects of Langmuir turbulence on the entrainment buoyancy flux in the ocean surface boundary layer. J. Phys. Oceanogr. 27, 2863–2886.

Li, Q., Reichl, B. G., Fox-Kemper, B., Adcroft, A. J., Belcher, S., Danabasoglu, G., Grant, A., Griffies, S. M., Hallberg, R. W., Hara, T., Harcourt, R., Kukulka, T., Large, W. G., McWilliams, J. C., Pearson, B., Sullivan, P., Roekel, L. V., Wang, P., Zheng, Z., 2019. Comparing ocean boundary vertical mixing schemes including Langmuir turbulence. Journal of Advances in Modeling Earth Systems (JAMES) 11 (11), 3545–3592.

Madec, G., Bourdallé-Badie, R., Bouttier, P.-A., Bricaud, C., Bruciaferri, D., Calvert, D., Chanut, J., Clementi, E., Coward, A., Delrosso, D., Ethé, C., Flavoni, S., Graham, T., Harle, J., Iovino, D., Lea, D., Lévy, C., Lovato, T., Martin, N., Masson, S., Mocavero, S., Paul, J., Rousset, C., Storkey, D.,

Storto, A., Vancoppenolle, M., Oct. 2017. NEMO ocean engine. Revision 8625 from SVN repository.
URL https://doi.org/10.5281/zenodo.1472492

Madec, G., Imbard, M., 1996. A global ocean mesh to overcome the North Pole singularity. Climate Dynamics 12 (6), 381–388.

Mahadevan, A., D'Asaro, E., Lee, C., Perry, M. J., 2012. Eddy-driven stratification initiates North Atlantic spring phytoplankton blooms. Science 337 (6090), 54–58.

Marshall, J., Schott, F., 1999. Open-ocean convection: Observations, theory and models. Rev. Geophys. 37, 1, 1–64.

McWilliams, J. C., Sullivan, P. P., Moeng, C.-H., 1997. Langmuir turbulence in the ocean. J. Fluid Mech. 334, 1–30.

Megann, A., Storkey, D., Aksenov, Y., Alderson, S., Calvert, D., Graham, T., Hyder, P., Siddorn, J., Sinha, B., 2014. GO5.0: The joint NERC-Met Office NEMO global ocean model for use in coupled and forced applications. Geoscientific Model Development 7, 1069–1092.
URL https://doi.org/10.5194/gmd-7-1069-2014

Merryfield, W. J., Holloway, G., Gargett, A. E., 1999. A global ocean model with double-diffusive mixing. Journal of Physical Oceanography 29 (6), 1124–1142.

Parkinson, C. L., Cavalieri, D. J., 2008. Arctic sea ice variability and trends, 1979–2006. Journal of Geophysical Research: Oceans 113 (C7).

Pearson, B., Fox-Kemper, B., Bachman, S., Bryan, F., 2017. Evaluation of scale-aware subgrid mesoscale eddy models in a global eddy-rich model. Ocean Modelling 115, 42–58.

Rae, J., Hewitt, H., Keen, A., Ridley, J., West, A., Harris, C., Hunke, E., Walters, D., 2015. Development of the Global Sea Ice 6.0 CICE configuration for the Met Office Global Coupled model. Geoscientific Model Development 8 (7), 2221–2230.

Ridley, J. K., Blockley, E. W., Keen, A. B., Rae, J. G., West, A. E., Schroeder, D., 2018. The sea ice model component of HadGEM3-GC3.1. Geoscientific Model Development 11 (2), 713–723.

Rodgers, K. B., Aumont, O., Fletcher, S. M., Plancherel, Y., Bopp, L., Montégut, C. D. B., Iudicone, D., Keeling, R., Madec, G., Wanninkhof, R., 2014. Strong sensitivity of Southern Ocean carbon uptake and nutrient cycling to wind stirring. Biogeosciences 11 (15), 4077–4098.

Sakai, S., 1989. Rossby-Kelvin instability: A new type of ageostrophic instability caused by a resonance between Rossby waves and gravity waves. J. Fluid Mech. 202, 149–176.

Simmons, H. L., Jayne, S. R., Laurent, L. C. S., Weaver, A. J., 2004. Tidally driven mixing in a numerical model of the ocean general circulation. Ocean Modelling 6 (3-4), 245–263.

Steele, M., Morley, R., Ermold, W., 2001. PHC: A global ocean hydrography with a high-quality Arctic Ocean. Journal of Climate 14 (9), 2079–2087.

Stone, P. H., 1970. On non-geostrophic baroclinic instability: Part II. J. Atmos. Sci. 27, 721–726.

Storkey, D., Blaker, A. T., Mathiot, P., Megann, A., Aksenov, Y., Blockley, E. W., Calvert, D., Graham, T., Hewitt, H. T., Hyder, P., Kuhlbrodt, T., Rae, J. G. L., Sinha, B., 2018. UK Global Ocean GO6 and GO7: A traceable hierarchy of model resolutions. Geoscientific Model Development Discussions 2018, 1–43.
URL `https://www.geosci-model-dev-discuss.net/gmd-2017-263/`

Sullivan, P. P., McWilliams, J. C., 2018. Frontogenesis and frontal arrest of a dense filament in the oceanic surface boundary layer. Journal of Fluid Mechanics 837, 341–380.

Swapna, P., Roxy, M., Aparna, K., Kulkarni, K., Prajeesh, A., Ashok, K., Krishnan, R., Moorthi, S., Kumar, A., Goswami, B., 2015. The IITM Earth system model: Transformation of a seasonal prediction model to a long-term climate model. Bulletin of the American Meteorological Society 96 (8), 1351–1367.

Swart, S., Thomalla, S., Monteiro, P., 2015. The seasonal cycle of mixed layer dynamics and phytoplankton biomass in the Sub-Antarctic Zone: A high-resolution glider experiment. Journal of Marine Systems 147, 103–115. URL `http://www.sciencedirect.com/science/article/pii/S0924796314001535`

Thomas, L. N., 2005. Destruction of potential vorticity by winds. J. Phys. Oceanogr. 35, 2457–2466.

Thomas, L. N., Taylor, J. R., Ferrari, R., Joyce, T. M., 2013. Symmetric instability in the Gulf Stream. Deep Sea Res. II 91, 96–110.

Thompson, A. F., Lazar, A., Buckingham, C., Naveira Garabato, A. C., Damerell, G. M., Heywood, K. J., 2016. Open-ocean submesoscale motions: A full seasonal cycle of mixed layer instabilities from gliders. Journal of Physical Oceanography 46 (4), 1285–1307.
URL `https://doi.org/10.1175/JPO-D-15-0170.1`

Thomson, R. E., Fine, I. V., 2003. Estimating mixed layer depth from oceanic profile data. Journal of Atmospheric and Oceanic Technology 20 (2), 319–329.

Turner, J., Kraus, E., 1967. A one-dimensional model of the seasonal thermocline I. A laboratory experiment and its interpretation. Tellus 19 (1), 88–97.

Weijer, W., Sloyan, B. M., Maltrud, M. E., Jeffery, N., Hecht, M. W., Hartin, C. A., van Sebille, E., Wainer, I., Landrum, L., 2012. The Southern Ocean and its climate in CCSM4. Journal of climate 25 (8), 2652–2675.

Willebrand, J., Barnier, B., Boning, C., Dieterich, C., Killworth, P., Le Provost, C., Jia, Y., Molines, J., New, A., 2001. Circulation characteristics in three eddy-permitting models of the North Atlantic. Progress In Oceanography 48, 123–161.

Williams, K. D., Harris, C. M., Bodas-Salcedo, A., Camp, J., Comer, R. E., Copsey, D., Fereday, D., Graham, T., Hill, R., Hinton, T., Hyder, P., Ineson, S., Masato, G., Milton, S. F., Roberts, M. J., Rowell, D. P., Sanchez, C., Shelly, A., Sinha, B., Walters, D. N., West, A., Woollings, T., Xavier, P. K., 2015. The Met Office Global Coupled model 2.0 (GC2) configuration. Geoscientific Model Development 8 (5), 1509–1524.

Zalesak, S. T., 1979. Fully multidimensional flux-corrected transport algorithms for fluids. Journal of computational physics 31 (3), 335–362.

	1/12°	1/4°	1°	2°
Time step (s)	360	1350	2700	5760
Eddy-induced transport coefficient (m^2s^{-1})	-	-	1000	2000
Lateral diffusion operator	Laplacian	Laplacian	Laplacian	Laplacian
Lateral diffusion coefficient (m^2s^{-1})	125	300 (150)	1000	2000
Lateral viscosity operator	Biharmonic	Biharmonic	Laplacian	Laplacian
Lateral viscosity coefficient (m^2s^{-1}, m^4s^{-1})	-1.25x10^{10}	-1.5x10^{11}	1x10^4 (2x10^4)	4x10^4

Table 1: Resolution-dependent parameters in the GO5 and GO6 simulations, with GO6 values in brackets where they differ from GO5. Coefficients are given as their maximum values; the lateral diffusion and viscosity coefficients decrease with the grid spacing (Laplacian operator) and with the cube of the grid spacing (biharmonic operator), and the eddy-induced transport coefficient varies spatially following Held and Larichev (1996).

Δs	$\Delta(\overline{H_d})$	$\Delta(\overline{H_d}^{\max})$	$\Delta(\overline{H_d}^{\min})$
1/4°	-5.4	-6.6	-1.5
1°	-8.8	-11.3	-2.3
2°	-10.1	-13.0	-2.3

Table 2: Global mean impact (%) of the SMLEP on mixed layer depths in the 1/4°, 1° and 2° GO5 simulations.

Configuration	$\Delta\sigma_\theta$	$\Delta(\overline{H_d})$	$\Delta(\overline{H_d}^{\max})$	$\Delta(\overline{H_d}^{\min})$
GO5	0.01	-1.9	-2.2	-0.4
GO5	0.03	-5.4	-6.6	-1.5
GC2	0.01	-1.6	-1.7	-1.2
GC2	0.03	-4.7	-5.4	-2.2

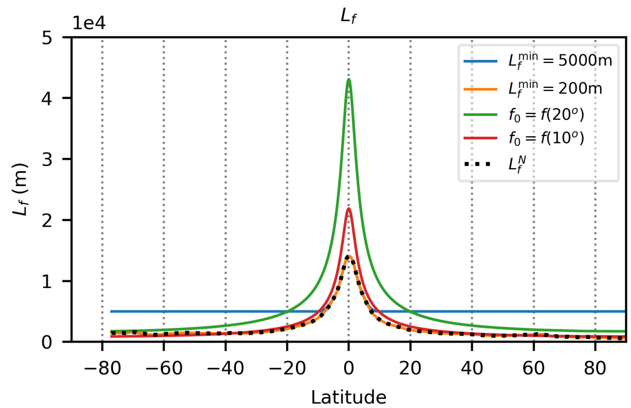
Table 3: Global mean impact (%) of the SMLEP on mixed layer depths in the 1/4° GO5 and GC2 simulations, as a function of the $\Delta\sigma_\theta$ finite-difference criterion used for H.

Configuration	$\overline{H_d}$		$\overline{H_d}^{\max}$		$\overline{H_d}^{\min}$	
	Mean (%)	Std. dev.	Mean (%)	Std. dev.	Mean (%)	Std. dev.
GO5	6 (9.8)	66	22 (16.9)	193	-2 (-2.8)	8
GO5 + SMLEP	2 (3.4)	62	10 (8.5)	180	-2 (-4.4)	8
GC2	-5 (-1.7)	39	-6 (3.1)	117	-4 (-8.9)	9
GC2 + SMLEP	-8 (-6.8)	33	-14 (-3.5)	98	-4 (-11.0)	10

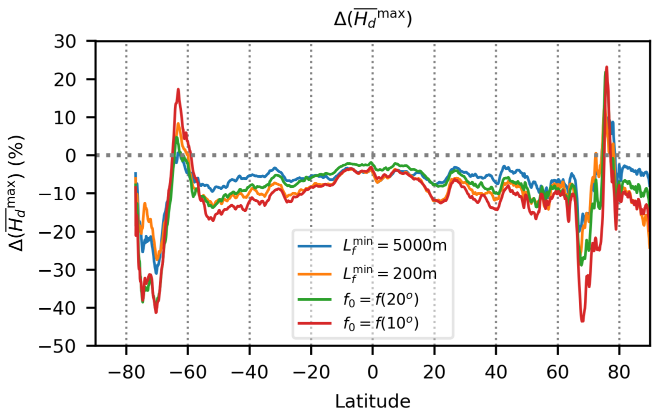
Table 4: Global mean mixed layer depth error (m), relative error (%) and error standard deviation (m) calculated using the climatology of de Boyer Montégut et al. (2004) in the 1/4° GO5 and GC2 simulations with and without the SMLEP. Figures in metres have been rounded to the nearest integer.

Configuration	Labrador	Greenland	Ross	Weddell
GO5	518 (185.2)	162 (53.7)	234 (161.3)	444 (362.0)
GO5 + SMLEP	446 (156.6)	110 (29.7)	71 (50.4)	512 (421.2)
GC2	255 (62.4)	-29 (1.7)	-69 (-41.8)	-38 (-30.5)
GC2 + SMLEP	118 (15.3)	-133 (-24.9)	-76 (-45.5)	-40 (-32.6)

Table 5: Winter (seasonal maximum) mean mixed layer depth error (m) and relative error (in brackets, %) calculated using the climatology of de Boyer Montégut et al. (2004) in the 1/4° GO5 and GC2 simulations with and without the SMLEP. The regions are defined using the NSIDC Arctic regional masks described by Cavalieri and Parkinson (2008) and Parkinson and Cavalieri (2008). Figures in metres have been rounded to the nearest integer.



(a)



(b)

Figure 1: Zonal averages of (a) the mixed layer frontal width L_f, calculated for the FK11 scheme using the year 6-30 average of L_f^m (3) in the 1/4° GO5 simulation without the SMLEP and for the approximate scheme using L_f^a (12), and (b) $\Delta(\overline{H_d}^{\max})$ (15), calculated as the difference between zonal averages of $\overline{H_d}^{\max}$ in the 1/4° GO5 simulations. The zonal average is taken along grid lines and will differ from an average along lines of latitude where the grid transitions from an isotropic Mercator grid to a bipolar grid north of 30°N.

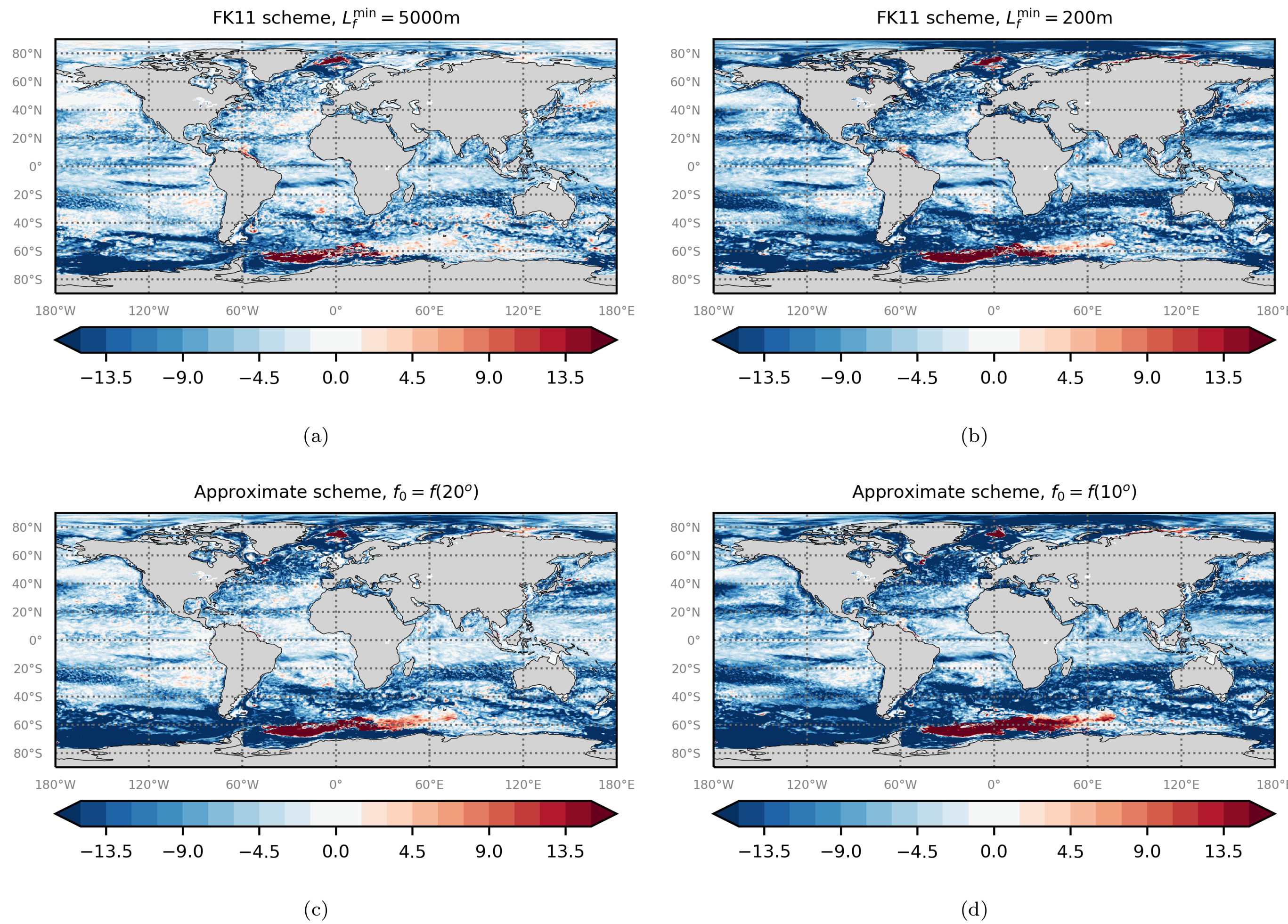


Figure 2: $\Delta(\overline{H_d}^{\max})$ in the 1/4° GO5 simulations using: the FK11 scheme with (a) $L_f^{\min} = 5000$m and (b) $L_f^{\min} = 200$m; and the approximate scheme with $L_0 = 5000$m and (c) $f_0 = f(20°)$, (d) $f_0 = f(10°)$.

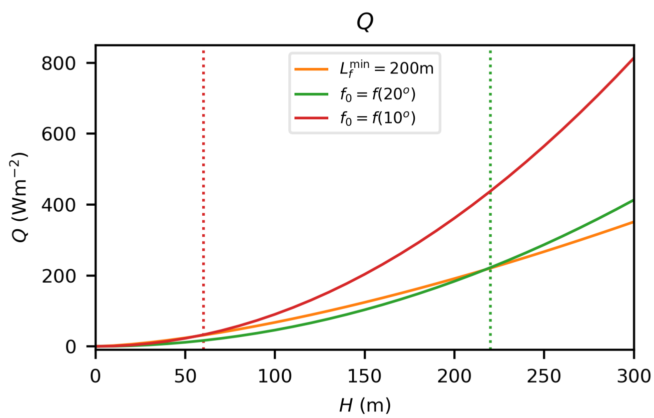
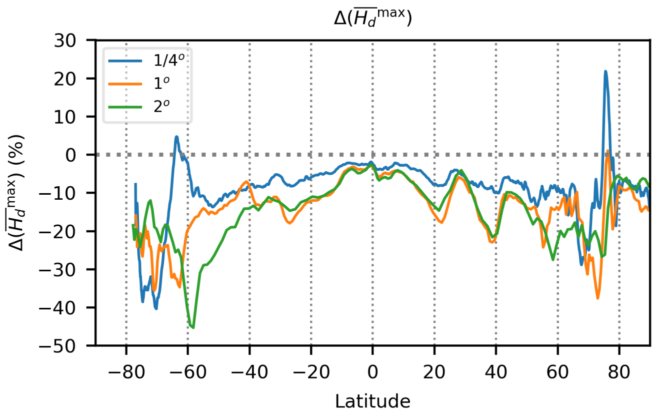
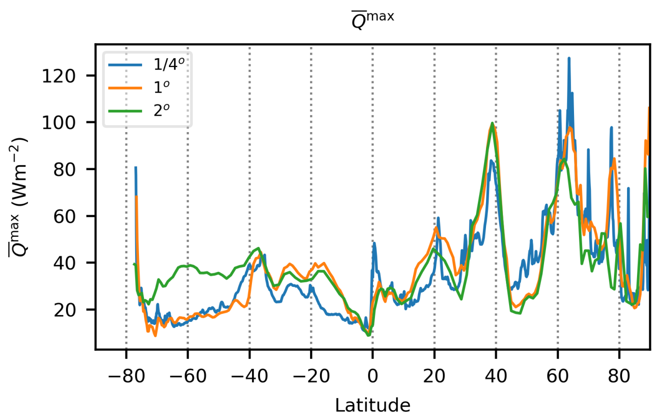


Figure 3: Scaling of the equivalent heat flux across the mixed layer Q, as defined by (14), using (11) for the FK11 scheme, (13) for the approximate scheme and assuming $R = 10^{-7}\text{kg}^2\text{m}^{-7}$. The dotted lines indicate where the $f_0 = f(20^o)$ (green line) and $f_0 = f(10^o)$ (red line) calculations of Q become larger than that of $L_f^{\min} = 200$m.



(a)



(b)

Figure 4: Zonal averages of (a) $\Delta(\overline{H_d}^{\max})$, calculated as the difference between zonal averages of $\overline{H_d}^{\max}$, and (b) the seasonal maximum of the equivalent heat flux across the mixed layer $\overline{Q}^{\max}$, in the 1/4°, 1° and 2° GO5 simulations. The zonal average is taken along grid lines and will differ from an average along lines of latitude where the grid transitions from an isotropic Mercator grid to a bipolar grid north of 30°N.

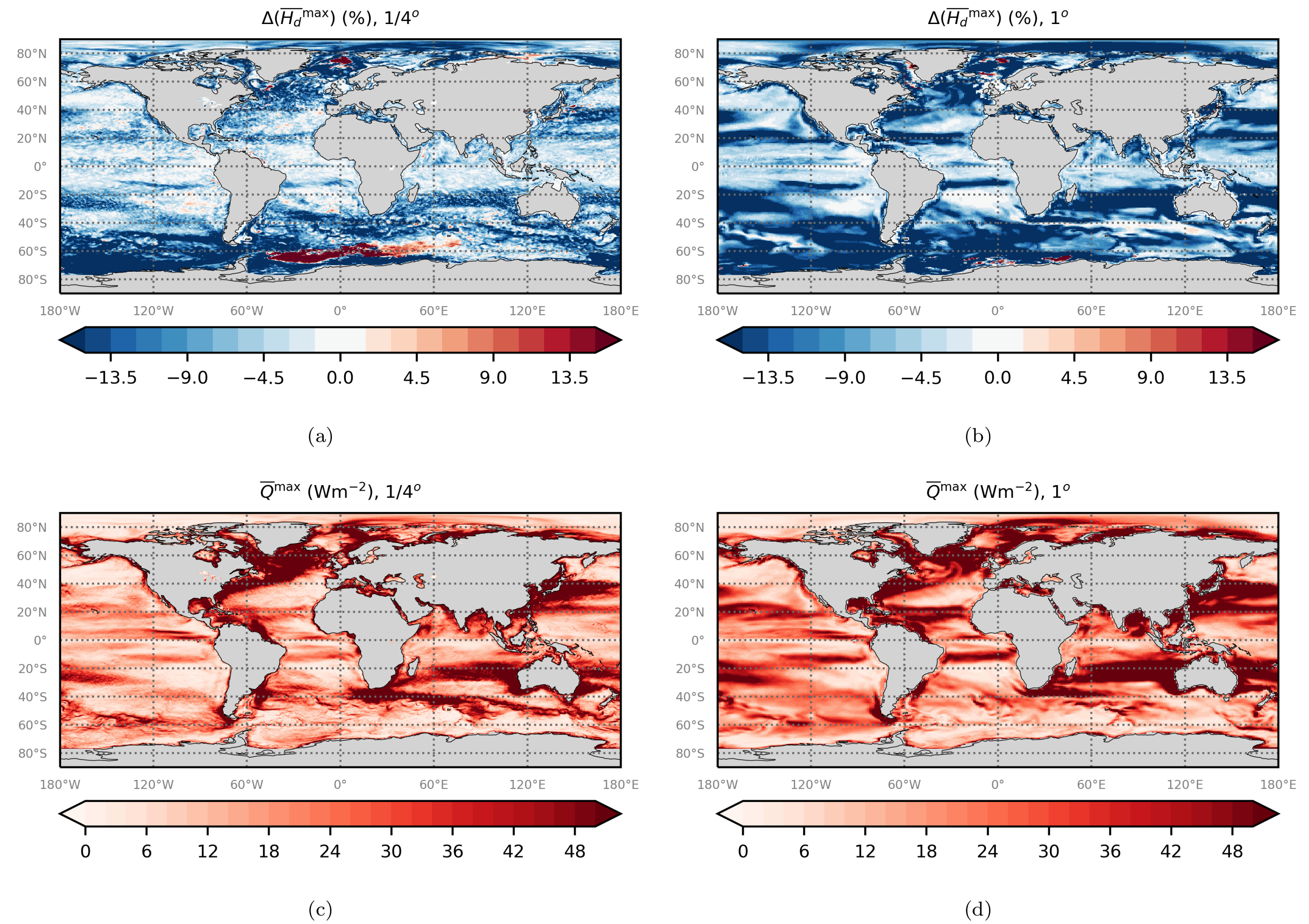


Figure 5: $\Delta(\overline{H_d}^{\max})$ (top) and $\overline{Q}^{\max}$ (bottom) in the (left) 1/4° and (right) 1° GO5 simulations. Panel (a) is identical to panel (c) of figure 2.

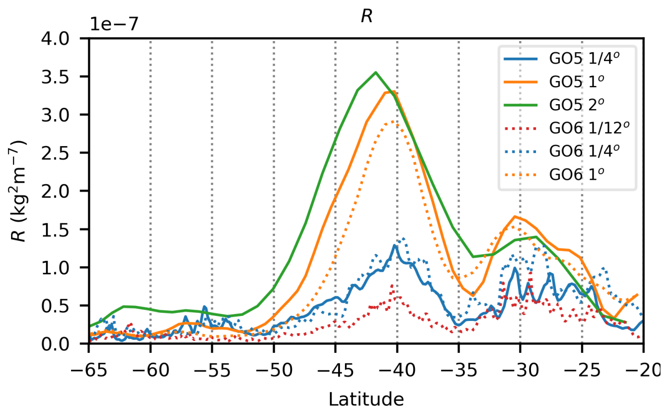
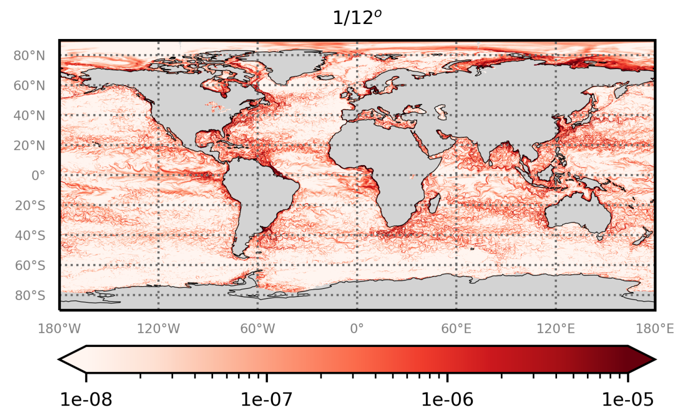
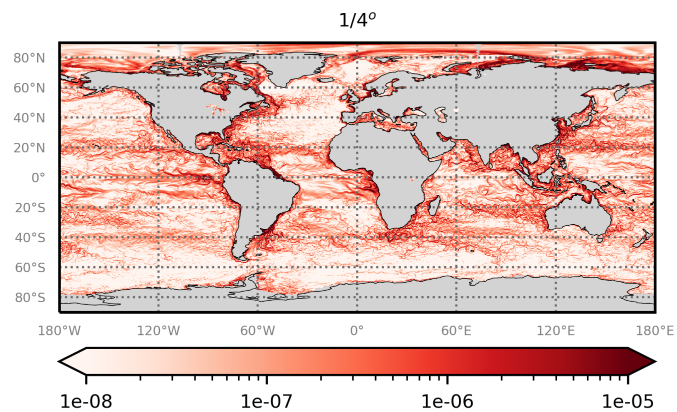


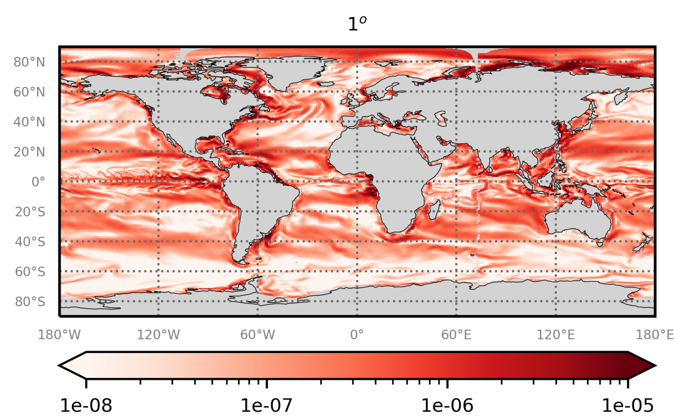
Figure 6: Zonal averages of resolution-scaled squared horizontal buoyancy gradient R, as defined by (8), over 20 - 65°S, 100 - 160°W calculated using instantaneous data for March, year 6 of the GO5 simulations without the SMLEP and the GO6 simulations.



(a)



(b)



(c)

Figure 7: Resolution-scaled squared horizontal buoyancy gradient R calculated using instantaneous data for March, year 6 of the (a) 1/12°, (b) 1/4° and (c) 1° GO6 simulations.

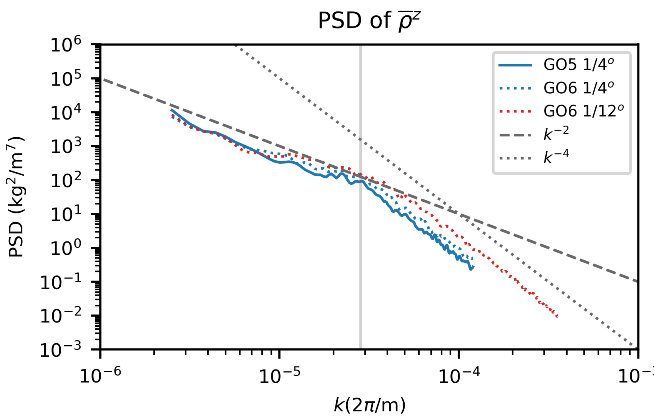


Figure 8: Average power spectral density of $\overline{\rho}^z$ over 20 - 65°S, 100 - 160°W calculated using instantaneous data for March, year 6 of the 1/4° GO5 simulation without the SMLEP and the 1/12°, 1/4° GO6 simulations. The grey lines are slopes of k^{-2} (dashed line), k^{-4} (dotted line) and $k \equiv 2°$ (solid line). To account for varying grid resolution along the meridional axis, zonal rows are treated as separate data segments from which the linear trend is removed, a Hanning window applied and the 1D power spectrum calculated. Cubic splines are then fitted to the individual spectra and meridionally averaged.

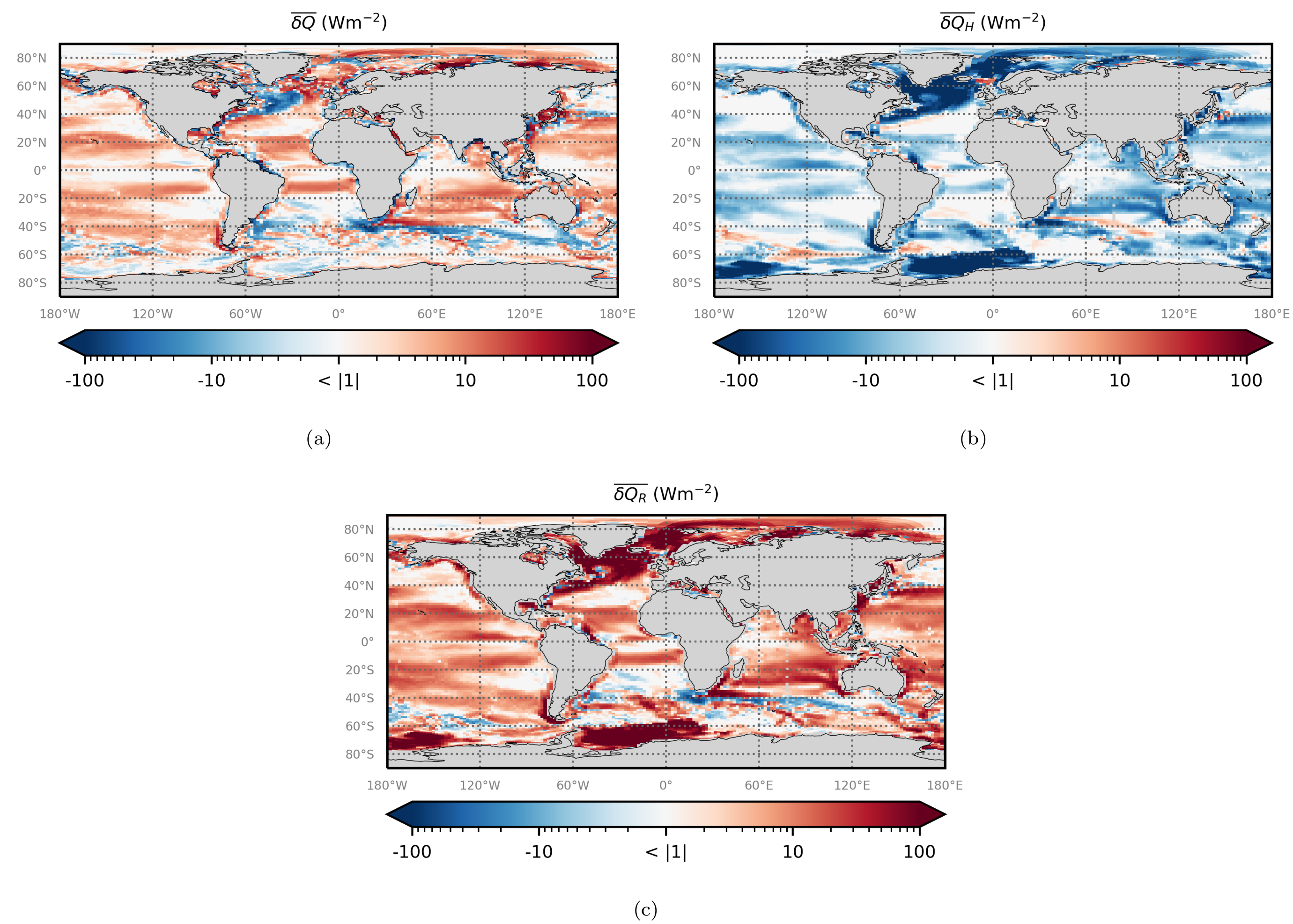
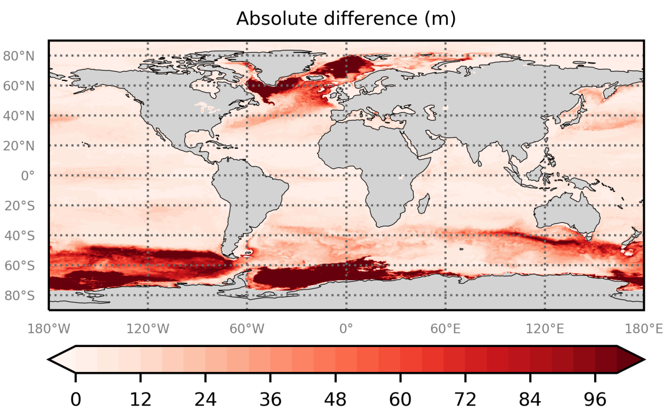
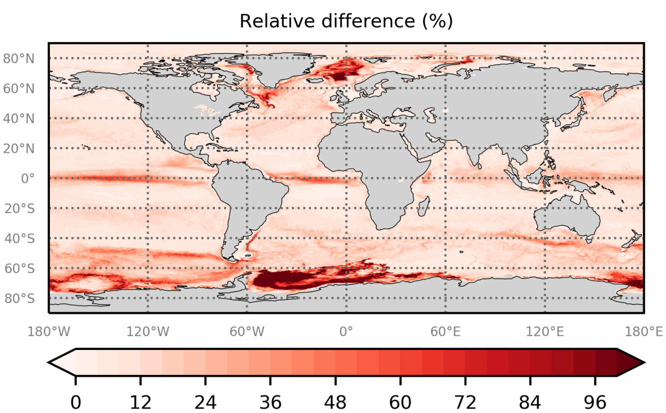


Figure 9: Terms in (17) calculated for the 1° GO5 simulation with respect to the 1/4° simulation and averaged over years 6-30: (a) $\overline{\delta Q}$, (b) $\overline{\delta Q_H}$ and (c) $\overline{\delta Q_R}$.



(a)



(b)

Figure 10: Absolute (a) and relative (b) difference between $\overline{H_d}^{\max}$ calculated using a $\Delta\sigma_\theta = 0.03\text{kgm}^{-3}$ criterion for H_d with respect to that calculated using a $\Delta\sigma_\theta = 0.01\text{kgm}^{-3}$ criterion, for the 1/4° GO5 simulation without the SMLEP.

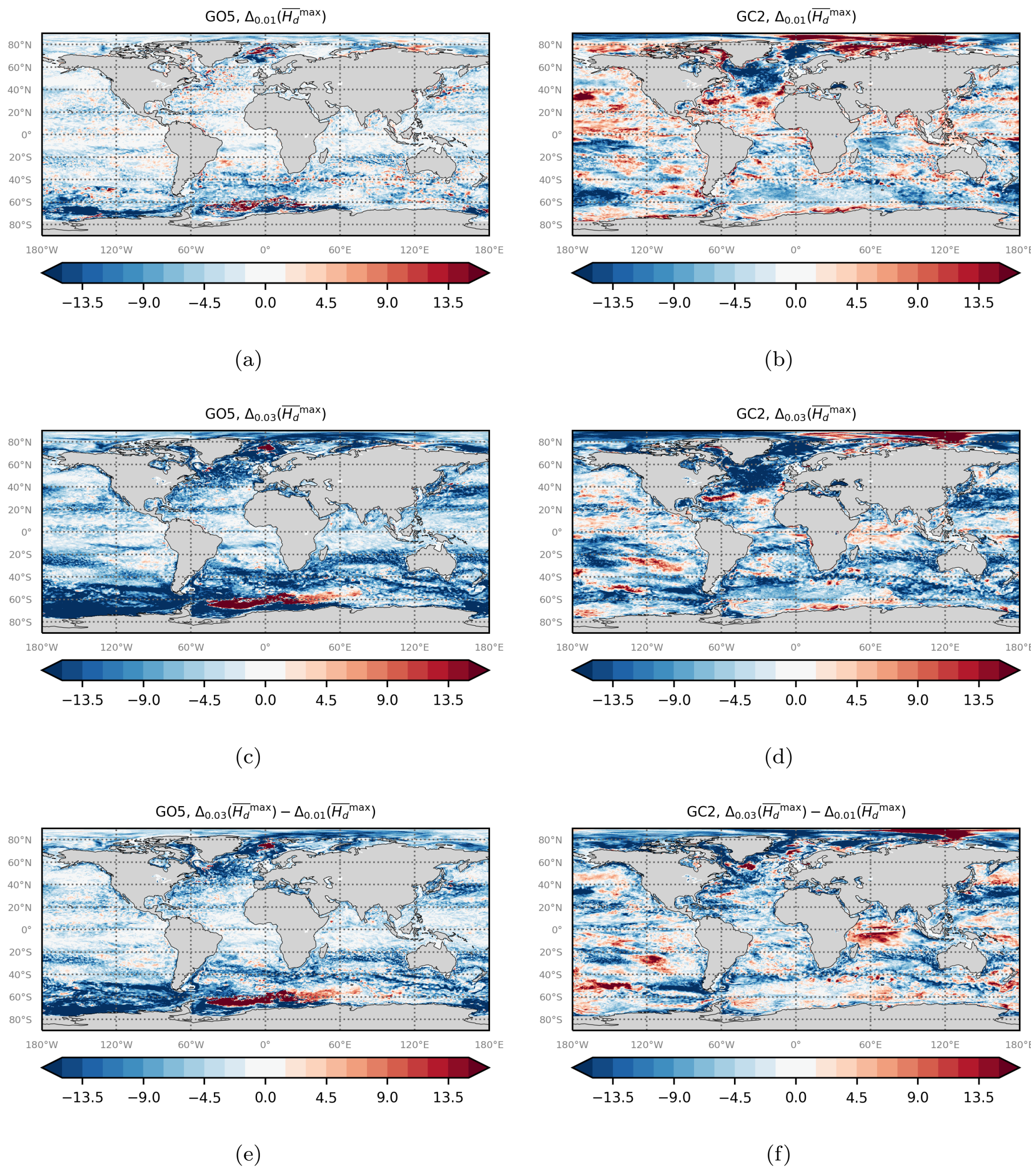


Figure 11: $\Delta_{0.01}(\overline{H_d}^{\max})$ (top), $\Delta_{0.03}(\overline{H_d}^{\max})$ (center) and $\Delta_{0.03}(\overline{H_d}^{\max}) - \Delta_{0.01}(\overline{H_d}^{\max})$ (bottom) for the 1/4° GO5 (left) and GC2 (right) simulations. Panel (c) is identical to panel (a) of figure 5 and panel (c) of figure 2.

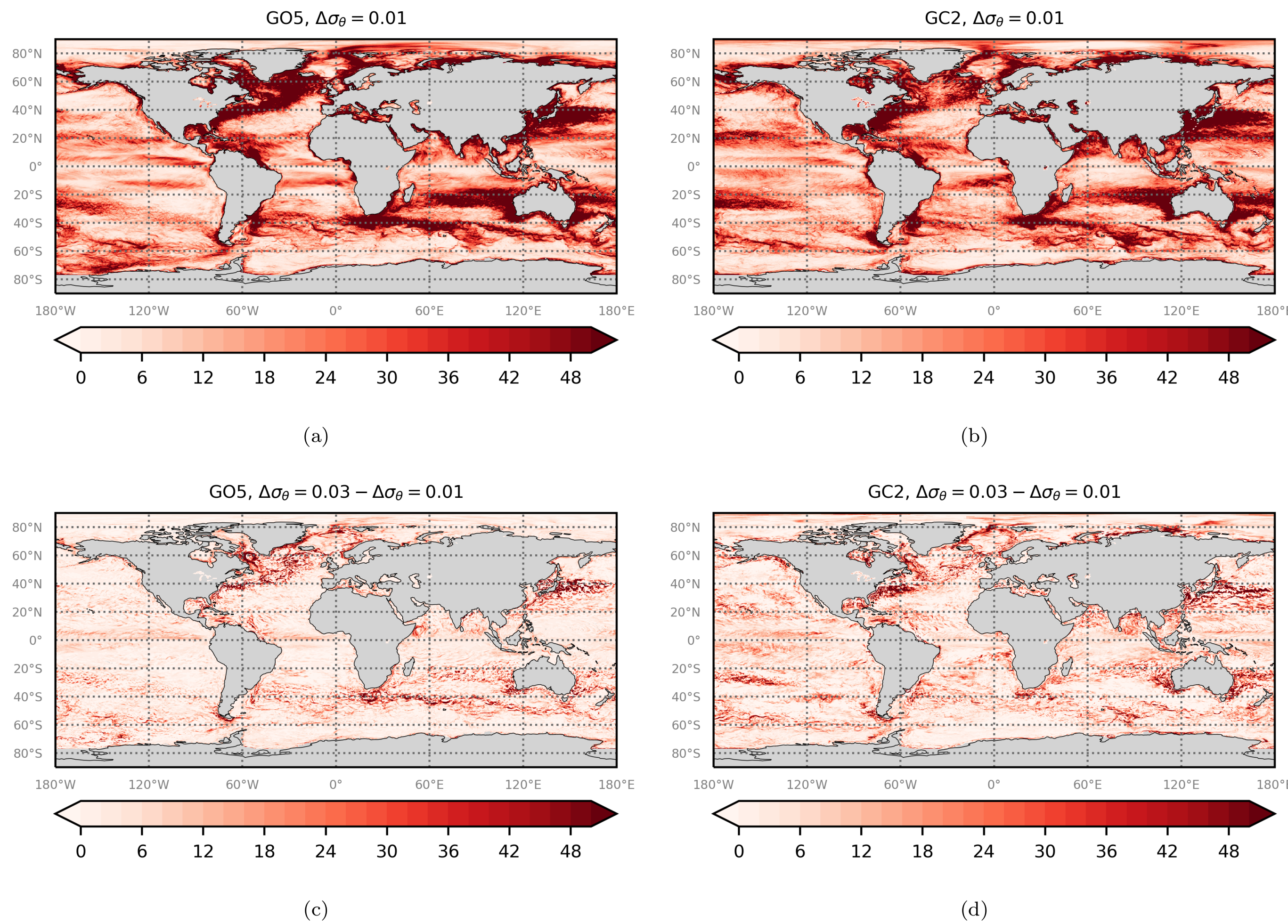


Figure 12: As for panels (a), (b), (e) and (f) of figure 11 but for $\overline{Q}^{\max}$ averaged over years 26-30.

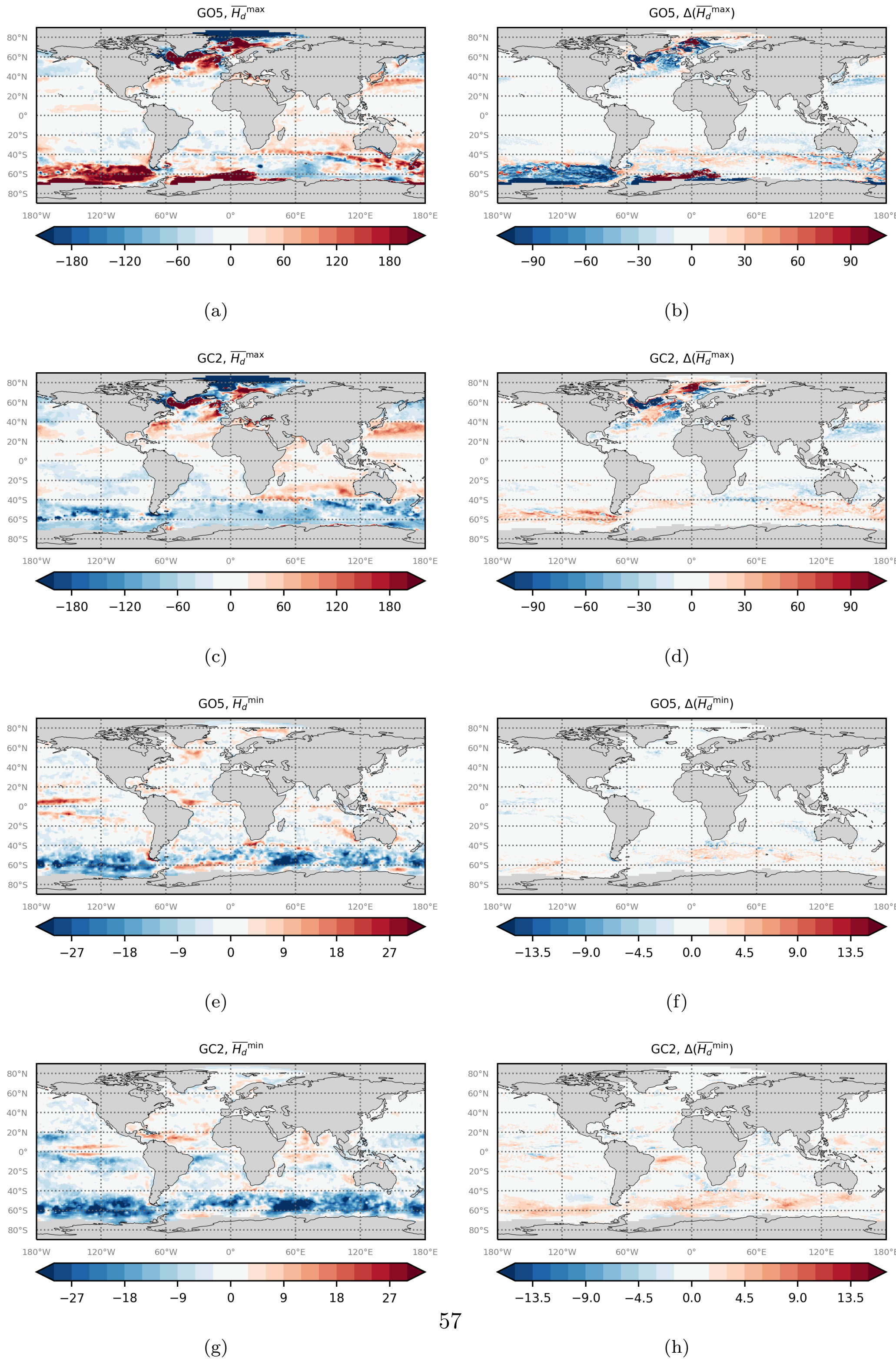


Figure 13: Mixed layer depth error (m) calculated using the climatology of de Boyer Montégut et al. (2004) in the 1/4° GO5 and GC2 simulations without the SMLEP (left) and change in magnitude of the error (such that negative and positive changes respectively indicate a reduction and increase in error) in the simulations with the SMLEP (right).