Ocean Modelling 39 (2011) 200–207

Contents lists available at ScienceDirect

Ocean Modelling

journal homepage: www.elsevier.com/locate/ocemod

Comparison of four mixed layer mesoscale parameterizations and the equation for an arbitrary tracer

V.M. Canuto [a,b,*], M.S. Dubovikov [a,c]

[a] NASA, Goddard Institute for Space Studies, 2880 Broadway, New York, NY 10025, United States
[b] Dept. of Applied Phys. and Math., Columbia University, New York, NY 10027, United States
[c] Center for Climate Systems Research, Columbia University, New York, NY 10025, United States

ARTICLE INFO

Article history:
Available online 20 April 2011

Keywords:
Mixed layer
Tracer equation
Eddy induced velocity

ABSTRACT

In this paper we discuss two issues, the inter-comparison of four mixed layer mesoscale parameterizations and the search for the eddy induced velocity for an arbitrary tracer. It must be stressed that our analysis is limited to mixed layer mesoscales since we do not treat sub-mesoscales and small turbulent mixing.

As for the first item, since three of the four parameterizations are expressed in terms of a stream function and a residual flux of the RMT formalism (residual mean theory), while the fourth is expressed in terms of vertical and horizontal fluxes, we needed a formalism to connect the two formulations. The standard RMT representation developed for the deep ocean cannot be extended to the mixed layer since its stream function does not vanish at the ocean's surface.

We develop a new RMT representation that satisfies the surface boundary condition. As for the general form of the eddy induced velocity for an arbitrary tracer, thus far, it has been assumed that there is only the one that originates from the curl of the stream function. This is because it was assumed that the tracer residual flux is purely diffusive.

On the other hand, we show that in the case of an arbitrary tracer, the residual flux has also a skew component that gives rise to an additional bolus velocity. Therefore, instead of only one bolus velocity, there are now two, one coming from the curl of the stream function and other from the skew part of the residual flux. In the buoyancy case, only one bolus velocity contributes to the mean buoyancy equation since the residual flux is indeed only diffusive.

Published by Elsevier Ltd.

1. Introduction

In this work we discuss two issues: the intercomparison of four available mixed layer ML mesoscale parameterizations and whether the eddy induced velocity for buoyancy can also represent tracers other than buoyancy, for example, passive tracers such CO_2, CFC, etc., that form part of climate studies. We study mixed layer mesoscales only, with no reference to sub-mesoscales and small scale turbulent mixing which require parameterizations not discussed here.

As for the first item, three of the four parameterizations are expressed in terms of the stream function $\mathbf{\Psi}$ and residual flux $\mathbf{F}_r$ of the residual mean theory RMT formalism (Aiki et al., 2004; Ferrari et al., 2008, 2010, cited as A4, F8, 10), while the fourth one (Canuto et al., submitted for publication, cited as C11) is expressed in terms of the vertical–horizontal mesoscale fluxes (F_v, $\mathbf{F}_H$).[1] To carry out the model intercomparison, we need to translate the C11 (F_v, $\mathbf{F}_H$) formulation into the corresponding formulation in terms of ($\mathbf{\Psi}$, $\mathbf{F}_r$).

[1] Killworth (2005, K5) was the first to argue that in the ML flows occur mostly within on horizontal planes and thus the natural representation of the mesoscale tracer flux is in terms of the horizontal $\mathbf{F}_H$ and vertical F_v component. K5 solved the linear mesoscale dynamic equations and showed that F_v is a skew flux, i.e., its divergence yields an horizontal advection with a bolus-like velocity $\mathbf{u}^*$, while $\mathbf{F}_H$ is of the diffusion type with a mesoscale diffusivity κ_M. Given the linear character of the model, K5 was unable to derive the strength of either $\mathbf{u}^*$ and/or of κ_M. Recently, K5's analysis was extended to include the non-linear terms (C11) and the form of $\mathbf{F}_H$ and F_v *for an arbitrary tracer* in terms of the large scale fields was derived. When the tracer was the buoyancy field, the parameterization was assessed in several ways, e.g., z-profile of the eddy kinetic energy vs. WOCE data, surface eddy kinetic energy vs. T/P altimetry data, dependence of the vertical flux on the mean velocity against eddy resolving simulation data, etc. The ($\mathbf{\Psi}$, $\mathbf{F}_r$) representation has the advantage of facilitating the matching with the ocean interior at the bottom of the mixed layer while the (F_v, $\mathbf{F}_H$) representation has a different advantage. Since the dynamic equation for the EKE (eddy kinetic energy, see e.g., Boning and Budich, 1992) shows that F_v acts as a source of EKE, one can model the surface eddy kinetic energy by averaging the vertical buoyancy flux F_v over the mixed layer and then assess the result against the T/P data (Scharffenberg and Stammer, 2010).

* Corresponding author at: NASA, Goddard Institute for Space Studies, 2880 Broadway, New York, NY 10025, United States. Tel.: +1 212 6785571; fax: +1 212 6785560.

E-mail address: vcanuto@giss.nasa.gov (V.M. Canuto).

1463-5003/$ - see front matter Published by Elsevier Ltd.
doi:10.1016/j.ocemod.2011.04.008

Since the standard form of ($\mathbf{\Psi}$, $\mathbf{F}_r$) cannot be used in the ML since it does not satisfy the boundary condition $\mathbf{\Psi}(0) = 0$, we developed a new RMT formulism valid in the ML. The final result, Eq. (21), expresses (F_v, $\mathbf{F}_H$) in terms of ($\mathbf{\Psi}$, $\mathbf{F}_r$).

The results of the four models intercomparison can be summarized as follows: (a) the F8 bolus velocity does not entail ML re-stratification which is known to exist, (b) the A4, F10 bolus velocities induce re-stratification but at the lowest order in the smallness parameter h/H, A4,F10 are not different (h, H are the ML and ocean depths, respectively), (c) using the $\mathbf{\Psi}$ of A4, F8,10, we construct the corresponding $\mathbf{F}_r$'s; we reproduce the F8 result while the $\mathbf{F}_r$ of A4, F10 are new since they were not given in the original work, (d) in both A4, F10, $\mathbf{F}_r$ is diffusive with a mesoscale diffusivity that vanishes at the bottom of the ML, as expected, (e) however, their values at the surface is h/H times smaller than the value obtained by Zhurbas and Oh (2003), and finally, (f) only the C11 model accounts for wind and mean flow, which affect both the mesoscale fluxes and their kinetic energy.

Concerning the parameterization of an arbitrary tracer, we obtain the following results. In addition to a diffusive component, the residual flux exhibits a new feature, a *skew component,* which gives rise to an additional bolus velocity. There are therefore two mesoscale advection terms: one due to the bolus velocity originating from the stream function and the other from the bolus velocity originating from the skew part of the residual flux. The common assumption that there is only one bolus velocity is therefore no longer tenable (in the case of buoyancy, only the bolus velocity from the stream function contributes to the mean buoyancy equation).

2. Inapplicability of the standard RMT to the ML

Consider the model independent dynamical equation for the mean buoyancy $b = -g\rho_0^{-1}\rho$ (e.g., Ferreira et al., 2005, Eq. (1))

$$\partial_t \bar{b} + \overline{\mathbf{U}} \cdot \nabla \overline{\mathbf{b}} + \nabla \cdot \mathbf{F}(\overline{\mathbf{b}}) = -\nabla \cdot \mathbf{F}_{\mathbf{SM}}(\overline{\mathbf{b}}) - \partial_{\mathbf{z}} \mathbf{F}_{\mathbf{ss}} + \mathbf{Q} \tag{1}$$

Here, $\overline{\mathbf{U}} = (\overline{\mathbf{u}}, \overline{\mathbf{w}})$ is the mean velocity and $\mathbf{F}(b) = \overline{\mathbf{U}'b'}$ is the 3D mesoscale buoyancy flux with horizontal-vertical components $\mathbf{F}_H(b) = \overline{\mathbf{u}'b'}, F_v(b) = \overline{w'b'}$, $\mathbf{U}' = (\mathbf{u}', w')$ is the mesoscale velocity field and ∇ is the 3D nabla operator; the overbar stands for an ensemble average.[2] The first and second terms on the rhs represent the contribution due to sub-mesoscales and small scale (ss) turbulence (which we write for completeness but which we do not treat in this work), while Q stands for sources and sinks. The RMT decomposition of the buoyancy flux $\mathbf{F}(b)$ into isopycnal–diapycnal components is as follows (Andrews and McIntyre, 1976; Treguier et al., 1997; Plumb and Ferrari, 2005; Ferreira et al., 2005; F8):

$$\mathbf{F}(b) \equiv \overline{\mathbf{U}'b'} = \mathbf{\Psi} \times \nabla \bar{b} + \mathbf{F}_r(b) \tag{2}$$

where the *stream function pseudo-vector* $\mathbf{\Psi}$ and the *residual vector flux* $\mathbf{F}_r$ are defined as follows:

$$\mathbf{\Psi} = -\frac{\mathbf{F}(b) \times \nabla \bar{b}}{|\nabla \bar{b}|^2} = \frac{1}{|\nabla \bar{b}|^2}\left[(F_v \nabla_H \bar{b} - N^2 \mathbf{F}_H) \times \mathbf{e}_z - \mathbf{F}_H \times \nabla_H \bar{b}\right] \tag{3}$$

$$\mathbf{F}_r(b) = \frac{\mathbf{F}(b) \cdot \nabla \bar{b}}{|\nabla \bar{b}|^2} \nabla \bar{b} = \frac{\nabla_H \bar{b} + N^2 \mathbf{e}_z}{|\nabla \bar{b}|^2}\left[\mathbf{F}_H \cdot \nabla_H \bar{b} + F_v N^2\right] \tag{4}$$

where $\mathbf{e}_z$ is the vertical unit vector and N is the Brunt–Vaisala frequency. The first term in (2) representing the isopycnal component of the buoyancy flux, has the form of a skew flux (Griffies, 1998) and its divergence leads to an advection:

$$\nabla \cdot (\mathbf{\Psi} \times \nabla \bar{b}) = \mathbf{U}^+ \cdot \nabla \bar{b} \tag{5a}$$

where $\mathbf{U}^+$ is the eddy induced or bolus velocity:

$$\mathbf{U}^+ = \nabla \times \mathbf{\Psi}, \quad \nabla_H \cdot \mathbf{u}^+ + \partial_z w^+ = 0 \tag{5b}$$

It follows that the mesoscale buoyancy flux $\mathbf{F}(b)$ contributes to Eq. (1) as an advection and a diffusion:

$$\nabla \cdot \mathbf{F}(b) = \mathbf{U}^+ \cdot \nabla \bar{b} + \nabla \cdot \mathbf{F}_r \tag{5c}$$

In a fully adiabatic ocean, that is, one with no diabatic ML, the residual flux $\mathbf{F}_r$ is negligible and thus the main mesoscale effect is represented by the first, advective, term in (5c). As McDougall and McIntosh (2001) showed, at the ocean's surface the stream function satisfies the boundary condition $\mathbf{\Psi}(0) = 0$. We concentrate on the horizontal component of this condition:

$$\mathbf{\Psi}_H(0) = 0 \tag{6a}$$

since (6a) ensures the vanishing of the vertical component of the eddy induced velocity (5b) at the ocean's surface:

$$w^+(0) = \mathbf{e}_z \cdot [\nabla_H \times \mathbf{\Psi}_H(0)] = 0 \tag{6b}$$

Furthermore, condition (6a) leads to the vanishing of the vertical component of the isopycnal flux at the surface:

$$\mathbf{\Psi}(0) \times \nabla \bar{b} \cdot \mathbf{e}_z = 0 \tag{6c}$$

Since the vertical component of the full flux also vanishes at the surface:

$$\mathbf{F}(0) \cdot \mathbf{e}_z \equiv F_v(0) = 0 \tag{7}$$

an analogous boundary condition must be satisfied by the residual flux:

$$\mathbf{F}_r(0) \cdot \mathbf{e}_z = 0 \tag{8}$$

How does the presence of the diabatic ML affect conditions (6)–(8)? To answer the question, we consider $\mathbf{\Psi}_H$ near surface. Since the last term in (3) does not contribute to $\mathbf{\Psi}_H$ and the first term is very small, strictly, it vanishes at $z = 0$ because of (7), we consider the term:

$$\mathbf{\Psi}_H(z) = -\frac{N^2(z)\mathbf{F}_H(z) \times \mathbf{e}_z}{N^4(z) + |\nabla_H \bar{b}|^2} \tag{9a}$$

Near the surface, the horizontal flux $\mathbf{F}_H$ does not vanish as it follows from the observational result by Zhurbas and Oh (2003):

$$\mathbf{F}_H(0) = -\kappa_s \nabla_H \bar{b}, \quad \kappa_s \equiv \kappa_M(0) = C\ell K^{1/2}(0) \tag{9b}$$

who arrived at it using data from the Global Drifter Program/Surface Velocity Program. In (9b), $\kappa_M(z)$ is the mesoscale diffusivity and $C = 1.02 \pm 0.13$; furthermore $\ell = \min(r_d\, L_R)$ where r_d is the Rossby deformation radius and L_R is the Rhines scale. The surface mesoscale eddy kinetic energy $K(0)$ can be obtained from the T/P data (Scharffenberg and Stammer, 2010). Since in the ocean, even near the surface, one has $s < 1$ (typical isopycnal slopes below the ML are of the order of 10^{-3} while in the ML they are about an order of magnitude larger) from (9a,b) we conclude that condition (6a) and therefore (6b,c), are not satisfied. In addition, the interpretation of the divergence of the skew flux (5a) as an advection becomes problematic. In fact, in the limit $s \ll 1$, from (9a,b) we obtain that the vertical bolus velocity becomes:

[2] The average in the mesoscale flux (1) stems from the non-linear term in the equation for the instantaneous buoyancy field when averaged over a coarse resolution grid cell of horizontal scale ~100 km. However, averaging over the grid cell in (1) is not sufficient in, say, testing a mesoscale flux parameterization using high resolution simulations. The reason is that the diagnosed fluxes are random functions of the large scale fields which are the ones averaged over the grid cell. To obtain deterministic functions from high resolution data, one needs to ensemble average the random functions. If the large scale fields are stationary and/or homogeneous for sufficiently long time and/or within large area, the diagnosed instantaneous fluxes averaged over the grid cell may be further averaged over the corresponding time and/or space intervals instead of ensemble averaging. Below, we imply instantaneous averages over the grid cell together with ensemble averages.

$$w^+(0) = \mathbf{e}_z \cdot [\nabla_H \times \mathbf{\Psi}_H(0)] = -\nabla_H \cdot (\kappa_s \mathbf{s}) \tag{9c}$$

which is a diffusive not advective term.

Next, let us consider the limiting case $N^2(0) = 0$. Although from (9a) it follows that the (6a) is satisfied, the following difficulty arises. In the ML, the variables $|\nabla_H \bar{b}| \sim 10^{-8}\ s^{-2}$ and $\mathbf{F}_H(z) = -\kappa_M \nabla_H \bar{b}$, are almost z-independent while $N^2(z)$ increases from its assumed surface value $N^2(0) = 0$ to $N^2(-h) \sim 10^{-5}\ s^{-2}$ at the bottom of the ML. It follows that the function $\mathbf{\Psi}_H$ vs. N^2 given in relation (9a) has a maximum at $N^2 = |\nabla_H \bar{b}|$, at a depth denoted by z_m:

$$z = z_m \quad N^2 = \nabla_H b \sim 10^{-8}\ s^{-2} \ll N^2(-h) \sim 10^{-5}\ s^{-2} \tag{10a}$$

Substituting this relation into (9b,a), we obtain:

$$\mathbf{\Psi}_H(z_m) \approx \frac{1}{2}\kappa_s \frac{\nabla_H \bar{b}}{|\nabla_H \bar{b}|} \times \mathbf{e}_z \tag{10b}$$

Using this result, we compute the average of the bolus velocity from the surface to the depth z_m:

$$\langle \mathbf{u}^+ \rangle_{z_m} = \mathbf{e}_z \times < \partial_z \mathbf{\Psi} >_{z_m} = \mathbf{e}_z \times \mathbf{\Psi}(z_m)/z_m = \frac{1}{2}\kappa_s \frac{\nabla_H \bar{b}}{|\nabla_H \bar{b}|} z_m^{-1} \tag{10c}$$

Thus, near the surface, the bolus velocity is estimated to be:

$$z = 0: \quad \langle \mathbf{u}^+ \rangle_{z_m} = \frac{1}{2}\frac{\kappa_s}{z_m} \sim 10\ ms^{-1} \tag{10d}$$

where we have taken $\kappa_s \sim 10^3\ m^2\ s^{-1}$and $z_m \sim 100$ m. Eq. (10d) is the difficulty we alluded to earlier since it represents an unphysical result. In the next section we show that the following modification of the RMT functions:

$$\mathbf{\Psi}, \mathbf{F_r} \to \widetilde{\mathbf{\Psi}}, \widetilde{\mathbf{F}}_\mathbf{r} \tag{10e}$$

solves the problem by ensuring the compliance of conditions (6a,c) and how the stream function $\widetilde{\mathbf{\Psi}}$ now varies smoothly from zero at the surface to $|\mathbf{\Psi}(-h)| \sim 1\ m^2\ s^{-1}$at the bottom of the ML. This behavior dispenses with the need to use artificial tapering schemes (Griffies et al., 2005) whose arbitrariness resulted in a widely varying characteristics of the simulated flows (Gnanadesikan et al., 2007; F8). It must also be noted that over the years, several authors argued for the need of the new (tilde) variables, e.g., Treguier et al. (1997), Held and Schneider (1999), Plumb and Ferrari (2005) and F8 even though the tilde itself was not explicitly used.

3. Formulation of an RMT for the mixed layer

In order to satisfy $\mathbf{\Psi}(0) = 0$ while keeping (2), we carry out the following transformations:

$$\mathbf{\Psi} \to \widetilde{\mathbf{\Psi}} = \Psi + \mathbf{A}, \quad \mathbf{F}_r \to \widetilde{\mathbf{F}}_r = \mathbf{F}_r - \mathbf{A} \times \nabla \bar{b} \tag{11}$$

Here, $\mathbf{A}$ is an arbitrary function of space and time for whose determination we follow the following principles. Since there is no reason to modify the RMT formulated for the deep ocean, we search for a $\widetilde{\mathbf{\Psi}}$ that is almost identical to $\mathbf{\Psi}$ in the deep ocean but satisfies the condition $\mathbf{\Psi}(0) = 0$. This in turn requires that the function $\mathbf{A}$ vanish in the deep ocean and that it cancels $\mathbf{\Psi}$ at the surface. Since the solution of the problem is not unique, additional conditions must be imposed. We choose $\widetilde{\mathbf{\Psi}}$ to be a 2D pseudo-vector and following F10, we introduce the vector **Y:**

$$\widetilde{\mathbf{\Psi}} = \mathbf{Y} \times \mathbf{e}_z \tag{12a}$$

Thus, the eddy induced velocity defined in Eq. (5b) with $\mathbf{\Psi} \to \widetilde{\mathbf{\Psi}}$, is expressed as follows:

$$\mathbf{u}^+ = \partial_z \mathbf{Y}, \quad w^+ = -\nabla \cdot \mathbf{Y} \tag{12b}$$

The general form of $\mathbf{A}$ that satisfies both the above requirements is the following:

$$|\nabla_H \bar{b}|^2 \mathbf{A} = N^2 \mathbf{F}_r \times \mathbf{e}_z + \mathbf{F}_H \times \nabla_H \bar{b} + \mathbf{D} \tag{13}$$

where $\mathbf{D}$ is an arbitrary 2D pseudo-vector that satisfies the condition $\mathbf{D}(0) = 0$ and that vanishes in the interior. It is worth mentioning that F8 used only the first term in (13). Next, we consider the $\mathbf{D} = 0$ case and show that it does not satisfy physical requirements that, in addition to the boundary conditions, must be met. We then suggest an expression for $\mathbf{D}$ that solves the problem.

3.1. The $\mathbf{D} = 0$ case

Substituting (11) and (13) into Eqs. (3) and (4), we obtain the following results[3]:

$$\mathbf{Y} = -|\nabla \bar{b}|^{-2} N^2 \mathbf{F}_H^{tr} + |\nabla_H \bar{b}|^{-2} F_v \nabla_H \bar{b} \tag{14}$$

$$\widetilde{\mathbf{F}}_r = \mathbf{F}_H^\ell + |\nabla_H \bar{b}|^{-2} F_v N^2 \nabla_H \bar{b} + |\nabla_H \bar{b}|^2 |\nabla \bar{b}|^{-2} \mathbf{F}_H^{tr} \tag{15}$$

At the surface, both F_v and $\mathbf{F}_H^{tr}$ vanish, the former by construction, the latter because relation (9b) shows that at $z = 0$ the flux $\mathbf{F}_H$ is directed along the horizontal gradient of the mean buoyancy and thus, has no transverse component. From relations (12a) and (14), it then follows the desired relation $\widetilde{\mathbf{\Psi}}(0) = 0$ which, in terms of $\mathbf{Y}(z)$, becomes $\mathbf{Y}(0, -H) = 0$. Next, if we multiply (14) scalarly by $\nabla_H \bar{b}$, we obtain the following two relations:

$$F_v = \mathbf{Y}^\ell \cdot \nabla_H \bar{b}, \quad N^2 \mathbf{F}_H^{tr} = -|\nabla \bar{b}|^2 \mathbf{Y}^{tr} \tag{16}$$

which, once substituted into (15), yield the following result:

$$\widetilde{\mathbf{F}}_r = \mathbf{F}_H^\ell + N^2 \mathbf{Y}^\ell + |\nabla_H \bar{b}|^2 N^{-2} \mathbf{Y}^{tr} \tag{17}$$

This relation connects the stream function to the residual flux and shows that the parameterization of the two variables cannot be chosen independently.[4] In the next sections we will discuss the use of analogous relations to study the A4, F8, 10, C11 models. Finally, we note that in relations (16) and (17), when N^2 vanishes, there is a singularity unless we have that:

$$\mathbf{Y}^{tr}(N^2 = 0) = 0 \tag{18}$$

As we discuss in Section 4, this condition is indeed satisfied in F8 but not in A4 and F10. Thus, to avoid singularities in the latter cases, we need to include an appropriate function $\mathbf{D}$ into (13).

3.2. The $\mathbf{D} \neq 0$ case

Let us consider the choice:

$$|\nabla \bar{b}|^2 \mathbf{D} = N^{-2} |\nabla_H \bar{b}|^2 \mathbf{e}_z \times \mathbf{F}_H^{tr} \tag{19}$$

which is allowed only if $\mathbf{F}_H^{tr}$ vanishes not only at $z = 0$ for any N but also when $N = 0$ for any z. In particular, in the C11 model discussed in Section 4, this condition is indeed satisfied. Then, instead of relations (14) and (15), we now have:

$$\mathbf{Y} = -N^{-2} \mathbf{F}_H^{tr} + F_v |\nabla_H \bar{b}|^{-2} \nabla_H \bar{b} \tag{20a}$$

$$\widetilde{\mathbf{F}}_r = \mathbf{F}_H^\ell + F_v N^2 |\nabla_H \bar{b}|^{-2} \nabla_H \bar{b} \tag{20b}$$

Thus, instead of (16) and (17), we now have the final relations:

$$\widetilde{\mathbf{\Psi}} = (\mathbf{Y}^{tr} + \mathbf{Y}^\ell) \times \mathbf{e}_z, \quad \mathbf{Y}^{tr} = -N^{-2} \mathbf{F}_H^{tr},$$

$$\mathbf{Y}^\ell = \frac{F_v(b)}{|\nabla_H \bar{b}|^2} \nabla_H \bar{b}, \quad \widetilde{\mathbf{F}}_r = \mathbf{F}_H^\ell(b) + N^2 \mathbf{Y}^\ell \tag{21}$$

[3] To simplify the notation, we have introduced the transverse and longitudinal components of an *arbitrary horizontal vector* $\mathbf{V}$ which are defined as follows:$\mathbf{V} = \mathbf{V}^{tr} + \mathbf{V}^\ell$, $\left|\nabla_H \bar{b}\right|^2 \mathbf{V}^\ell = (\mathbf{V} \cdot \nabla_H \bar{b}) \nabla_H \bar{b}$.

[4] It is instructive to consider a zonal flow when $\nabla_H \bar{b} = \partial_y \bar{b} \mathbf{e}_y$, $\overline{\mathbf{u}'b'} = \overline{v'b'} \mathbf{e}_y$. Then, from (14) and (15) we obtain $\widetilde{\mathbf{\Psi}} = \widetilde{\Psi} \mathbf{e}_x$, $\widetilde{\mathbf{F}}_r = \widetilde{F}_r \mathbf{e}_y$, $\widetilde{\mathbf{F}}_r = \overline{v'b'} + N^2 \widetilde{\mathbf{\Psi}}$. The last relation links the stream function to the residual flux and coincides with the results of Held and Schneider (1999) who suggested it using heuristic arguments.

The problem of finding the stream function and the residual flux corresponding to a given vertical and horizontal fluxes, is solved. Specifically, the transverse component of $\mathbf{F}_H$ and F_v yield the stream function $\widetilde{\mathbf{\Psi}}$; the longitudinal component of $\mathbf{F}_H$ then yields the residual flux. Relations (21) show that the stream function and the residual flux depend on each other and so do their parameterizations. To highlight the difference between $\mathbf{\Psi}$ and $\widetilde{\mathbf{\Psi}}$, we rewrite them as follows:

$$\textbf{Old RMT}: \quad \mathbf{\Psi} = aN^{-4}F_v\nabla_H\overline{b}\times\mathbf{e}_z - aN^{-2}\mathbf{F}_H\times\mathbf{e}_z - aN^{-4}\mathbf{F}_H\times\nabla_H\overline{b} \tag{22}$$

$$\textbf{New RMT}: \quad \widetilde{\mathbf{\Psi}} = \widetilde{a}N^{-4}F_v\nabla_H\overline{b}\times\mathbf{e}_z - N^{-2}\mathbf{F}_H\times\mathbf{e}_z - \widetilde{a}N^{-4}(\mathbf{F}_H\cdot\mathbf{s})\times\nabla_H\overline{b}\times\mathbf{e}_z \tag{23}$$

where $a = (1+s^2)^{-1}$ and $\widetilde{a} = s^{-2}$. Within a down-gradient model the last term in (22) vanishes. Since $F_v(0) = 0$, the first term in (22) vanishes while the second term does not vanish since $\mathbf{F}_H(0)\neq 0$, see relation (9b). Thus, $\mathbf{\Psi}(0)\neq 0$, as already discussed. On the other hand, using a down-gradient model of the type $\mathbf{F}_H(b) = -\kappa_M\nabla_H\overline{b}$, the last two terms in (23) cancel each other out while the first term vanishes at $z = 0$. This results in $\widetilde{\mathbf{\Psi}}(0) = 0$, as required.

4. Comparison of four parameterizations

4.1. F8. model

This model assumes that between the ML of thickness h and the deep interior, there is a transition layer (TL) of thickness $D < h$. In the ML, the stream function $\widetilde{\mathbf{\Psi}}$ is assumed to be linear in z and to vanish at the surface. In the TL, $\widetilde{\mathbf{\Psi}}$ is parameterized as a quadratic function of z. At the boundary between ML and TL, $\widetilde{\mathbf{\Psi}}$ and its z derivative are required to be continuous while at $-z_I = H + D$, which is the boundary between TL and the interior, $\widetilde{\mathbf{\Psi}}$ and its z-derivative are required to match those of the GM model; the horizontal buoyancy flux is assumed to be of the down-gradient type in all three regimes ML, TL and deep ocean. This assumption allowed the authors to get rid of the vertical and transverse components of $\mathbf{\Psi}$ before the transformation (11) was carried out. This feature, in turn made the second and third term in (13) redundant and thus the definitions of the modified stream function and residual flux have the forms given by Eqs. (12a), (14), and (15). On this basis, the model yields (see Eqs. (25) and (28) of F8):

$$\mathbf{Y}(z) = \kappa_{GM}G(z)\frac{\nabla_H\overline{b}}{N_I^2}, \quad \widetilde{\mathbf{F}}_r = -\kappa_{GM}\left[1 - G(z)\frac{N^2(z)}{N_I^2}\right]\nabla_H\overline{b} \tag{24}$$

where $G(z)$ is given in Eq. (26) of F8. As one can see, (24) satisfies relation (17) in which $\mathbf{Y}^{tr} = 0$, as we discussed below Eq. (17). To make (24) in the ML and TL more transparent, we write the function $G(z)$ in the limit $D \ll h$ which is in accordance with F8's estimates of D and h given in their Fig. 5a and b. In addition, we assume that within the TL, the z-derivative of N^2 is constant (in the ML, F8 assume that $N^2 \ll N_I^2$ is constant). This implies that in the TL, $D^{-1} \equiv -N_I^{-2}\partial_z N^2$. Adopting a rigid lid approximation, relation (26) of F8 becomes:

$$-h \leqslant z \leqslant 0: \quad G(z) \approx -\frac{3}{2}\frac{z}{h}; \quad -(h+D) \leqslant z \leqslant -h: \quad G(z) \approx -\frac{3}{2}\frac{z}{h} - \frac{(z+h)^2}{2D^2} \tag{25}$$

Substituting (24) and (25) into (12b) and since the main contribution to the z-derivative comes from the differentiation of $G(z)$, we obtain that in the ML:

$$\text{ML}: \quad \mathbf{u}^+ \approx -\frac{3}{2}\kappa_{GM}\frac{\nabla_H\overline{b}}{hN_I^2} \tag{26}$$

which does not depend on z. Using $N^2 = \partial\overline{b}/\partial z$, from Eq. (1) we obtain:

$$\frac{\partial N^2}{\partial t} = -\partial_z(\mathbf{u}^+\cdot\nabla\overline{b}) - \partial_z(\nabla\cdot\widetilde{\mathbf{F}}_r) + \cdots \tag{27}$$

Because both $\mathbf{u}^+$ *and* $\nabla\overline{b}$ *are z- independent, the eddy induced advection does not contribute to the ML stratification.* This result was one of the motivations for the F10 model we discuss next.

4.2. A4 and F10 models

These models suggest the following differential equations for $\mathbf{Y}(z)$:

$$\text{A4}: \quad C^{-1}\partial_{zz}^2\mathbf{Y} = -\nabla_H\overline{b}, \quad \text{F10}: \quad (c^2\partial_{zz}^2 - N^2)\mathbf{Y} = -\kappa_{GM}\nabla_H\overline{b} \tag{28a, b}$$

Since at any depth, the solutions of (28a,b) depend on $\nabla_H\overline{b}$ at all depths and since $\nabla_H\overline{b}$ has different directions at different depths, $\mathbf{Y}$ has both longitudinal and transverse components. As it follows from (17), in the $\mathbf{D} = 0$ case the transverse component of $\mathbf{Y}$ leads to singularities in $\mathbf{F}_r$. Thus, the A4, F10 models need the choice (19) which avoids the singularities, as Eqs. (21) show.

As one can see, the difference between the two models is the term $N^2(z)$. In the opinion of the authors of F10, in the A4 model without such a term, "*information about the background stratification, and hence about the mixed layer, would be lost.*" Below, we study analytically the A4, F10 models and reach the following conclusions: (1) to the first order in h/H, in both A4 and F10, the ML stream function does not depend on the stratification N^2 and, (2) such a dependence is present in F10 but it comes into play only to higher orders in h/H (h, H are the ML and ocean depths). Thus, F10 differs from A4 in those locations where $h \sim H$, while in the rest of the ocean the two models are quite similar. To carry out the analysis, we employ $N^2(z) = N_I^2 = \text{const.}$ in the ocean's interior and $N_{ML}^2 < N_I^2$. As for the horizontal buoyancy gradients in the ML and interior, we assume that they are different $\nabla_H\overline{b}_M \neq \nabla_H\overline{b}_I$ but z-independent in their respective regimes. F10 suggest that the speed c in (28b) be chosen as the first baroclinic velocity and thus, to the main order in h/H, we have $c = N_I H/\pi$. We have found that the ML solution of Eq. (28b) with the boundary condition $\mathbf{Y}(0, -H) = 0$ is given by[5]:

$$\mathbf{Y} = -\kappa_{GM}\zeta(a\mathbf{A} + \mathbf{B}\zeta), \quad \zeta = \pi z/H \tag{29}$$

where the two vectors **A, B** and the constant a are defined as follows:

$$N_I^2\mathbf{A} = \nabla_H\overline{b}_I, \quad N_I^2\mathbf{B} = \frac{1}{2}\nabla_H\overline{b}_M, \quad a = \tanh\pi/2 \tag{30}$$

In (29), we kept the term $\sim\zeta$ since it is the only one that contributes to the re-stratification of the ML, as we show below. In the A4 model, $a = \pi/2$. In both models, a was determined by solving the boundary-value problem between the surface and the bottom. *To the main order in h/H, these results do not depend on the ML stratification N.* The bolus velocity and its z-derivative contributing to the ML re-stratification, are obtained from (12a) and (29):

$$\mathbf{u}^+ = -\pi\kappa_{GM}H^{-1}(a\mathbf{A} + 2\mathbf{B}\zeta), \quad -\partial_z\mathbf{u}^+\cdot\nabla_H\overline{b} = 4\pi^2|\mathbf{B}|^2\kappa_{GM}\frac{N_I^2}{H^2} > 0 \tag{31}$$

[5] In the Eady case corresponding to $N^2(z) = N_I^2 = \text{const.}$ throughout the whole ocean treated in F10, the solution of Eq. (28b) is given by Eqs. (38) and (39) of F10. When expanded in power of z/H, the result $\mathbf{Y}_{Eady} \approx -\kappa_{GM}\mathbf{A}\zeta(\tanh\pi/2 + \zeta/2)$ coincides with (29) since in the Eady case $2\mathbf{A} = \mathbf{B}$.

In contrast to the F8 model, the second of (31) leads to a ML re-stratification.

Though in A4 and F10 the residual flux was not presented, it can be found on the basis of our last relation (21) and the solutions of (28a,b). Though A4 and F10 do not give the residual flux, it can be found using their stream function. First, we study $\mathbf{F}_{\mathrm{H}}^{\ell}$ and consider the top of the ocean interior where $\widetilde{\mathbf{F}}_{\mathrm{r}} \approx \mathbf{F}_{\mathrm{r}} \approx 0$. From the fourth of (21), we then obtain:

$$z = z_{\mathrm{I}} \quad \mathbf{F}_{\mathrm{H}}^{\ell} \approx -N^2 \mathbf{Y}^{\ell} \tag{32a}$$

On the other hand, since inside the ML the variables $\mathbf{u}'$, b' and $\nabla_{\mathrm{H}}\overline{b}$ are almost z-independent, in the region $z \geqslant z_{\mathrm{I}}$ we may use (9b) and extend it all the way to the bottom of the ML. Thus, using the approximate relation:

$$z \geqslant z_{\mathrm{I}}: \quad \mathbf{F}_{\mathrm{H}}^{\ell} \approx -\kappa_{\mathrm{s}} \nabla_{\mathrm{H}} \overline{b} \tag{32b}$$

and substituting it into (32a), we can estimate the surface value of the mesoscale diffusivity:

$$\kappa_{\mathrm{s}} \approx N^2 \mathbf{Y}^{\ell} \cdot \nabla_{\mathrm{H}} \overline{b} / \left| \nabla_{\mathrm{H}} \overline{b} \right|^2 \Big|_{\mathrm{I}} \tag{32c}$$

where all the variables are computed at the bottom of the mixed layer. Substituting (32b,c) into the last of (21) we obtain:

$$\mathrm{A4, F10}: \quad \widetilde{\mathbf{F}}_{\mathrm{r}} = -\kappa_{\mathrm{r}}(z) \nabla_{\mathrm{H}} \overline{b} \tag{33a}$$

$$\kappa_{\mathrm{r}} \approx \kappa_{\mathrm{s}} \left(1 - \frac{\mathbf{Y} \cdot \nabla_{\mathrm{H}} \overline{b}}{\mathbf{Y}_{\mathrm{I}} \cdot \nabla_{\mathrm{H}} \overline{b}_{\mathrm{I}}} \frac{|\nabla_{\mathrm{H}} \overline{b}_{\mathrm{I}}|^2}{|\nabla_{\mathrm{H}} \overline{b}|^2} \frac{N^2(z)}{N_{\mathrm{I}}^2} \right) \approx \kappa_{\mathrm{s}} \left(1 + \frac{z}{h} \frac{N^2(z)}{N_{\mathrm{I}}^2} \right) \tag{33b}$$

where in the last step we used Eqs. (29) and (30). Thus, the residual flux Eq. (33) is represented by a horizontal diffusion with a z-dependent mesoscale diffusivity which equals κ_{s} at the surface and vanishes at $z = -\mathrm{h}$. To evaluate κ_{s}, we substitute (29) and (30) into (32c) with the results:

$$\frac{\kappa_{\mathrm{s}}(\mathrm{F10})}{\kappa_{\mathrm{GM}}} \sim \pi[\tanh(\pi/2)](\mathrm{h/H}), \quad \frac{\kappa_{\mathrm{s}}(\mathrm{A4})}{\kappa_{\mathrm{GM}}} \sim \frac{1}{2}\pi^2 (h/H) \tag{34a}$$

Since relations (34a) were obtained for the particular case when in the interior the buoyancy gradient is z independent, they should be viewed as estimates which we indicate by $\sim$. These results show that in the case of a sufficiently shallow ML where $h/H \ll 0.2$, we have:

$$\mathrm{A4,\ F10}: \quad \kappa_{\mathrm{s}} \ll \kappa_{\mathrm{GM}} \tag{34b}$$

While the commonly used value $\kappa_{\mathrm{GM}} \sim 10^3\ \mathrm{m^2\,s^{-1}}$ (e.g., Ferreira et al., 2005) is consistent with the second relation (9b), use of the typical values $r_{\mathrm{d}} = 30$ km and $K = 10^{-2}\ \mathrm{m^2\,s^{-2}}$ in relation (34b) entails a surface diffusivity much lower than what is observed. This conclusion deserves some comments. We arrived at (34a,b) thanks to relations (21) that allows us to relate the stream function of the A4–F10 models to the horizontal flux which in turn we related to (9b) derived from observations. Since at the time A4 and F10 were presented, relation (21) was not known, the translation of $\mathbf{\Psi}$ into a horizontal flux and then the comparison with the data based relation (9b), was not possible, conclusions (34) could not be reached and went unnoticed. In conclusion:

(a) the F8 eddy induced velocity does not entail ML re-stratification which is known to exist,
(b) the A4 and F10 eddy induced velocities yield re-stratification but, in spite of their formal difference, at the lowest order in the small parameter h/H, neither model exhibits a dependence of the re-stratification on the ML $N^2(z)$. Therefore, A4 and F10 are not different,
(c) the residual flux was not presented in A4 and F10. We have shown that in both A4,F10 it is of the diffusive type with a mesoscale diffusivity that vanishes at the bottom of the ML,
(d) however, the value of the surface diffusivity in such models is roughly h/H times smaller than what Zhurbas and Oh (2003) have derived using drifter data.

4.3. C11 model[6]

The C11 results for the horizontal and vertical buoyancy fluxes are:

$$\mathbf{F}_{\mathrm{H}}(b) = -\kappa_{\mathrm{M}} \nabla_{\mathrm{H}} \overline{b}, \quad F_{\mathrm{v}}(b) = -\boldsymbol{\kappa} \cdot \nabla_{\mathrm{H}} \overline{b} \tag{35a}$$

where:

$$\boldsymbol{\kappa} = \kappa_{\mathrm{M}} z \widehat{\mathbf{F}}(z) \tag{35b}$$

$$f r_{\mathrm{d}}^2 \widehat{\mathbf{F}}(z) = (\widehat{\mathbf{u}} + \widetilde{\mathbf{u}}) \times \mathbf{e}_z, \quad \widetilde{\mathbf{u}}(z) = \overline{\mathbf{u}}(\mathbf{z}) - \mathbf{u_d}$$

$$z \widehat{\mathbf{u}}(z) = \int_0^z \overline{\mathbf{u}}(z') dz', \quad \langle \cdot \rangle \equiv \int_{-H}^{-h} \cdot K^{1/2}(z) \mathrm{d}z \Big/ \int_{-H}^{-h} K^{1/2}(z) \mathrm{d}z \tag{35c}$$

The variable $\mathbf{u}_{\mathrm{d}}$ represents the mesoscale "drift velocity" which in the f-plane has the following expression:

$$\mathbf{u}_{\mathrm{d}} = \langle \overline{\mathbf{u}} \rangle - \frac{1}{2} f r_{\mathrm{d}}^2 \mathbf{e}_z \times < \partial_z \mathbf{s} > \tag{35d}$$

The above relations exhibit an interesting physical feature. Since mesoscale eddies move with their own drift velocity $\mathrm{u_d}$, the effective mean velocity is the one in the frame co-moving with the mesoscales and that is why the mean velocity enters as in the second relation in (35c). Finally, the mesoscale diffusivity is given by:

$$\kappa_{\mathrm{M}}(z) = f(\overline{u}, K) \ell K^{1/2}(z) \tag{36a}$$

where:

$$f(\overline{u}, K) = \left[1 + \frac{3}{4K} |\widetilde{\mathbf{u}}(z)|^2 \right]^{-1}, \quad \ell = \min(r_{\mathrm{d}}, L_{\mathrm{R}}) \tag{36b}$$

When the eddy kinetic energy K is larger than that of the mean flow, see Fig. 7 of Scharffenberg and Stammer (2010), the function $f(\overline{u}, K)$ is close to unity.

To present the C11 parameterization in terms of the RMT formalism, we first remark that from (35) and (21) we have:

$$\mathbf{F}_{\mathrm{H}}^{\mathrm{tr}} = 0, \quad \mathbf{Y}^{\mathrm{tr}} = 0, \quad \mathbf{Y}^{\ell} = -\boldsymbol{\kappa}^{\ell} = -\frac{\boldsymbol{\kappa} \cdot \nabla_{\mathrm{H}} \overline{b}}{|\nabla_{\mathrm{H}} \overline{b}|^2} \nabla_{\mathrm{H}} \overline{b} \tag{37}$$

where $\boldsymbol{\kappa}^{\ell}$ is the longitudinal component of $\boldsymbol{\kappa}$, see footnote 3. Using (35b), $\mathbf{Y}$ acquires the form:

$$\mathbf{Y} = -z \kappa_{\mathrm{M}} \frac{\widehat{\mathbf{F}} \cdot \nabla_{\mathrm{H}} \overline{b}}{|\nabla_{\mathrm{H}} \overline{b}|^2} \nabla_{\mathrm{H}} \overline{b} \tag{38a}$$

To link this result with the GM model, we rewrite it as follows:

$$\mathbf{Y} = T(z) \mathbf{Y}_{\mathrm{GM}}(-h), \quad T(z) = -z \frac{N^2(-h)}{|\nabla_{\mathrm{H}} \overline{b}|^2} \widehat{\mathbf{F}} \cdot \nabla_{\mathrm{H}} \overline{b} \tag{38b}$$

where the function $T(z)$ satisfies the conditions

$$T(0) = 0, \quad T(-h) = 1 \tag{38c}$$

[6] In the case of buoyancy, the C11 parameterization was assessed in several ways: z-profile of the eddy kinetic energy vs. WOCE data, surface eddy kinetic energy vs. T/P altimetry data, dependence of the vertical flux on the full mean velocity field (not only its geostrophic part) which is affected by wind, overall assessment against eddy resolving simulation data, etc.

While the first relation is obvious, the second condition comes from the fact that at the bottom of the ML, the smallness of the measured diapycnal diffusivities (Ledwell et al., 1998, 2011) implies the vanishing of *the diapycnal flux:*

$$F_d = F_V + N^{-2}\mathbf{F}_H \cdot \nabla_H \bar{b} \tag{39}$$

Two considerations are in order here: (1) relations (38b,c) are analogous to the default tapering scheme (Griffies et al., 2005; Gnanadesikan et al., 2007) with the difference that the *tapering function* $T(z)$ is no longer arbitrary but given by the model, as the second relation in (38b) shows and, (2) we did not impose that at the bottom of the ML the stream function $\mathbf{Y}$ matches that of GM, and yet it does, as the first of (38b) shows. As for the residual flux, use of (35a,b) in the last of (21), yields the following results:

$$\widetilde{\mathbf{F}}_r = -\kappa_r \nabla_H \bar{b}, \quad \kappa_r = \kappa_M[1 - \widetilde{T}(z)], \quad \widetilde{T}(z) = \frac{N^2(z)}{N^2(-h)} T(z) \tag{40a}$$

in which the diffusivity κ_r vanishes at the bottom of the ML. Eqs. (38b) also show that in the C11 model, the profile of κ_r and the tapering function $T(z)$ relate to one another while within the default tapering scheme they are independent. Finally, we notice that in most of the ML, N^2 is sufficiently small and thus from (40a) we have:

$$z > -h \quad \kappa_r \approx \kappa_M \tag{40b}$$

a relation that we use in the next section.

For further analysis, it is convenient to decompose $\widehat{\mathbf{F}}$ into $\mathbf{F}_1$ and $\mathbf{F}_2$ components that are governed by the deep ocean and the ML respectively:

$$\widehat{\mathbf{F}} = \mathbf{F}_1 + \mathbf{F}_2, \quad \mathbf{F}_1 = -f^{-1} r_d^{-2} \mathbf{u}_d \times \mathbf{e}_z, \quad \mathbf{F}_2(z) = f^{-1} r_d^{-2} (\widehat{\mathbf{u}} + \overline{\mathbf{u}}) \times \mathbf{e}_z \tag{41a}$$

and consequently:

$$\mathbf{Y} = \mathbf{Y}_1 + \mathbf{Y}_2, \quad \mathbf{Y}_{1,2} = -\kappa_M z \mathbf{F}^{\ell}_{1,2} \tag{41b}$$

Since both terms in (35d) are of the same order of magnitude $\sim H^{-1}\mathbf{A}$, where $\mathbf{A}$ is given in (30), use of the relation $r_d = N_I H/\pi|f|$ yields (the subscript I refers to the interior):

$$\mathbf{Y}_1 \sim -\kappa_M \frac{z}{H}(\mathbf{n} \cdot \mathbf{A})\mathbf{n}, \quad \mathbf{n} \equiv |\nabla_H \bar{b}|^{-1} \nabla_H \bar{b} \tag{41c}$$

Thus, the first term of (29) and (41b) are rather similar. As for $\mathbf{Y}_2$, we distinguish three cases.

4.3.1. No wind, geostrophic flow

Using the thermal wind relation, $\overline{\mathbf{u}} = \mathbf{z}\mathbf{f}^{-1}\mathbf{e}_z \times \nabla_H \overline{\mathbf{b}}_M = 2\widehat{\mathbf{u}}$ and $r_d = N_I H/\pi|f|$, we obtain the following expression (the suffix "g" stands for geostrophic):

$$\mathbf{Y}_{2g} = -\frac{3}{2}\kappa_M \frac{\nabla_H \bar{b}_M}{N_I^2}\zeta^2, \quad \zeta = \frac{\pi z}{H} \tag{41d}$$

which, after using (30), is close to the second term in (29). Thus, in the absence of wind stresses, A4, F10 and C10 yield similar results.

4.3.2. Wind

In the presence of wind, the situation changes significantly since there are wind stresses and horizontal gradients of the surface pressure which, in turn, result in a *geostrophic surface velocity* which, as shown in C11, dominates the production of ML eddy kinetic energy. In the presence of a geostrophic surface velocity (denoted by the subscript sv), the term $\mathbf{Y}_2$ has an additional term that is not present in the A4, F10 models:

$$\mathbf{Y}_{sv} = -\frac{\kappa_M}{2fr_d^2} z(\overline{\mathbf{u}}_{\mathbf{gs}} \times \mathbf{e}_z \cdot \mathbf{n})\mathbf{n} \tag{42a}$$

where $\overline{\mathbf{u}}_{\mathbf{gs}}$ is the geostrophic component of the mean surface velocity.

4.3.3. Ekman layer

The presence of an Ekman layer, represented here for simplicity by an Ekman spiral, contributes a term given by:

$$\mathbf{Y}_E = |\nabla_H \bar{b}|^{-2} F_E \nabla_H \bar{b}, \quad F_E = A(\eta)\boldsymbol{\tau} \cdot \nabla_H \bar{b} + B(\eta)\nabla_H \bar{b} \times \boldsymbol{\tau} \cdot \mathbf{e}_z \tag{42b}$$

where $\boldsymbol{\tau}$ is the wind stress, $\eta = z/\delta_E$, δ_E is the thickness of the Ekman layer and:

$$\begin{aligned} A(\eta) &= \frac{\kappa_M}{r_d^2 f^2 \rho_0}(1 - e^{\eta}\cos\eta) + \eta e^{\eta}(\sin\eta - \cos\eta) \\ B(\eta) &= \frac{\kappa_M}{r_d^2 f^2 \rho_0}(f/|f|)[\eta(\cos\eta + \sin\eta) + sin\eta]e^{\eta} \end{aligned} \tag{42c}$$

Thus, the complete form of the stream function is given by:

$$\mathbf{Y} = \mathbf{Y}_1 + \mathbf{Y}_{2g} + \underbrace{\mathbf{Y}_{sv} + \mathbf{Y}_E}_{New\ terms} \tag{43}$$

While the terms $\mathbf{Y}_1, \mathbf{Y}_{2g}$ in (41c,d) are similar to the corresponding terms in A4–F10, the terms $\mathbf{Y}_{sv}$, $\mathbf{Y}_E$ given in (42a,b) are absent in A4, F10. One should distinguish $\mathbf{Y}_{sv}$, $\mathbf{Y}_E$ from the terms appearing in the mean buoyancy equation due to the presence of wind dependent terms in the mean velocity. While the latter stem from the direct effect of the wind stresses on the mean momentum equation, $\mathbf{Y}_{sv}$, $\mathbf{Y}_E$ originate from the non-linear interaction of mesoscales and mean velocities in the dynamic equations for the mesoscale velocities.

In summary, the main difference between C11 and A4, F10 is that in the former case, the surface diffusivity coincides with κ_M, Eq. (36), in agreement with observations, whereas in A4, F10 the surface diffusivity is much smaller than what the data indicate. The C11 model yields profiles of $\widetilde{\boldsymbol{\Psi}}$, $\widetilde{\mathbf{F}}_r$ that are analogous to those of the default tapering schemes with the difference that the profiles are no longer arbitrary but determined within the model.

5. The case of an arbitrary tracer

In the previous sections we considered the ML mesoscale parameterization for buoyancy. However, OGCMs time step tracers such as temperature, salinity, CO_2, CFC, etc. which cannot be represented as a buoyancy field. In what follows we show that the eddy induced velocity for buoyancy and arbitrary tracers are different. We begin with the equation for an arbitrary tracer:

$$\partial_t \overline{\tau} + \overline{\mathbf{U}} \cdot \nabla \overline{\tau} + \nabla \cdot \mathbf{F}(\tau) = -\nabla \cdot \mathbf{F}_{\mathbf{SM}} - \partial_{\mathbf{z}} \mathbf{F}_{\mathbf{ss}} + \mathbf{Q} \tag{44a}$$

where the rhs has the same meaning as in Eq.(1). Using the relations (35) for the case of a tracer:

$$\mathbf{F}_H(\tau) = -\kappa_M \nabla_H \overline{\tau}, \quad F_v(\tau) = -\boldsymbol{\kappa} \cdot \nabla_H \overline{\tau} \tag{44b}$$

we have:

$$\mathbf{F}(\tau) = -\kappa_M \nabla_H \overline{\tau} - (\boldsymbol{\kappa} \cdot \nabla_H \overline{\tau})\mathbf{e}_z \tag{44c}$$

Next, consider the second term in the rhs of (44c) which we rewrite as follows:

$$\begin{aligned} -(\boldsymbol{\kappa} \cdot \nabla_H \overline{\tau})\mathbf{e}_z &= -(\boldsymbol{\kappa} \cdot \nabla \overline{\tau})\mathbf{e}_z = -(\boldsymbol{\kappa} \times \mathbf{e}_z) \times \nabla \overline{\tau} - \boldsymbol{\kappa}\frac{\partial \overline{\tau}}{\partial z} \\ &= \mathbf{F}_{skew}(\tau) - \boldsymbol{\kappa}\frac{\partial \overline{\tau}}{\partial z} \end{aligned} \tag{45a}$$

where we have defined the *skew flux*:

$$\mathbf{F}_{skew}(\tau) \equiv -(\boldsymbol{\kappa} \times \mathbf{e}_z) \times \nabla \overline{\tau} \tag{45b}$$

Eq. (44c) then becomes:

$$\mathbf{F}(\tau) = -\kappa_{\mathbf{M}} \nabla_{\mathbf{H}} \overline{\tau} - \boldsymbol{\kappa}\frac{\partial \overline{\tau}}{\partial \mathbf{z}} + \mathbf{F}_{\mathbf{skew}}(\tau) \equiv \mathbf{F}_{\mathbf{diff}}(\tau) + \mathbf{F}_{\mathbf{skew}}(\tau) \tag{45c}$$

Substituting (45c) into (44a), we obtain:

$$\partial_t\overline{\tau} + (\overline{\mathbf{U}} + \mathbf{U}_*)\cdot\nabla\overline{\tau} + \nabla\cdot\mathbf{F}_{\mathbf{diff}}(\tau) = -\nabla\cdot\mathbf{F}_{\mathbf{SM}} - \partial_z\mathbf{F}_{\mathbf{ss}} + \mathbf{Q} \tag{45d}$$

where the eddy induced velocity $\mathbf{U}_* = (\mathbf{u}_*, w_*)$ is given by:

$$\mathbf{u}_* = -\frac{\partial\boldsymbol{\kappa}}{\partial z}, \quad w_* = \nabla_{\mathrm{H}}\cdot\boldsymbol{\kappa} \tag{45e}$$

We must remark that in the buoyancy case, Eq.(45d) has the form:

$$\partial_t\overline{b} + (\overline{\mathbf{U}} + \mathbf{U}^+)\cdot\nabla\overline{\mathbf{b}} + \nabla\cdot\mathbf{F}_{\mathbf{diff}}(\mathbf{b}) = -\nabla\cdot\mathbf{F}_{\mathbf{SM}} - \partial_z\mathbf{F}_{\mathbf{ss}} + \mathbf{Q} \tag{46a}$$

where ($\nabla\cdot\mathbf{U}^+ = 0$):

$$\mathbf{u}^+ \equiv \frac{\partial\mathbf{Y}}{\partial z} = -\frac{\partial\boldsymbol{\kappa}^\ell}{\partial z}, \quad w^+ = \nabla_{\mathrm{H}}\cdot\boldsymbol{\kappa}^\ell \tag{46b}$$

Since:

$$\mathbf{u}^+ = -\frac{\partial\boldsymbol{\kappa}^\ell}{\partial z} = -\frac{\partial\boldsymbol{\kappa}}{\partial z} + \frac{\partial\boldsymbol{\kappa}^{\mathrm{tr}}}{\partial z} = \mathbf{u}_* - \mathbf{u}_{**} \tag{46c}$$

it follows that

$$\mathbf{u}_* \equiv \mathbf{u}^+ + \mathbf{u}_{**} \tag{46d}$$

Therefore, the so-called "*residual velocities*" for buoyancy and arbitrary tracers in Eqs. (46a) and (45d) are given by:

$$\mathbf{u}_{\mathrm{res}}(b) = \overline{\mathbf{u}} + \mathbf{u}^+, \quad \mathbf{u}_{\mathbf{res}}(\tau) = \mathbf{u}_{\mathbf{res}}(\mathbf{b}) + \mathbf{u}_{**} \tag{46e}$$

The common assumption (e.g., Ferreira and Marshall, 2006) that there is only one bolus velocity $\mathbf{u}^+$ *is no longer valid* since in the case of an arbitrary tracer the true bolus velocity is (46d) in which the new component $\mathbf{u}_{**}$ contributes only to tracers other than buoyancy since in the latter case:

$$\mathbf{u}_{**}\cdot\nabla_{\mathrm{H}}\overline{b} = 0 \tag{46f}$$

It is important to note that the same results are obtained using the RTM formalism. In fact, substituting Eqs. (45c,b), (12a) and (37) into the definition of the tracer residual flux:

$$\widetilde{\mathbf{F}}_{\mathrm{r}}(\tau) = \mathbf{F}(\tau) - \widetilde{\boldsymbol{\Psi}}\times\nabla\overline{\tau} \tag{47a}$$

we obtain:

$$\widetilde{\mathbf{F}}_{\mathrm{r}}(\tau) = \mathbf{F}_{\mathrm{diff}}(\tau) + \widetilde{\mathbf{F}}_{\mathrm{skew}}(\tau) \tag{47b}$$

where:

$$\widetilde{\mathbf{F}}_{\mathrm{skew}}(\tau) = -\boldsymbol{\kappa}^{\mathrm{tr}}\times\nabla\overline{\tau} \tag{47c}$$

Therefore, within the C11 model, the residual flux has not only a diffusive component as all previous models, but also a skew one which gives rise to the additional bolus velocity $\mathbf{U}_{**}$. Substituting $\mathbf{F}(\tau) = \widetilde{\boldsymbol{\Psi}}\times\nabla\overline{\tau} + \widetilde{\mathbf{F}}_{\mathrm{r}}(\tau)$ into (44a), together with (12a), (37) and (47b,c), we obtain the equation:

$$\partial_t\overline{\tau} + (\overline{\mathbf{U}} + \mathbf{U}^+ + \mathbf{U}_{**})\cdot\nabla\overline{\tau} + \nabla_{\mathbf{H}}\cdot\mathbf{F}_{\mathbf{diff}} = -\nabla\cdot\mathbf{F}_{\mathbf{SM}} - \partial_z\mathbf{F}_{\mathbf{ss}} + \mathbf{G} \tag{47d}$$

which, because of (46d), is identical to (45d).

6. Summary and conclusions

In this work, we have compared four recent ML mesoscale parameterizations. In C11, the parameterization was presented in terms of the horizontal and vertical mesoscale fluxes $\mathbf{F}_{\mathrm{H}}$ and F_{v}. While $\mathbf{F}_{\mathrm{H}}$ is given by a down-gradient diffusion, F_{v} is a skew flux whose divergence yields an advection. The other three parameterizations, F8,10 and A4 were formulated in terms of a stream function and a residual flux of the residual mean theory RMT, though the last two modeled only the stream function. For this reason, in the first part, we compared the four parameterizations for the buoyancy flux only. To compare the C11 results with the other three models, we had to translate the $\mathbf{F}_{\mathrm{H}}$ and F_{v} fluxes into the mean residual theory formalism, i.e., a stream function and a residual flux. We began with discussing the fact that the definition of the latter used for the adiabatic ocean interior, must be modified in the diabatic ML in order to satisfy the homogeneous boundary condition of the stream function. The appropriate modifications and the new stream function and residual flux were discussed in Section 3. The expressions for the modified $\widetilde{\boldsymbol{\Psi}}$ and $\widetilde{\mathbf{F}}_{\mathrm{r}}$ in terms of the *V–H* fluxes are given in Eq. (21). These relations allowed us not only to compare C11 with F8,10 and A4, but will be instrumental in matching the ML mesoscale parameterization with the one in the adiabatic ocean. We shall treat this problem in subsequent work.

In the absence of the wind stresses, we have shown that the stream functions of A4, F10 and C11 are similar. In the case of *strong winds*, the C11 model yields a stream function that includes the Ekman and surface velocities whereas these two contributions are absent in the A4, F8, 10 models. The inclusion of the Ekman flow is important since eddy resolving simulations (Maltrud et al., 1998) showed that the surface eddy kinetic energy is contributed quite substantially by the Ekman flow. As for the re-stratification effects of mesoscales, we have shown that to the main order in the parameter h/H (which is small everywhere except in deep convective regimes), the models A4, F10 give the same result and neither of them exhibits a dependence on the ML stratification.

An interesting result emerges from the new stream function and residual flux, namely that they are related to each other, see Eq. (21), and, thus, parameterization of one affects the other, a feature that did not exist in the deep ocean where the residual flux is negligible.

We have checked that relation (21) is satisfied in the F8 and C11 models while the A4, F10 do not provide an explicit parameterization of the residual flux. We have shown that in order to satisfy (21), the A4 and F10 models require a surface mesoscale diffusivity that is much smaller than the results of the observations by Zhurbas and Oh (2003).

Concerning the parameterization of an arbitrary tracer, we have found the following:

(a) for OGCMs that employ vertical-horizontal fluxes in the ML, the appropriate arbitrary tracer equation is given by an equation of the form (44a) with the fluxes given by Eqs. (44b).
(b) for OGCMs that employ the RMT in the ML, the arbitrary tracer equation is Eq. (45d)
(c) within the TRM formalism, and in the case of an arbitrary tracer, the eddy induced velocity $\mathbf{u}^+$ computed as the curl of the stream function does not represent the full eddy induced velocity which is given by $\mathbf{u}_*$.

Acknowledgements

The authors would like to thanks two anonymous Referees for several suggestions that greatly helped improve the original manuscript.

References

Aiki, H., Jacobson, T., Yamagata, T., 2004. Parameterizing ocean eddy transports from surface to bottom. Geophys. Res. Lett. 31, L19302, doi:10.1029/2004GL020703.

Andrews, D.G., McIntyre, M.E., 1976. Planetary waves in horizontal and vertical shear: the generalized Eliassen–Palm relation and the mean zonal acceleration. J. Atmos. Sci. 33, 2031–2048.
Boning, C.W., Budich, R.G., 1992. Eddy dynamics in a primitive equation model. Sensitivity to horizontal resolution and friction. J. Phys. Oceanogr. 22, 361–381.
Canuto, V.M., Dubovikov, M.S., Leboissetier, A., submitted for publication. Mixed layer mesoscales: kinetic energy, diffusivity, dynamical tapering, new bolus velocity. Model development and assessment with T/P, WOCE and Drifter. Ocean Modell. (cited as C11).
Ferrari, R., McWilliams, J.C., Canuto, V.M., Dubovikov, M.S., 2008. Parameterization of eddy fluxes near ocean boundaries. J. Climate 21, 2770–2789.
Ferrari, R., Griffies, S.M., Nurser, A.J., Vallis, G.K., 2010. A boundary-value problem for the parameterized mesoscale eddy transport. Ocean Modell. 32, 143–156.
Ferreira, D., Marshall, J.M., 2006. Formulation and implementation of a residual mean ocean circulation model. Ocean Modell. 13, 86–107.
Ferreira, D., Marshall, J.M., Heimbach, P., 2005. Estimating eddy stresses by fitting dynamics to observations using a residual-mean ocean circulation model and its adjoint. J. Phys. Oceanogr. 35, 1891–1910.
Gnanadesikan, R., Griffies, S.M., Samuels, B.L., 2007. Effects in a climate model of slope tapering in neutral physics schemes. Ocean Modell. 16, 1–16.
Griffies, S.M., 1998. The Gent–McWilliams skew flux. J. Phys. Oceanogr. 28, 831–841.
Griffies, S.M., Gnanadesikan, A., Dixon, K.W., Dunne, J.P., Gerdes, R., Harrison, M.J., Rosati, A., Russell, J.L., Samuels, B.L., Spelman, M.J., Winton, M., Zhang, R., 2005. Formulation of an ocean model for global climate simulations. Ocean Sci. 1, 45–79.
Held, I.M., Schneider, T., 1999. The surface branch of the zonally averaged mass transport circulation in the troposphere. J. Atmos. Sci. 56, 1688–1697.
Killworth, P.D., 2005. Parameterization of eddy effects on mixed layers and tracer transport: a linearized eddy perspective. J. Phys. Oceanogr. 35, 1717–1725.
Ledwell, J.R., Watson, A.J., Law, C.S., 1998. Mixing of a tracer released in the pycnocline of a subtropical gyre. J. Geophys. Res. 103, 21499–21529.
Ledwell, J.R., St Laurent, L.C., Girton, J.B., Toole, J.M., 2011. Diapycal mixing in the Antarctic Circumpolar Current. J. Phys. Oceanogr. 41, 241–246.
Maltrud, M.E., Smith, R.D., Semtner, A.J., Malone, R.C., 1998. Global eddy resolving simulation driven by 1985–1995 atmospheric winds. J. Geophys. Res. 103, 30825–30853.
McDougall, T.J., McIntosh, P.C., 2001. The temporal-residual-mean velocity. Part II: Isopycnal interpretation and tracer and momentum equation. J. Phys. Oceanogr. 31, 1222–1246.
Plumb, R.A., Ferrari, R., 2005. Transformed Eulerian-mean theory. Part I: Nonquasi geostrophic theory for eddies on a zonal-mean flow. J. Phys. Oceanogr. 35, 165–174.
Scharffenberg, M.G., Stammer, D., 2010. Seasonal variations of the geostrophic flow field and of eddy kinetic energy inferred from TOPEX/POSEIDON and Jason-1Tandem Mission Data. J. Geophys. Res. 115, C2. doi:10.1029/2008JC005242.
Treguier, A.M., Held, I.M., Larichev, V.D., 1997. Parameterization of quasi geostrophic eddies in primitive equation ocean models. J. Phys. Oceanogr. 27, 567–580.
Zhurbas, V., Oh, I.S., 2003. Lateral diffusivity and Lagrangean scales in the Pacific Ocean as derived from drifter data. J. Geophys. Res. 108 (C5), 3141. doi:10.1029/2002JC001596.