Properties of the lateral mesoscale eddy-induced transport in a high-resolution ocean model: Beyond the flux-gradient relation

Yueyang Lu[a] , Igor Kamenkovich[a] , and Pavel Berloff[b]

[a] *Rosenstiel School of Marine and Atmospheric Sciences, University of Miami, Miami, Florida*

[b] *Department of Mathematics, Imperial College London, London, United Kingdom*

Corresponding author: Yueyang Lu, yueyang.lu@miami.edu

ABSTRACT: Lateral mesoscale eddy-induced tracer transport is traditionally represented in coarse-resolution models by the flux-gradient relation. In its most complete form, the relation assumes the eddy tracer flux as a product of the large-scale tracer concentration gradient and an eddy transport coefficient tensor. However, several recent studies reported that the tensor has significant spatio-temporal complexity and is not uniquely defined, that is, it is sensitive to the tracer distributions and to the presence of non-divergent ("rotational") component of the eddy flux. These issues could lead to significant biases in the representation of the eddy-induced transport. Using a high-resolution tracer model of the Gulf Stream region, we examine the diffusive and advective properties of lateral eddy-induced transport of dynamically passive tracers, re-evaluate the utility of the flux-gradient relation, and propose an alternative approach based on modeling the local eddy forcing by a combination of diffusion and generalized eddy-induced advection. Mesoscale eddies are defined by a scale-based spatial filtering, which leads to the importance of new eddy-induced terms, including eddy-mean covariances in the eddy fluxes. The results show that the biases in representing these terms are noticeably reduced by the new approach. A series of targeted simulations in the high-resolution model further demonstrates that the approach outperforms the flux-gradient model in reproducing the stirring and dispersing effect of eddies. Our study indicates potential to upgrade the traditional flux-gradient relation for representing the eddy-induced tracer transport.

1. Introduction

Mesoscale eddies, broadly defined here as deviations from large-scale fields, profoundly impact the ocean by redistributing dynamically active and passive tracers. Eddies modify the large-scale currents by transporting momentum (Waterman et al. 2011), maintain the mean stratification in the Southern Ocean by transporting buoyancy (Marshall and Speer 2012) and potentially influence the climate by transporting heat and carbon (Jayne and Marotzke 2002; Sallée et al. 2012). Mesoscale eddies remain largely unresolved in most current climate models (Meijers 2014; Hewitt et al. 2020), and thus their large-scale effects need to be parameterized.

The most common method of parameterizing eddy effects on tracers is the flux-gradient relation, which represents eddy tracer fluxes as a product of the large-scale tracer concentration gradient and a transport coefficient tensor (Taylor 1921; Bachman and Fox-Kemper 2013). The tensor characterizes the eddy-induced diffusion and eddy-induced advection (Griffies 1998), and is considered to be a function of the flow properties such as eddy energy (Eden and Greatbatch 2008; Marshall and Adcroft 2010) and eddy mixing length scale (Prandtl 1925; Ferrari and Nikurashin 2010). The large-scale and eddy fields are commonly separated using a long-term time (Reynolds) averaging (Gent and McWilliams 1990) or a basin-scale spatial averaging (Abernathey et al. 2013; Bachman and Fox-Kemper 2013; Klocker and Abernathey 2014), which simplifies the eddy fluxes and reduces variations in the transport tensor. Other studies define eddies as deviations from spatially low-pass filtered ("coarsened") fields (e.g., Bachman et al. 2017; Aluie et al. 2018; Haigh et al. 2021a,b). This study uses the latter definition which can more accurately account for spatial and temporal variability in the large-scale fields but leads to additional complexity in the tensor.

The mesoscale eddy-induced transport, as well as the transport tensor, is mainly two-dimensional due to the joint effect of stratification and rotation. In the interior ocean, mesoscale currents move primarily along isopycnal surfaces (i.e., neutral surfaces). Thus, it is convenient to study the along-isopycnal eddy flux separately from the cross-isopycnal flux in isopycnal coordinates. For example, the widely used Gent-McWilliams parameterization scheme (Gent and McWilliams 1990, hereafter GM) approximates the time-mean isopycnal eddy transport as a combination of the Redi isopycnal diffusion (Redi 1982, hereafter Redi) and the advection by mass (layer thickness) eddy-induced velocity (EIV). In the mixed layer, the tracer distribution is influenced horizontally by eddy stirring while is homogenized vertically by turbulent mixing (Ferrari et al. 2008). In this study we focus on

the lateral eddy transport defined as horizontal in the mixed layer and isopycnal in the near-adiabatic interior.

In eddy-resolving numerical simulations, the transport tensor can be diagnosed from eddy fluxes (Abernathey et al. 2013; Bachman and Fox-Kemper 2013; Bachman et al. 2015). Recent studies, however, found inherent complexity of the tensor: ubiquitous negative (anti-diffusive) eigenvalues of its symmetric component and strong variability in space and time of all its elements (Bachman et al. 2015, 2020; Haigh et al. 2020, 2021a,b; Haigh and Berloff 2021; Kamenkovich et al. 2021; Sun et al. 2021). These properties cast doubts on the accuracy of using a positive, stationary and isotropic diffusivity for eddy parameterization in oceanic components of climate models (Meijers 2014), especially for the simulation of the transient tracer field.

Even more concerning is the non-uniqueness of the transport coefficient tensor. The tensor is not uniquely defined because of the non-divergent ("rotational") component in the eddy fluxes. The component does not directly affect tracer distributions, but can have a large magnitude (Marshall and Shutts 1981; Jayne and Marotzke 2002; Griesel et al. 2009; Kamenkovich et al. 2021; Sun et al. 2021) and thus imprint heavily on the tensor. The separation of rotational and divergent flux components is possible but it causes ambiguity due to the dependence on boundary conditions and its intrinsic spatial nonlocality (Fox-Kemper et al. 2003; Maddison et al. 2015). In addition, no matter whether the rotational flux component is retained or removed, the diagnosed tensor is generally not unique for a given flow as evidenced by its sensitivity to the initial tracer distributions, i.e., tracer dependence (Bachman et al. 2015, 2020; Kamenkovich et al. 2021; Sun et al. 2021). This undesirable property contradicts the fundamental assumption of the flux-gradient relation that the eddy transport coefficient is a quantity inherent to the flow.

The reported complexity and non-uniqueness raise serious concerns on interpretation and utility of the flux-gradient relation in situations where the temporal and spatial variations in the eddy-induced diffusion are important. The inherent spatio-temporal variability implies errors in the instantaneous eddy flux represented by only a time- and/or space-invariant eddy diffusivity. Since the tracer distributions in coarse-resolution models are very sensitive to the eddy diffusivity (Danabasoglu and McWilliams 1995; Gent et al. 2002; Kuhlbrodt et al. 2012; Gnanadesikan et al. 2015), the non-uniqueness can also lead to biases in simulating different tracers. However, the significance of such biases is uncertain because most previous studies only used a highly simplified

form (isotropic, time-invariant) of the diffusivity, partly due to the numerical instabilities induced by negative diffusivity (Leonard 1997; Trias et al. 2020). Overall, although the above evidence does not prove that the "classical" flux-gradient relation is wrong or inaccurate, it provides motivation for exploring modifications and extensions of the century-old formalism.

This study examines the advective and diffusive properties of lateral mesoscale eddy-induced tracer transport. Mesoscale eddies are defined by a high-pass spatial filter, resulting in several distinct components of the eddy tracer flux. We further explore the relative importance of these components for tracer distributions and the properties of corresponding transport tensors. The conclusions question the utility of the flux-gradient model and motivate us to to explore other methods of representing the eddy effects on tracers. We propose a new approach that directly models the local eddy forcing as a combination of eddy-induced diffusion and generalized eddy-induced advection, which helps to alleviate some of the deficiencies of the flux-gradient model. The skills of the two approaches in reproducing the eddy-induced stirring are further evaluated in a series of targeted tracer simulations with the full and truncated eddy terms. All simulations in this study are carried out on the fine grid, which avoids numerical errors arising from re-discretization of all terms onto a coarser grid. This study, therefore, is not a direct attempt to develop a new eddy parameterization scheme for coarse-resolution models.

This paper is organized as follows. Section 2 describes the off-line tracer model used to perform the tracer simulations. Section 3 describes the key theories and formulations of the tracer experiments, eddy forcing, and the eddy-representing approaches. Section 4 discusses properties of the lateral eddy fluxes and corresponding transport tensors. Section 5 discusses properties of the new approach. Section 6 presents results of the targeted experiments. Conclusions and discussions are offered in section 7.

2. Model

Simulations in this study were performed in an off-line tracer model, which uses previously computed velocities and layer thicknesses (Kamenkovich et al. 2017, 2021). The reference solution was obtained in a separate online simulation with the Hybrid Coordinate Ocean Model (HYCOM; Bleck 2002). The simulation (Mensa et al. 2013) covers the Gulf Stream region (28.78°N - 45.72°N, 81.44°W - 50°W) and spans over 1.5 years; only the last 365 days were used in this

study. The model uses a Mercator horizontal grid with 1/48° resolution. The vertical grid has 30 hybrid layers: isopycnal in the interior ocean, z-levels near the surface, and sigma-coordinates in the shallow coastal regions. The reference velocities and layer thicknesses were saved every 12 hours and were interpolated in time with a step of 1 hour in the off-line model. As in Kamenkovich et al. (2017), we used a purely advective version of the code, that is, without explicit vertical tracer diffusion below the surface mixed layer. Tracer concentration within the mixed layer is vertically homogenized on each time step.

The simulated currents contain a realistic Gulf Stream, fully resolved mesoscale currents, and partially resolved submesoscale currents that are mainly restricted to the surface mixed layer (Mensa et al. 2013). The main advantage of the off-line formulation is the ability to carry out computationally efficient sensitivity runs with modified (e.g., spatially filtered) advection. Kamenkovich et al. (2017) compared online and off-line simulations of idealized tracer releases with the same off-line code used here and concluded that the off-line tracer patch dispersion stays within 2 percent of the online simulations. They used daily velocities and layer thicknesses with a 1/12° spatial resolution, instead of the 12-hour fields at 1/48° resolution used here. The off-line model solves the following equation for the tracer in each model layer:

$$\frac{\partial (hc)}{\partial t} + \nabla \cdot (\mathbf{U}c) + \partial_s (wc) = A_h \nabla \cdot (h \nabla c)\,, \tag{1}$$

where c is the tracer concentration, h is the layer thickness, $\mathbf{U}$ is the lateral thickness flux ($\mathbf{u}h$) within the layer, and ∇ is the lateral gradient. A_h is the explicit lateral diffusivity that represents effects of unresolved subgrid mixing; its value is 0.02 m s^{-1} $\cdot \Delta x$ in the reference simulation, with Δx being the horizontal grid spacing. w is the volume flux through layer interfaces, which is vertical within the mixed layer and diapycnal in the interior ocean; we will refer to it as w-flux hereafter. Operator $\partial_s \mathbf{F} = \mathbf{F}_t - \mathbf{F}_b$ is the difference between fluxes through the top ($\mathbf{F}_t$) and the bottom ($\mathbf{F}_b$) of the layer. The w-flux is diagnosed from the mass conservation within each layer

$$\partial_s w = -\frac{\partial h}{\partial t} - \nabla \cdot \mathbf{U}\,. \tag{2}$$

Figure 1 shows the simulated tracer concentration field and upper-ocean speed, as well as the meridional cross-sections of temperature and layer interfaces. The tracer was initialized with a

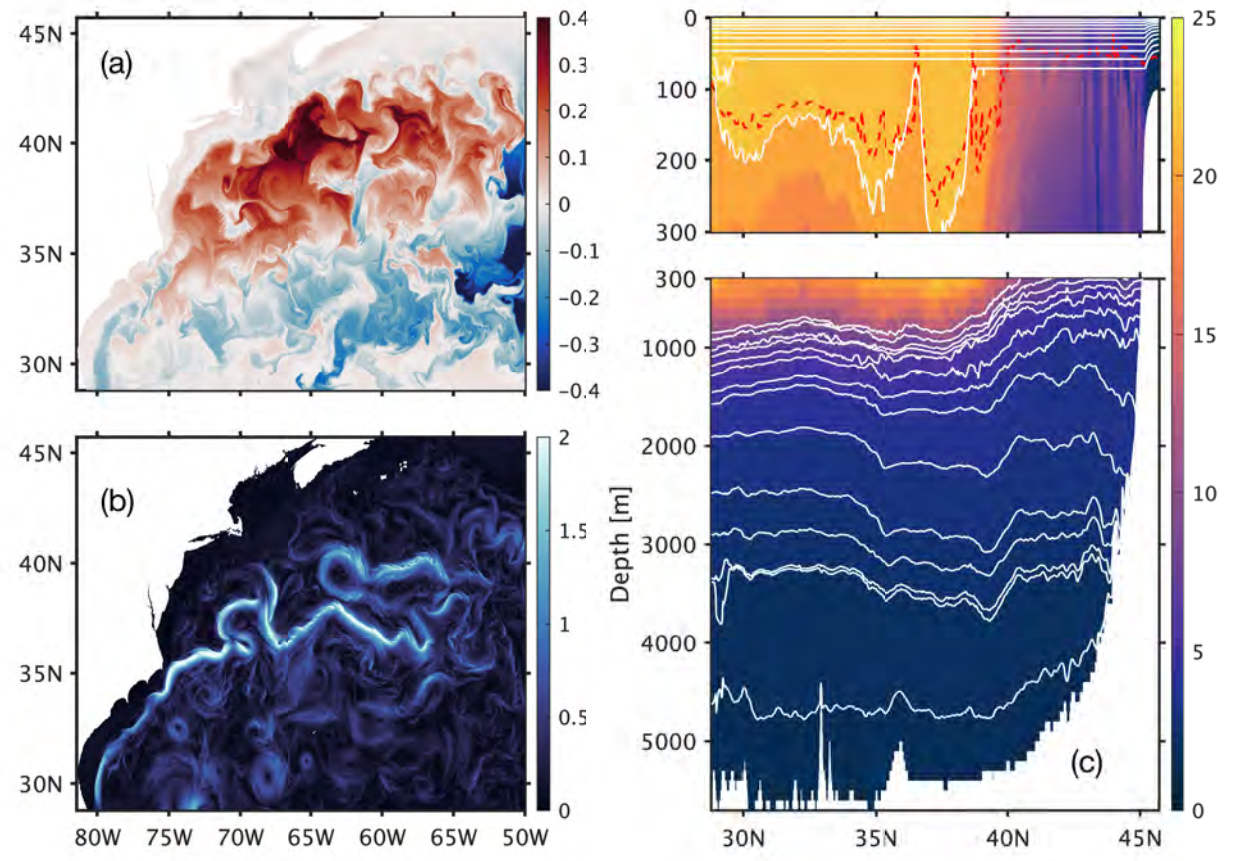


FIG. 1. Model fields. (a) The tracer concentration anomalies vertically averaged within the mixed layer after 2 months of evolution in the off-line model. The initial distribution (c_{t1}) is subtracted. (b) Flow speed from the online model averaged in the mixed layer at Day 9 of Year 9 that corresponds to (a). (c) Temperature section along 305°E at Day 9 of Year 9, overlapped with layer bottom interfaces (white solid lines) and mixed layer depth (red dashed line). Only interfaces of layers 1-10 and 16-29 are plotted in the upper and lower subplots, respectively.

sinusoidal meridional distribution ranging from unity at the southernmost points to nearly zero at the northernmost points (Appendix A). After 60 days of integration, the tracer deviates significantly from the initial distribution and forms anomalies, due to both the large-scale and eddy-induced stirring.

3. Effects of eddies and their representation

This study will analyze the eddy effects on tracers and evaluate two different approaches to representing these effects. We define the eddy effects in section 3a, discuss the flux-gradient relation in section 3b and derive the new model of eddy effects in section 3c, and discuss tracer experiments in section 3d.

a. Tracer "eddy forcing"

We first separate the lateral isopycnal thickness fluxes and layer thicknesses into the large-scale (angle brackets) and mesoscale (primes) components

$$\mathbf{U} = \langle \mathbf{U}(x,y,z,t) \rangle + \mathbf{U}'(x,y,z,t)\,, \qquad h = \langle h(x,y,z,t) \rangle + h'(x,y,z,t)\,. \tag{3}$$

The large-scale component of the w-flux, $\langle w \rangle^{\dagger}$, is calculated from the mass conservation

$$\partial_s \langle w \rangle^{\dagger} = -\frac{\partial \langle h \rangle}{\partial t} - \nabla \cdot \langle \mathbf{U} \rangle\,, \tag{4}$$

and the eddy part is $w' = w - \langle w \rangle^{\dagger}$. In this study the low-pass filter $\langle ... \rangle$ is a nominal 2° (101^2 horizontal grid points) boxcar spatial averaging. This definition is different from the common Reynolds (long-term temporal or zonal) averaging, and allows for a direct quantification of the interaction between different spatial scales (Aluie et al. 2018; Garabato et al. 2022). The superscript of $\langle w \rangle^{\dagger}$ is to distinguish it from the the directly filtered flux $\langle w \rangle$, because the spatial filter does not commute precisely with differential operators on a sphere (Aluie 2019). We verified their difference to be small and hereafter drop the superscript. Since submesoscale is partially resolved in our model, the "eddies" include both mesoscale and submesoscale anomalies. Note, however, that although the submesoscales are ubiquitous in the surface mixed layer, they are generally very weak in deeper layers (Mensa et al. 2013). We confirm that the conclusions are the same when using a medium-pass filter (Capet et al. 2008) to extract the mesoscale, and we continue to use "mesoscale" and "eddies" interchangeably.

Inserting (3) and (4) into the tracer equation (1) then leads to

$$\frac{\partial (\langle h \rangle c)}{\partial t} + \nabla \cdot (\langle \mathbf{U} \rangle c) + \partial_s (\langle w \rangle c) - A_h \nabla \cdot (\langle h \rangle \nabla c) = -\mathcal{D}_e\,, \tag{5}$$

where the tracer "eddy forcing" $\mathcal{D}_e(x, y, z, t)$ includes all the eddy effects on tracer,

$$\mathcal{D}_e = \frac{\partial(h'c)}{\partial t} + \nabla \cdot (\mathbf{U}'c) + \partial_s(w'c) - A_h \nabla \cdot (h' \nabla c), \tag{6}$$

and is the main object of this study. Note that $\mathcal{D}_e$ is a function of the full tracer concentration c including both the large-scale and eddy parts. An alternative and equivalent form of $\mathcal{D}_e$ can be derived by combining (2), (4), and (6):

$$\mathcal{D}_e = h' \frac{\partial c}{\partial t} + \mathbf{U}' \cdot \nabla c + (\partial_s(w'c) - c\partial_s w') - A_h \nabla \cdot (h' \nabla c)\,. \tag{7}$$

This study will focus only on the lateral components of the eddy forcing $\mathcal{D}_e$, namely, the lateral eddy flux divergence $\nabla \cdot (\mathbf{U}'c)$ in (6) and the lateral advection by eddy volume flux $\mathbf{U}' \cdot \nabla c$ in (7). We will use the flux-gradient relation and a new approach to model the two terms, respectively. We do not consider the eddy tendency and vertical/diapycnal terms, $\partial_t(ch') + \partial_s(w'c)$ in $\mathcal{D}_e$, because they cannot be easily expressed by the flux-gradient relation. Although in a closed domain they can be incorporated into a divergent flux (Sun et al. 2021; Haigh et al. 2021b), the appropriate boundary conditions for the corresponding Poisson problem do not exist in a general case. In addition, the w-flux is induced not only by mesoscale eddies but also by dynamically different processes, such as buoyancy mixing due to breaking of internal gravity waves. Finally, vertical terms are difficult to quantify in the mixed layer of our idealized simulations. The neglect of the eddy tendency and vertical/diapycnal terms can, however, cause biases in tracer evolution, which can be particularly significant in situations where these terms and $\nabla \cdot (\mathbf{U}'c)$ are both large and partially balance each other in $\mathcal{D}_e$. For this reason, we anticipate that truncating $\mathcal{D}_e$ to $\mathbf{U}' \cdot \nabla c$ will lead to lower biases in tracer evolution than truncating $\mathcal{D}_e$ to $\nabla \cdot (\mathbf{U}'c)$, and our tracer experiments will confirm this.

Note also that $\mathcal{D}_e$ does not include the flux of eddy tracer anomalies by the large-scale currents $\langle \mathbf{U} \rangle c'$, because this term is included into the left-hand side of (6) for consistency with the high-resolution tracer experiments in section 3d. For the sake of completeness and comparability with previous studies (e.g., Haigh et al. 2020), we will discuss properties of $\langle \mathbf{U} \rangle c'$ wherever appropriate, and confirm that our main conclusions hold true for that term as well.

b. The flux-gradient relation: transport tensor and its reduced form

The lateral eddy tracer flux $\mathbf{F}_e = \mathbf{U}'c$ in the eddy forcing can be conventionally represented by the eddy transport tensor $\mathbf{K}$ via the flux-gradient relation:

$$\mathbf{F}_e = -\langle h \rangle \mathbf{K} \nabla \langle c \rangle, \tag{8}$$

where the space- and time-dependent $\mathbf{K}$ is a 2×2 tensor that characterizes the properties of lateral eddy-induced transport. Within this framework, the eddy forcing $\mathcal{D}_e$ is approximated by $\widehat{\mathcal{D}} = -\nabla \cdot (\langle h \rangle \mathbf{K} \nabla \langle c \rangle)$ with the hat denoting a parametric model.

The tensor can be split into the symmetric and antisymmetric parts that represent physically distinct transport processes (Griffies 1998). The symmetric part $\mathbf{S}$ stands for a diffusive process affecting the domain-integrated tracer variance. It can be further modified to outline the anisotropy of diffusion:

$$\mathbf{S} = \frac{1}{2}(\mathbf{K} + \mathbf{K}^T) = \begin{pmatrix} S_{11} & S_{12} \\ S_{12} & S_{22} \end{pmatrix} = \mathbf{I}_R \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \mathbf{I}_R^T, \tag{9}$$

where $\mathbf{I}_R$ is a rotation matrix for the diffusion angle θ that defines the coordinates (eigenvectors) along which the Fickian diffusions occur with corresponding diffusivities (eigenvalues) $\lambda_{1,2}$ (Haigh et al. 2021a). The antisymmetric part $\mathbf{A}$ corresponds to an eddy-induced advection of the large-scale tracer concentration (Griffies 1998; Haigh et al. 2021b):

$$\mathbf{A} = \frac{1}{2}(\mathbf{K} - \mathbf{K}^T) = \begin{pmatrix} 0 & -A \\ A & 0 \end{pmatrix}, \tag{10}$$

where $\langle h \rangle A$ is the streamfunction for the eddy-induced advective flux $\mathbf{u}_c^* \langle h \rangle = \hat{\mathbf{z}} \times \nabla(\langle h \rangle A) = [-\partial_y(\langle h \rangle A), \partial_x(\langle h \rangle A)]$, with $\mathbf{u}_c^*$ being the corresponding tracer EIV.

In this study, the full tensor $\mathbf{K}$ will be used mainly for diagnostic analysis. In tracer simulations, we will use a reduced transport tensor $\mathbf{K}_{red}$ that combines isotropic diffusion and eddy-induced advection:

$$\mathbf{K}_{red} = \begin{pmatrix} K_{iso} & -A_{red} \\ A_{red} & K_{iso} \end{pmatrix}, \tag{11}$$

where K_{iso} is an isotropic diffusivity and A_{red} corresponds also to a tracer EIV $\mathbf{u}^*_{c-red} = \hat{\mathbf{z}} \times \nabla(\langle h \rangle A_{red})/\langle h \rangle$. The main reason of introducing $\mathbf{K}_{red}$ is that the full anisotropic tensor $\mathbf{K}$ causes numerical instability, which is not surprising given the persistent presence of negative eigenvalues $\lambda_{1,2}$ (Leonard 1997; Trias et al. 2020; Haigh and Berloff 2021). Note that the anisotropy of eddy-induced transport can still be partly captured by the advective part of $\mathbf{K}_{red}$, because $\mathbf{u}^*_{c-red}$ can always lead to tracer spreading along a well-defined direction.

There are several challenges of using the flux-gradient relation. A significant one is the presence of the rotational (non-divergent) component in eddy tracer fluxes. This component does not affect tracer distributions but is generally an order of magnitude larger than the divergent component (Jayne and Marotzke 2002; Griesel et al. 2009; Kamenkovich et al. 2021; Sun et al. 2021), leading to large values in the corresponding tensor. Thus, the biases in the parameterized eddy fluxes, inevitable when the tensor is simplified or approximated, can also be very large. This justifies the need to remove the rotational component from eddy fluxes when estimating the tensor, which is typically done by the Helmholtz decomposition. However, the definition of the rotational component is "nonlocal" and non-unique, because the decomposition depends on boundary conditions for both the divergent and rotational components, and these boundary conditions are arbitrary (Fox-Kemper et al. 2003; Maddison et al. 2015). Therefore, the dominance of rotational eddy flux leads to an inherent source of ambiguity in the estimate of the tensor. In addition, the Helmholtz decomposition is computationally expensive and is impractical if one wants to obtain the full spatio-temporal variability of the tensor. For example, the decomposition technique used in this study, as will be described later, costs about 2 CPU hours to solve for the divergent component of each instantaneous eddy flux.

The tracer dependence is another important source of uncertainty in modeling eddy fluxes. Although the origins of this uncertainty remain poorly understood, it hints at the need to revisit the flux-gradient relation. In addition, some components of the eddy forcing, such as the tendency and vertical/diapycnal terms that are not in a flux form, cannot be easily represented by the relation in a general case. These components are traditionally ignored in studies with long-term time averaging definition of eddies (e.g., Gent et al. 1995), but can be large under a more relevant non-Reynolds definition (e.g., Sun et al. 2021). These challenges of the flux-gradient relation will be further outlined by our results in section 4.

Note that the eddy fluxes in this study are unfiltered in order for preserving their divergence on the fine grid for the tracer experiments and for consistency to the eddy forcing $\mathcal{D}_e$. The results do not, however, change qualitatively, if $\langle \mathbf{F}_e \rangle$ is considered instead.

c. The generalized-advective-diffusive approach

Besides the above issues of the flux-gradient relation, there are two more fundamental properties of the eddy-induced stirring that motivate us to seek a new approach. Firstly, some components of the eddy flux cannot be expected to be successfully modeled by down-gradient diffusion. For example, the mixing length theory (Taylor 1921; Prandtl 1925), on which the flux-gradient model is based, relates tracer anomalies to the mean tracer gradient and water parcel excursion, $c' \sim -l'_j \partial_j \langle c \rangle$, which yields the definition of eddy transport tensor via $u'_i c' \sim -u'_i l'_j \partial_j \langle c \rangle$. Here the eddy velocity u'_i and excursion l'_j are assumed to be correlated, and indices i and j denote spatial coordinates. Flux $\langle u_i \rangle c'$ can also be expressed similarly as $\langle u_i \rangle c' \sim -\langle u_i \rangle l'_j \partial_j \langle c \rangle$. However, the eddy stirring of mean tracer contour, $\mathbf{U}' \langle c \rangle$, is not consistent with this framework, because the theory only gives the relation between c' and $\partial_j \langle c \rangle$, not between $\langle c \rangle$ and $\partial_j \langle c \rangle$. This flux term is clearly "advective" in nature, but cannot be represented by the advection of the antisymmetric tensor because of the divergent $\mathbf{U}'$. Therefore, a different "eddy-induced advection" is needed.

Secondly, it is the eddy forcing (e.g., flux divergence) that directly appears in the tracer budget and determines tracer evolution. The flux-gradient relation, which models the eddy flux, involves a non-local, ill-defined problem due to the rotational eddy flux component. This nonlocal property not only causes uncertainty in determining the tensor, but is also intuitively suspicious: why should mesoscale fluxes in the ocean interior depend on the remote boundary conditions? These considerations argue for modeling the tracer eddy forcing directly. George et al. (2021) attempted a similar idea by modeling eddy heat flux divergence from sea surface height anomalies with a data-driven method.

Here, we formulate a new approach with generalized eddy-induced advection (GEIA). The new model is motivated by the aforementioned properties of the flux-gradient relation and our own results, as is further discussed in section 5. We start from modifying the flux-gradient relation in two major ways. We choose to model the eddy forcing instead of the eddy flux, thus avoiding the discussed issues associated with calculating the rotational flux component. Then, to deal with

the part of the eddy forcing that cannot be represented by the flux-gradient relation, we add a new advection term, $\mathbf{U}^{\chi} \cdot \nabla\langle c \rangle$:

$$\widehat{\mathcal{D}} = -\nabla \cdot (\langle h \rangle \mathbf{K} \nabla \langle c \rangle) + \mathbf{U}^{\chi} \cdot \nabla \langle c \rangle \,, \tag{12}$$

where vector $\mathbf{U}^{\chi}$ is a free parameter, which is divergent and independent of $\mathbf{A}$. This term is a key difference from the flux-gradient relation in which the eddy-induced advective fluxes are non-divergent and determined uniquely by $\mathbf{A}$.

The generic, divergent form of $\mathbf{U}^{\chi}$ also makes the formulation purely local. To see that, we expand the first term on the right hand side of (12) and collect all advective terms into a generalized eddy-induced advective (GEIA) flux $\boldsymbol{\chi}$:

$$\widehat{\mathcal{D}} = -\langle h \rangle \{ S_{11} \partial_x \partial_x \langle c \rangle + 2 S_{12} \partial_x \partial_y \langle c \rangle + S_{22} \partial_y \partial_y \langle c \rangle \} + \boldsymbol{\chi} \cdot \nabla \langle c \rangle \,, \tag{13}$$

where the detailed derivation is provided in Appendix B. The GEIA flux $\boldsymbol{\chi}$ is treated as an independent parameter that needs to be determined from tracer distributions, because it includes the free parameter $\mathbf{U}^{\chi}$. This new formulation is local because all parameters are outside of the spatial gradient, and thus can be diagnosed directly from the local eddy forcing. It eliminates the ambiguity associated with solving for the divergent eddy flux component.

Given the large number of parameters needed in (13) and to provide a direct comparison to the reduced tensor $\mathbf{K}_{red}$, this study will consider a simplified form of (13) with isotropic diffusion and GEIA. As discussed previously, the simulations with a full diffusive tensor $\mathbf{S}$ suffer from numerical instabilities, which further justifies our choice on isotropic diffusivity here. Specifically, we set $S_{11} = S_{22} = \kappa$, $S_{12} = 0$ in (13) and get:

$$\widehat{\mathcal{D}} = -\langle h \rangle \kappa \nabla^2 \langle c \rangle + \boldsymbol{\chi} \cdot \nabla \langle c \rangle \,, \tag{14}$$

where $\kappa(x, y, z, t)$ [m^2 s^{-1}] is an isotropic diffusivity and $\boldsymbol{\chi}(x, y, z, t)$ [m^2 s^{-1}] incorporates all eddy-induced advective terms: $\boldsymbol{\chi} = -\nabla (\langle h \rangle \kappa) + \mathbf{u}_c^* \langle h \rangle + \mathbf{U}^{\chi}$. This new model can still lead to anisotropic transport of tracers because of the directional GEIA flux $\boldsymbol{\chi}$. Thus, some anisotropic properties of eddy-induced transport, such as the eddy-induced filamentation (Kamenkovich et al. 2021), can be

potentially captured, although a fully anisotropic formulation (13) would still be needed for better accuracy in future studies.

Equations (13) and (14) are counterparts of the full tensor $\mathbf{K}$ and the reduced tensor $\mathbf{K}_{red}$ in the flux-gradient approach, respectively. In what follows, we will explore applications of (14) and $\mathbf{K}_{red}$ to the tracer model and demonstrate advantages of the new approach.

Note that modeling the eddy forcing indicates reduction of the information available when diagnosing the parameters, e.g., direction and magnitude of eddy tracer fluxes will be missing. In particular, the regional eddy tracer transports can not be easily inferred from the approach. Our approach, however, is still justified because the eddy fluxes are contaminated by ill-defined rotational components.

d. Tracer experiments

The sensitivity experiments solve for tracer $\tilde{c}$ advected by the large-scale volume fluxes and forced by forcing $\mathcal{D}$:

$$\frac{\partial(\langle h\rangle\tilde{c})}{\partial t} + \nabla\cdot(\langle\mathbf{U}\rangle\tilde{c}) + \partial_s(\langle w\rangle\tilde{c}) - A_h\nabla\cdot(\langle h\rangle\nabla\tilde{c}) = -\mathcal{D}\,, \tag{15}$$

where all fields are defined on the high-resolution grid and $\tilde{c}$ will contain both large-scale and eddy parts. The forcing $\mathcal{D}$ will be either a truncated form of the full forcing $\mathcal{D}_e$ or its parametric representation $\widehat{\mathcal{D}}$ with (14) or (8) with $\mathbf{K}_{red}$. All the tracer experiments and corresponding $\mathcal{D}$ are listed in Table 1 and will be discussed in detail in section 6.

The benchmark for the solution $\tilde{c}$ is the full tracer concentration c of the reference simulation (1). It is obvious that $\tilde{c}$ will not equal c if $\mathcal{D} \neq \mathcal{D}_e$. Thus, the difference between $\tilde{c}$ and c measures the ability of the corresponding $\mathcal{D}$ to reproduce the eddy-induced stirring. Although theoretically one can aim at reproducing only the large-scale component of the tracer $\tilde{c} = \langle c\rangle$ by using $\mathcal{D} = \partial_t(h'c + \langle h\rangle c') + \nabla\cdot(\mathbf{U}'c + \langle\mathbf{U}\rangle c')$ (the w-flux and sub-grid mixing terms are ignored for simplicity) in (15), we do not pursue this objective for the following reason. Our experiments are carried out in a high-resolution model, which means that generation of small-scale tracer anomalies is inevitable. For example, the solution will have a non-trivial eddy component $\tilde{c}'$ even in the case of $\mathcal{D} = 0$, as will be seen in section 6. The anomalies $\tilde{c}'$ will in turn generate an additional eddy flux through large-scale currents, $\langle\mathbf{U}\rangle\tilde{c}'$, which will result in a solution $\tilde{c}$ diverting from $\langle c\rangle$. An

TABLE 1. Description of numerical experiments in the high-resolution off-line tracer model. Different forms of eddy forcing $\mathcal{D}$ are applied in (15). Note that $\tilde{c}$ is the "run-time" tracer solution, whereas c is the tracer diagnosed from the reference run.

Experiment	Description	Formulation
MEAN	Tracer transported by zero eddy flows (mean flows only)	$\mathcal{D} = 0$
EXP-ADV	By lateral eddy tracer advection	$\mathcal{D} = \mathbf{U}' \cdot \nabla \tilde{c}$
EXP-$\mathbf{K}_{red}$	By lateral eddy tracer flux divergence where the flux is represented by the reduced transport tensor	$\mathcal{D} = -\nabla \cdot (\langle h \rangle \mathbf{K}_{red} \nabla \langle \tilde{c} \rangle)$, where $\mathbf{K}_{red}$ is pre-estimated from $\mathbf{U}'c = -\langle h \rangle \mathbf{K}_{red} \nabla \langle c \rangle$
EXP-$\kappa\boldsymbol{\chi}$	By lateral eddy tracer advection represented by the generalized-advective-diffusive approach	$\mathcal{D} = -\langle h \rangle \kappa \nabla^2 \langle \tilde{c} \rangle + \boldsymbol{\chi} \cdot \nabla \langle \tilde{c} \rangle$, where κ and $\boldsymbol{\chi}$ are pre-estimated from $\mathbf{U}' \cdot \nabla c = -\langle h \rangle \kappa \nabla^2 \langle c \rangle + \boldsymbol{\chi} \cdot \nabla \langle c \rangle$
FULL	By full (mean plus eddy) flows	$\mathcal{D} = \partial_t (\tilde{c} h') + \nabla \cdot (\mathbf{U}' \tilde{c}) + \partial_s (w' \tilde{c}) - A_h \nabla \cdot (h' \nabla \tilde{c})$

alternative way is to solve for $\tilde{c}$ in a coarse-resolution offline model (Mana and Zanna 2014), but this task presents its own challenges in, for example, extrapolating and re-discretizing all physical fields onto a coarser grid (Patching 2022), and is beyond the scope of this study.

The eddy effects on tracer evolution in isopycnal coordinates have been studied by Gent et al. (1995), who parameterized the layer-thickness EIV $\overline{\mathbf{u}'h'}/\overline{h}$ using a long-term time (Reynolds) averaging $\overline{(...)}$. Our study is different in two aspects: i) we focus on the eddy tracer flux $\mathbf{U}'c$ instead of layer-thickness EIV in $\langle \mathbf{U} \rangle$, and ii) we use a more general scale-based filtering $\langle ... \rangle$. To clarify the differences and similarities between two approaches, we rewrite the tracer equation in the advective form by substituting (4) into (5), and ignore the vertical and small-scale mixing terms:

$$\frac{\partial c}{\partial t} + \frac{\langle \mathbf{U} \rangle}{\langle h \rangle} \cdot \nabla c = -\frac{\mathcal{D}_e}{\langle h \rangle}, \tag{16}$$

where the second term $\langle \mathbf{U} \rangle / \langle h \rangle = \langle \mathbf{u} \rangle + \mathbf{u}^*$ and the layer thickness EIV $\mathbf{u}^* = (\langle \mathbf{U} \rangle - \langle \mathbf{u} \rangle \langle h \rangle) / \langle h \rangle$. In this study, we assume that the large-scale layer-thickness flux $\langle \mathbf{U} \rangle$ is known and focus on the eddy forcing $\mathcal{D}_e$. The equation is the same as the equation (5) of Gent et al. (1995) with $\mathbf{u}^* = \overline{\mathbf{u}'h'}/\overline{h}$, $\mathcal{D}_e = \nabla \cdot \overline{\mathbf{U}'c'}$, and $\langle ... \rangle$ taken to be $\overline{(...)}$. In most modern coarse-resolution ocean models $\mathbf{u}^*$ is parameterized by the GM closure and $\mathcal{D}_e$ is parameterized by an isotropic Redi diffusion. However, Haigh et al. (2021b) showed that in a quasi-geostrophic model, $\mathbf{u}^*$ is significantly smaller than the tracer EIV $\mathbf{u}_c^*$ calculated from the advective tensor $\mathbf{A}$, indicating the importance of the advective part in $\mathcal{D}_e$. Our analysis will arrive at a similar conclusion and further demonstrate the need of

the new eddy-induced advection in our approach. As is explained in section 3c, the GEIA flux $\boldsymbol{\chi}$ includes $\mathbf{u}_c^*$, and thus $\boldsymbol{\chi} + \mathbf{u}^* \langle h \rangle$ describes the total advective effect of eddies on tracer.

Our $\boldsymbol{\chi}$ has some similarities with the "residual velocity" proposed by Pratt et al. (2016): they both measure the advective effect of eddies and both are able to contain diabatic terms. Nevertheless, we stress that they correspond to different processes: $\boldsymbol{\chi}$ describes the eddy forcing whereas the residual velocity is derived directly as the residual of the large-scale and eddy tracer flux.

Due to the tracer fluxes through the open boundaries of our domain, the global tracer mass, $\mathcal{M}_c$, is not conserved. Instead, $\mathcal{M}_c$ becomes part of the solution and will depend partly on the domain volume integral of the parametric model $\widehat{\mathcal{D}}(x, y, z, t)$. This integral determines the eddy-induced change in the global tracer mass, $\Delta\mathcal{M}_{c,eddy}$. For simplicity and a fair comparison of the different models $\widehat{\mathcal{D}}$, we force $\Delta\mathcal{M}_{c,eddy}$ to zero in the two eddy-representing experiments with (15). In both cases, only the large-scale flow is allowed to take tracers into or out of the domain. In the flux-gradient case, $\widehat{\mathcal{D}}$ is in a form of flux divergence, and thus the conservation of $\mathcal{M}_{c,eddy}$ can be guaranteed by requiring the eddy fluxes across all boundaries to be zero. We do this at both solid and open boundaries, because the eddy fluxes at open boundaries are generally unknown. In the case of the new approach, however, the conservation of $\mathcal{M}_{c,eddy}$ has to be enforced by different means because of the advection term $\mathbf{U}^{\chi} \cdot \nabla \langle c \rangle$. We will correct the eddy forcing (14) by subtracting its domain average at each grid point, to ensure $\Delta\mathcal{M}_{c,eddy} = 0$ when applying this approach in the tracer experiment. We confirm that the volume averaged forcing are less than 1% smaller than the forcing, so the effect of the correction on tracer distributions is negligible.

4. Properties of the lateral eddy transport

a. Calculation of the transport coefficient tensor

We calculate the full transport coefficient tensor $\mathbf{K}$ by inverting the flux-gradient relation (8) using two methods. One way is to obtain an exact solution (Haigh et al. 2020). Two independent tracers are needed in this case since $\mathbf{K}$ has four unknowns. As discussed by Kamenkovich et al. (2021) and Sun et al. (2021), the resulting tensor is different for each tracer pair. An alternative way is to use multiple tracers to form an overdetermined problem (Bachman et al. 2015), where the resulting tensor is an approximation that minimizes the mismatch between the reconstructed and original eddy fluxes in the tracer ensemble. Bachman et al. (2020) showed that this method leads

to an "optimal" solution of $\mathbf{K}$, judging by errors in reconstructing the eddy fluxes of temperature and salinity. Similarly, we will here use errors in reconstructing fluxes of tracers withheld from the inversion method.

To prepare the fields for calculating the tensors, 10 tracers were initialized with vertically uniform but horizontally different distributions (Appendix A). The tracer model (1) was then integrated over four consecutive 110-day segments with the first 20 days of each segment overlapped by the end of previous one, providing a one-year evolution for each tracer. Because of the indeterminate boundary conditions for the open boundaries, we use a minimization technique with Tikhonov regularization to remove the rotational eddy flux component (Li et al. 2006; Kamenkovich et al. 2021; Kamenkovich and Garraffo 2022). The method is able to obtain a decomposition without the need of explicitly specifying boundary conditions.

The reduced transport coefficient tensor $\mathbf{K}_{red}$ is also diagnosed from (8) using the same two inversion methods. The problem has only two unknowns so only one tracer is required for an exact solution.

b. Properties of the transport tensors

The complexity of the eddy transport tensor has been shown by several earlier studies (Bachman et al. 2020; Haigh et al. 2020; Kamenkovich et al. 2021; Sun et al. 2021). Below we briefly review these properties in our model.

Figure 2 shows the eigenvalues $\lambda_{1,2}$ of $\mathbf{K}$ estimated from the full eddy flux $\mathbf{F}_e$ and from its divergent component, overlapped with the diffusion angle θ. In agreement with previous studies (Bachman et al. 2020; Haigh et al. 2020; Kamenkovich et al. 2021; Sun et al. 2021), both $\lambda_{1,2}$ and θ exhibit remarkable spatial complexity. The magnitude and spatial structure of $\lambda_{1,2}$ change significantly when the rotational component is removed, and the magnitudes are over 10 times less than those estimated from the full eddy flux. The inhomogeneity of $\lambda_{1,2}$ and the root mean square (rms) values are reported in Table 2a. The inhomogeneity is quantified by the spatial standard deviation of the absolute value. The eigenvalues are mostly of opposite sign ("polarity" of the tensor), indicating that the eddy-induced diffusion is predominantly a filamentation process and thus highly anisotropic (Haigh et al. 2021a; Kamenkovich et al. 2021). For the advective part (Table 2b), the rms tracer EIV $\mathbf{u}_c^*$ is found to be one order of magnitude larger than the rms

TABLE 2. Inhomogeneity and root mean square (rms) values of (a) the diffusive ($\lambda_{1,2}$, K_{iso}, κ) and (b) the advective ($\mathbf{u}^*_c$, $\mathbf{u}^*_{c-red}$, $\boldsymbol{\chi}/\langle h\rangle$) parameters in different parametric models estimated from different eddy terms. Inhomogeneity is defined as the spatial standard deviation of absolute values. The statistics of λ_1 and λ_2 are similar and are averaged. The statistics of the fluxes are calculated from their norms. Parameters are diagnosed at day 100, layer 24, using tracers c_1-c_5. Note that $\mathbf{F}_1 = \mathbf{U}'\langle c\rangle$, $\mathbf{F}_2 = \mathbf{U}'c'$ and $\mathbf{F}_e = \mathbf{F}_1 + \mathbf{F}_2$, with the rotational components removed. For reference the inhomogeneity and rms values of $\langle\mathbf{U}\rangle/\langle h\rangle$ are 0.014 and 0.023 m s^{-1}, respectively.

(a) Diffusive parameters

	$\lambda_{1,2}$			K_{iso}			κ		
Eddy terms	$\mathbf{F}_e$	$\mathbf{F}_1$	$\mathbf{F}_2$	$\mathbf{F}_e$	$\mathbf{F}_1$	$\mathbf{F}_2$	$\mathbf{U}'\cdot\nabla c$	$\mathbf{U}'\cdot\nabla\langle c\rangle$	$\mathbf{U}'\cdot\nabla c'$
Inhomogeneity [m^2 s^{-1}]	3830	3760	466	1310	1270	236	1120	0.49	1120
rms [m^2 s^{-1}]	4230	4140	525	1510	1470	271	1220	0.56	1220

(b) Advective parameters

	$\mathbf{u}^*_c$			$\mathbf{u}^*_{c-red}$			$\boldsymbol{\chi}/\langle h\rangle$		
Eddy terms	$\mathbf{F}_e$	$\mathbf{F}_1$	$\mathbf{F}_2$	$\mathbf{F}_e$	$\mathbf{F}_1$	$\mathbf{F}_2$	$\mathbf{U}'\cdot\nabla c$	$\mathbf{U}'\cdot\nabla\langle c\rangle$	$\mathbf{U}'\cdot\nabla c'$
Inhomogeneity [m s^{-1}]	0.15	0.14	0.016	0.07	0.07	0.01	0.034	0.019	0.031
rms [m s^{-1}]	0.19	0.19	0.021	0.087	0.087	0.012	0.041	0.025	0.035

mean advective velocity $\langle\mathbf{U}\rangle/\langle h\rangle$ in (16), which results from the importance of a lateral eddy flux divergence discussed later.

The anisotropy of eddy-induced diffusion in the mid-latitude oceanic flows has been well known from Lagrangian studies (Sallée et al. 2008; Klocker et al. 2012a; Rypina et al. 2012; Kamenkovich et al. 2015). However, the Lagrangian diffusivity (tensor) is conceptually different from the locally defined $\mathbf{K}$. Although some studies (e.g., Riha and Eden 2011; Abernathey et al. 2013) found similarities between the two tensors, the asymptotically defined Lagrangian diffusivity is non-local in space and time, and it, therefore, cannot quantify the transient and local eddy effects. For example, the polarity of the Lagrangian diffusivity is particularly hard to capture, because the corresponding filamentation cannot proceed for long time. In addition, the tracer EIV $\mathbf{u}^*_c$ cannot be estimated by Lagrangian diffusivity.

Figure 3 shows the two elements of the reduced tensor $\mathbf{K}_{red}$. The inhomogeneity and rms statistics of K_{iso} and $\mathbf{u}^*_{c-red}$ are roughly one-third of those for $\mathbf{K}$ (Table 2). The reduced magnitudes

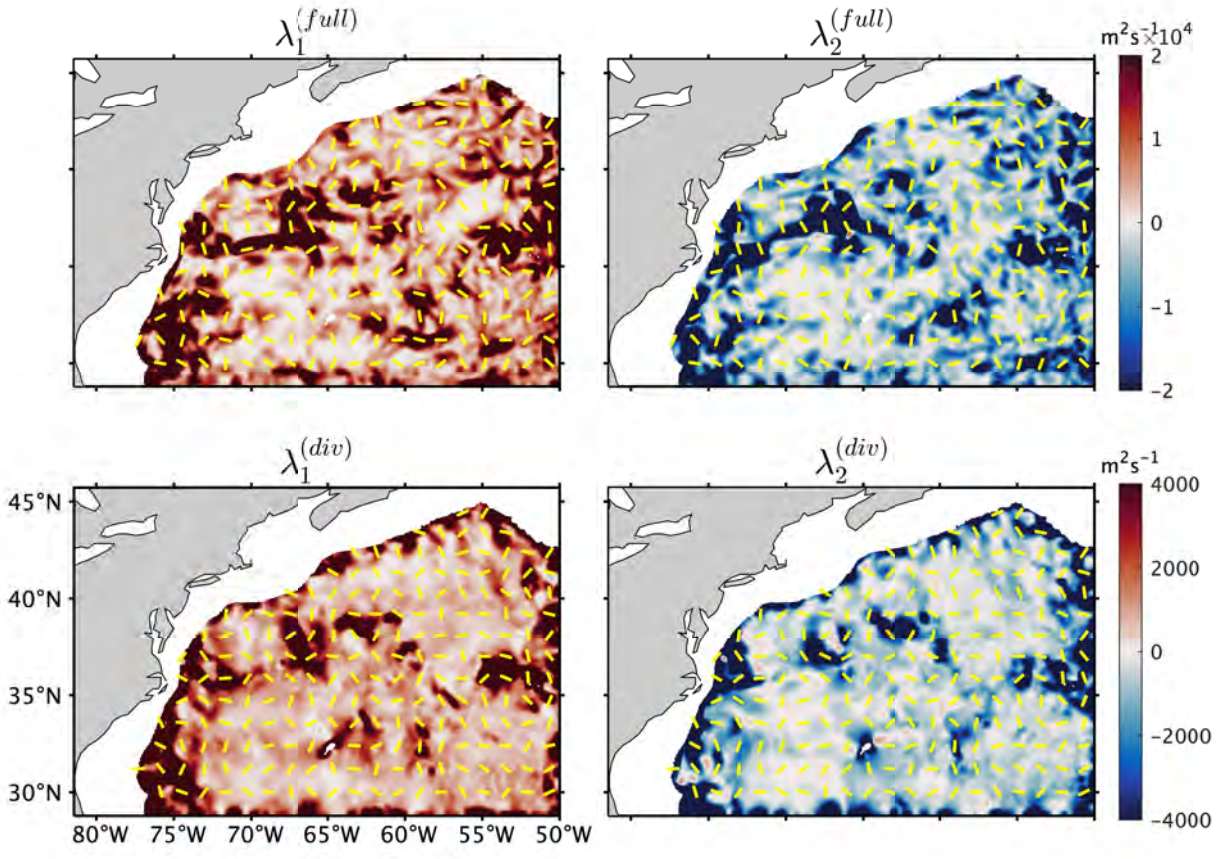


FIG. 2. Eigenvalues of **S** calculated from the total lateral eddy flux $\mathbf{F}_e$ (top) and its divergent flux component (bottom), overlapped with the direction of the major axes (direction of the maximal eigenvalue described by the angle θ) in yellow bars. The tensor is averaged over days 71-80 in layer 24 (~1500 m depth on average) and is overdetermined from five tracers (c_1-c_5). Note that the values in the top and bottom rows are an order of magnitude different. Parameters are smoothed by a 0.4°×0.4° boxcar filter for presentation.

are possibly due to the truncated anisotropic diffusion. Negative values are still common in K_{iso} (~40% of the domain), indicating transient anti-diffusive processes.

The transport tensor is also known to depend on the initial tracer distributions, although the tracers are stirred by the same flow (Bachman et al. 2015, 2020; Kamenkovich et al. 2021; Sun et al. 2021). This tracer dependence violates the main assumption that the tensor coefficient is a function of the flow only. This property is quantified here by calculating **K** ($\mathbf{K}_{red}$) from all possible tracer pairs (tracers), which results in an ensemble of the corresponding eigenvalues $\lambda_{1,2}$ (K_{iso}).

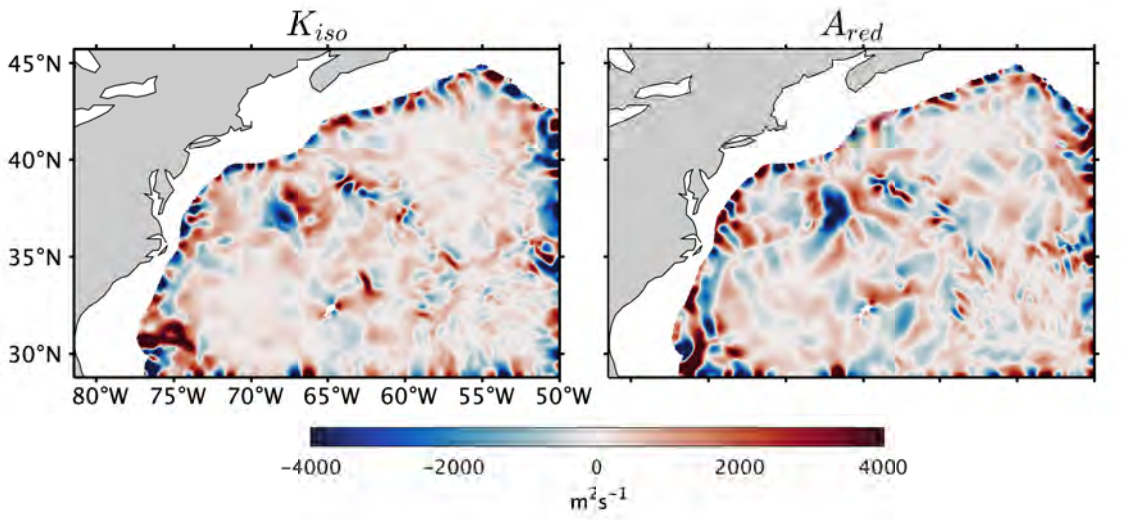


FIG. 3. Isotropic diffusivity K_{iso} and anti-symmetric part A_{red} of $\mathbf{K}_{red}$ (11) calculated from the divergent component of $\mathbf{F}_e$. The tensor is averaged over days 71-80 in layer 24, and calculated from tracers c_1-c_5. Parameters are spatially smoothed by a 0.4°×0.4° boxcar filter for presentation.

The resulting values of both $\lambda_{1,2}$ and K_{iso} change by more than 300% over the entire ensemble. This result questions the utility of the flux-gradient relation and the diffusion model.

c. Eddy flux components and their importance

The lateral eddy tracer flux $\mathbf{F}_e = \mathbf{U}'c$ is comprised of two components with distinct physical interpretations: $\mathbf{U}'\langle c\rangle$ and $\mathbf{U}'c'$. They represent the eddy advection of the large-scale tracer contours and the tracer anomalies, respectively. We now discuss their relative importance and properties of the corresponding transport tensors. For completeness, the properties of the third

eddy-induced flux, the flux of tracer anomalies by the large-scale flow $\langle \mathbf{U} \rangle c'$, are also discussed here.

Figure 4 (top) compares magnitudes of the divergences of these eddy flux terms with that of $\langle \mathbf{U} \rangle \langle c \rangle$. An intriguing result is that the divergence of the eddy advection of the large-scale tracer $\mathbf{U}' \langle c \rangle$ is largest among all four. To gain further insight into this term, we split it into two components:

$$\nabla \cdot (\mathbf{U}' \langle c \rangle) = \langle c \rangle \nabla \cdot \mathbf{U}' + \mathbf{U}' \cdot \nabla \langle c \rangle \,. \tag{17}$$

Within an isopycnal layer, the first term on the right-hand side represents the effect of volume convergence, that leads to the squeezing/stretching of the layer and its diapycnal mass exchanges with the adjacent isopycnal layers. We refer to this term as the "expansion term". In the mixed layer, the term corresponds to the vertical convergence of the volume flux. In both regimes, the term is the main cause of the exaggerated importance of $\nabla \cdot (\mathbf{U}' \langle c \rangle)$, because the second term on the right-hand side of (17) is of the same order of magnitude as the other eddy advection terms (Figure 4, bottom). There are two reasons for the dominance of the expansion term $\langle c \rangle \nabla \cdot \mathbf{U}'$: (i) $\langle c \rangle$ is generally larger than c', which explains why the expansion term is larger than $c' \nabla \cdot \mathbf{U}'$; (ii) the divergence of $\mathbf{U}'$ tends to be larger than the divergence of the large-scale volume flux $\langle \mathbf{U} \rangle$, which explains why the expansion term is larger than $c' \nabla \cdot \langle \mathbf{U} \rangle$. The expansion term can also be expected to explain a large part of the tracer dependence in $\mathbf{K}$, since $\langle c \rangle \nabla \cdot \mathbf{U}'$ strongly depends on the initial tracer distribution and would also change if a constant were added to c all over the domain. Note that the expansion term is zero in quasi-geostrophic studies (e.g., Haigh et al. 2020). The dominance of the expansion term $\langle c \rangle \nabla \cdot \mathbf{U}'$ and its dependence on $\langle c \rangle$ thus complicate application of the flux-gradient relation to the lateral (isopycnal) fluxes, especially $\mathbf{U}' \langle c \rangle$.

All the eddy forcing terms in Figure 4 remain large after being spatially filtered by $\langle ... \rangle$, meaning that they not only affect the full but also the large-scale tracer distribution. Remarkably, $\mathbf{U}' \langle c \rangle$, as well as its divergence, is still the largest eddy terms after the filtering. This property results from the lack of scale separation between the large-scale and eddy fields. Note that for the Reynolds decomposition, only the eddy-eddy term, $\mathbf{U}' c'$, remains non-zero after long-term time averaging.

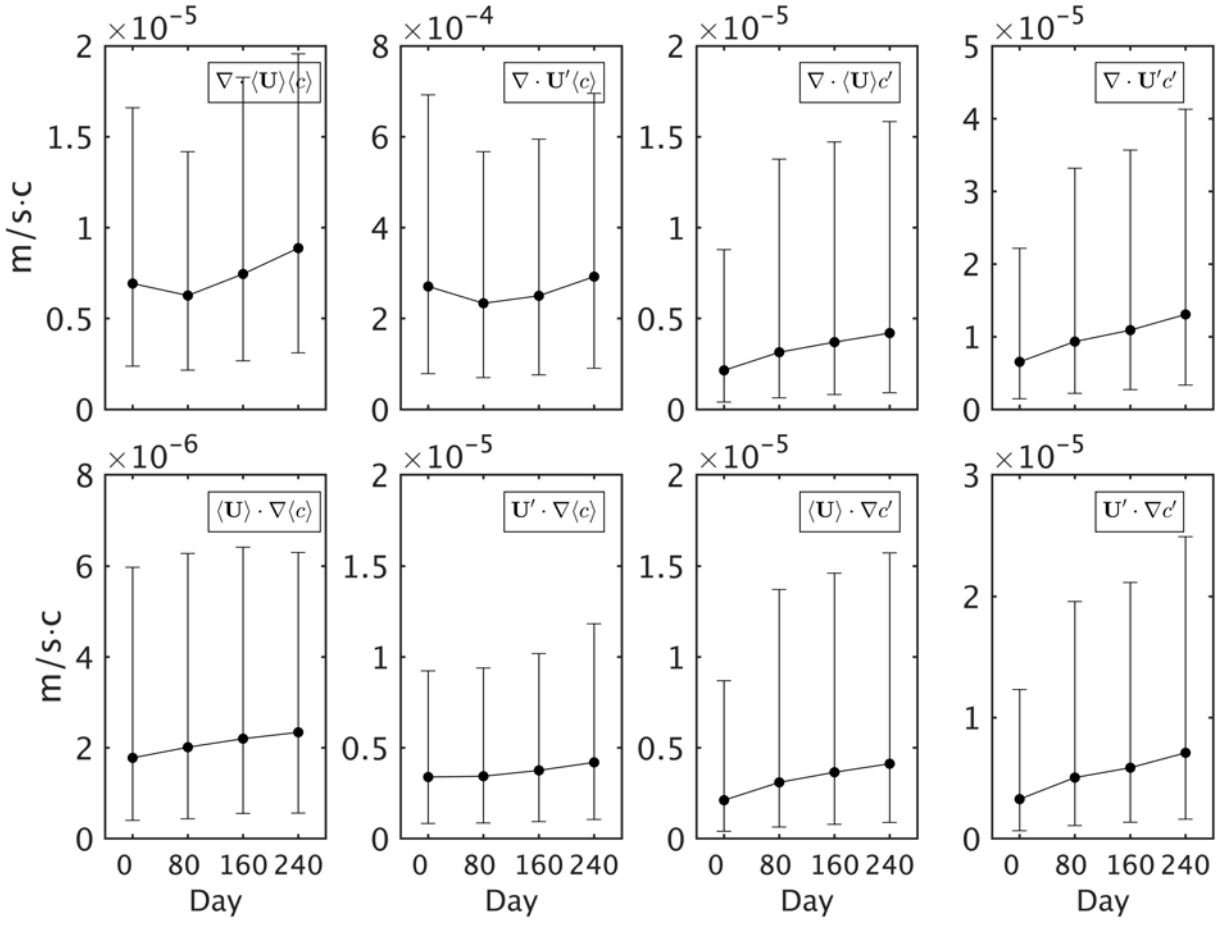


FIG. 4. Domain-median of absolute values of the divergence (top) and advection (bottom) for different lateral mean and eddy forcing terms over 8 months in layer 24. Bars denote the 20th-80th percentile range. Similar results are found in other layers. Initial distribution of the tracer is c_{t1} (Appendix A). Note that $\nabla \cdot (\mathbf{U}'\langle c \rangle)$ is an order of magnitude larger than the other forcing terms.

Because the flux-gradient relation is linear, we can further explore properties of $\mathbf{K}$ estimated from $\mathbf{U}'\langle c \rangle$ and $\mathbf{U}'c'$ separately:

$$\mathbf{F}_1 = \mathbf{U}'\langle c \rangle = -\langle h \rangle \mathbf{K}_1 \nabla \langle c \rangle, \quad \mathbf{F}_2 = \mathbf{U}'c' = -\langle h \rangle \mathbf{K}_2 \nabla \langle c \rangle \,. \tag{18}$$

The rotational component is removed from each of the flux components separately. Table 2a lists the inhomogeneity and rms magnitudes of both the eigenvalues and tracer EIVs of $\mathbf{K}_{1,2}$. Both variables are much larger for $\mathbf{K}_1$ than for $\mathbf{K}_2$, which is due to large mean values and fluctuations of

$\mathbf{F}_1$. Similarly, the reduced transport tensor $\mathbf{K}_{red}$ from $\mathbf{F}_1$ (Table 2) also has larger inhomogeneity and rms values than $\mathbf{K}_{red}$ from $\mathbf{F}_2$.

The tracer dependence of the tensors is quantified next. We first estimate an ensemble of tensors $\mathbf{K}_{1,2}$ from a set of tracer pairs, and calculate their eigenvalues and tracer EIVs. The tracer dependence is then defined as the ratio of the ensemble standard deviation to the absolute ensemble mean of these parameters. The results show that both the diffusion (Figure 5a) and advection (Figure 5b) tensor components are strongly tracer-dependent (>300% in most of the domain) for all flux components, although this dependence in $\mathbf{K}_2$ is weaker than in both $\mathbf{K}$ and $\mathbf{K}_1$. The tensors calculated from $\langle \mathbf{U} \rangle c'$ (not shown) have tracer dependence similar to $\mathbf{K}_2$. The tracer dependence in $\mathbf{K}_{red}$ is higher than $\mathbf{K}$ except for $\mathbf{F}_2$. The elevated tracer dependence in $\mathbf{F}_1$ can be explained by the expansion term being the leading source of tracer dependence. For example, if a constant is added to the tracer field, the tensor calculated by $-\mathbf{U}'\langle c \rangle \cdot [\nabla \langle c \rangle]^{-1}$ will change correspondingly, leading to unphysical dependence of the tensor on the tracer concentration. In agreement with Sun et al. (2021), we verified that tensors calculated from large-scale eddy fluxes, $\langle \mathbf{F}_e \rangle$, have very similar properties.

Because of tracer dependence, eddy fluxes reconstructed from $\mathbf{K}$ will always have biases regardless of the number of tracers used in the calculation. For $\mathbf{K}$ calculated from a large ensemble of tracers (overdetermined problem), the biases are inevitable for each of the eddy fluxes. For $\mathbf{K}$ calculated exactly from a tracer pair, the tracer dependence leads to biases for a third tracer. We use the relative error in the divergence of the eddy flux to examine the significance of these biases on tracer distribution, because it is the flux divergence that directly enters the tracer budget. We define the relative error ε that results from using a parametric model $\widehat{\mathcal{D}}$ to represent the eddy forcing $\mathcal{D}$ for a given tracer as

$$\varepsilon = \left| \frac{\mathcal{D} - \widehat{\mathcal{D}}}{\mathcal{D}} \right|, \tag{19}$$

where $\mathcal{D} = \nabla \cdot \mathbf{F}_e$ and $\widehat{\mathcal{D}} = -\nabla \cdot (\langle h \rangle \mathbf{K} \nabla \langle c \rangle)$. The relative error quantifies the biases in the reconstructed eddy forcing fields (e.g., in flux divergence) that are caused by the uncertainty (e.g., tracer dependence) in the parameters (e.g., $\mathbf{K}$) that represent the eddy forcing. It is zero only when $\mathbf{K}$ is calculated exactly from a pair of tracers that includes the given tracer.

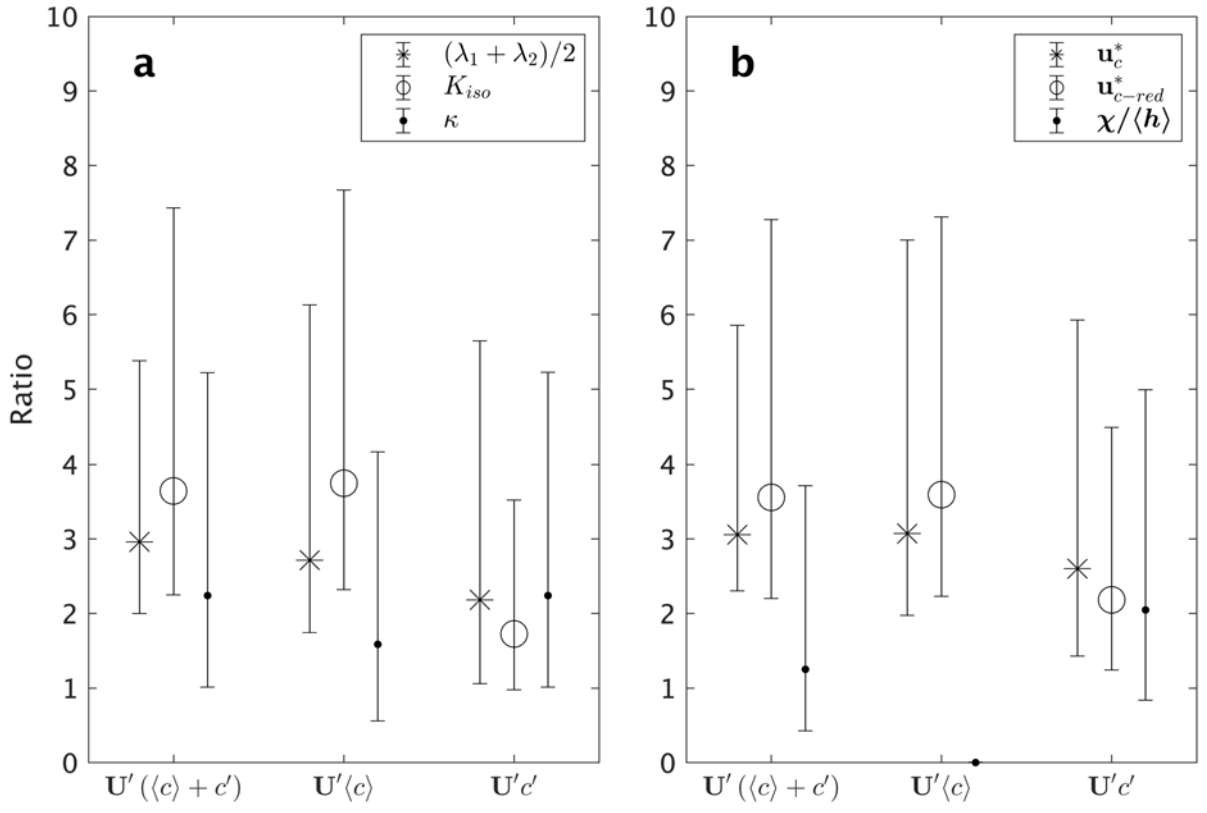


FIG. 5. Tracer dependence (ratio of the standard deviation to the absolute ensemble mean) of the (a) diffusive and (b) advective parameters of the two parametric models $\widehat{\mathcal{D}}$, estimated from different eddy forcing terms. Error bars denote the median and the 20th-80th percentile range of the ratio. (a) Diffusive parameters include the average of the eigenvalues $\lambda_{1,2}$ of $\mathbf{S}$, the isotropic diffusivity K_{iso}, and κ of the generalized-advective-diffusive model. (b) Advective parameters include the tracer EIVs associated with $\mathbf{A}$ and A_{red}, and the generalized advective flux $\boldsymbol{\chi}$ normalized by $\langle h\rangle$. All three ensembles of $\mathbf{K}_{red}$, $\mathbf{K}$, and κ and $\boldsymbol{\chi}$ contains 10 estimates, with $\mathbf{K}$ randomly chosen from all the 45 possible estimates (tracer pairs) and $\kappa\&\boldsymbol{\chi}$ randomly chosen from the 120 possible estimates (tracer triplets). For a flux, the ratios of its two horizontal components are averaged. The x-axis denotes different eddy terms from which the parameters are estimated. Note that tensors $(\mathbf{K}, \mathbf{K}_{1,2})$ are estimated from eddy fluxes $(\mathbf{F}, \mathbf{F}_{1,2})$ whereas $\kappa\&\boldsymbol{\chi}$ are estimated from eddy advections (e.g., $\mathbf{U}'\cdot\nabla c$). All parameters are diagnosed at day 100, layer 24. Other layers have similar results.

Figure 6 shows ε for $\mathbf{F}_e$ and $\mathbf{F}_{1,2}$, as a function of the number of tracers in the set $\{c_n, n = 1, ..., N\}$ used to calculate $\mathbf{K}$. The tensor is solved either exactly ($N = 2$) or by using the overdetermined method ($N > 2$). The results are shown for the tracer c_6, which is not included in the tensor calculations, but the conclusions are the same for all other tracers. The errors ε for different terms are larger than 100% in most of the domain and are not improved significantly with increasing N. This implies that the multi-tracer inversion cannot reduce the tracer dependence effectively. It is, therefore, almost certain that the tracer dependence of $\mathbf{K}$ has a significant effect on tracer transport through the errors in eddy forcing. We also see larger bias in the divergence of $\mathbf{U}'\langle c\rangle$, which is particularly problematic because it dominates lateral eddy fluxes. As discussed in section 3c, this flux term is inconsistent with the flux-gradient framework. We have now confirmed that it is one of the sources of bias in the relation.

5. Generalized-advective-diffusive approach

The results in the last section illustrate challenges of the flux-gradient approach and provide motivation for the generalized-advective-diffusive approach. In this section, we perform similar analyses of our new formulation (14), and show the improvement relative to the flux-gradient method.

a. Properties

One advantage of the new approach is that it can be used to represent any term in the eddy forcing, such as the lateral eddy tracer flux divergence $\nabla \cdot (\mathbf{U}'c)$, the lateral eddy advection $\mathbf{U}' \cdot \nabla c$, the vertical term, and eddy tendency. As mentioned in section 3a, this study focuses on the lateral eddy-induced transports only. This is for the sake of a direct and fair comparison with the flux-gradient approach, which can only be applied to the lateral eddy flux. The advection form of the lateral term $\mathbf{U}' \cdot \nabla c$ is anticipated to be more accurate than the flux divergence form $\nabla \cdot \mathbf{U}'c$. for the following reason. In section 4, we showed that a large portion of the lateral eddy flux divergence is devolved into the "expansion" term, i.e., $\langle c\rangle \nabla \cdot \mathbf{U}' \gg \mathbf{U}' \cdot \nabla \langle c\rangle$. Using the advection form will thus eliminate the need to account for this expansion term and avoids its elevated sensitivity to $\langle c\rangle$. We will later confirm this by analyzing the corresponding tracer dependence and errors in the the represented eddy forcing. We leave the consideration of the vertical/tendency terms for a future

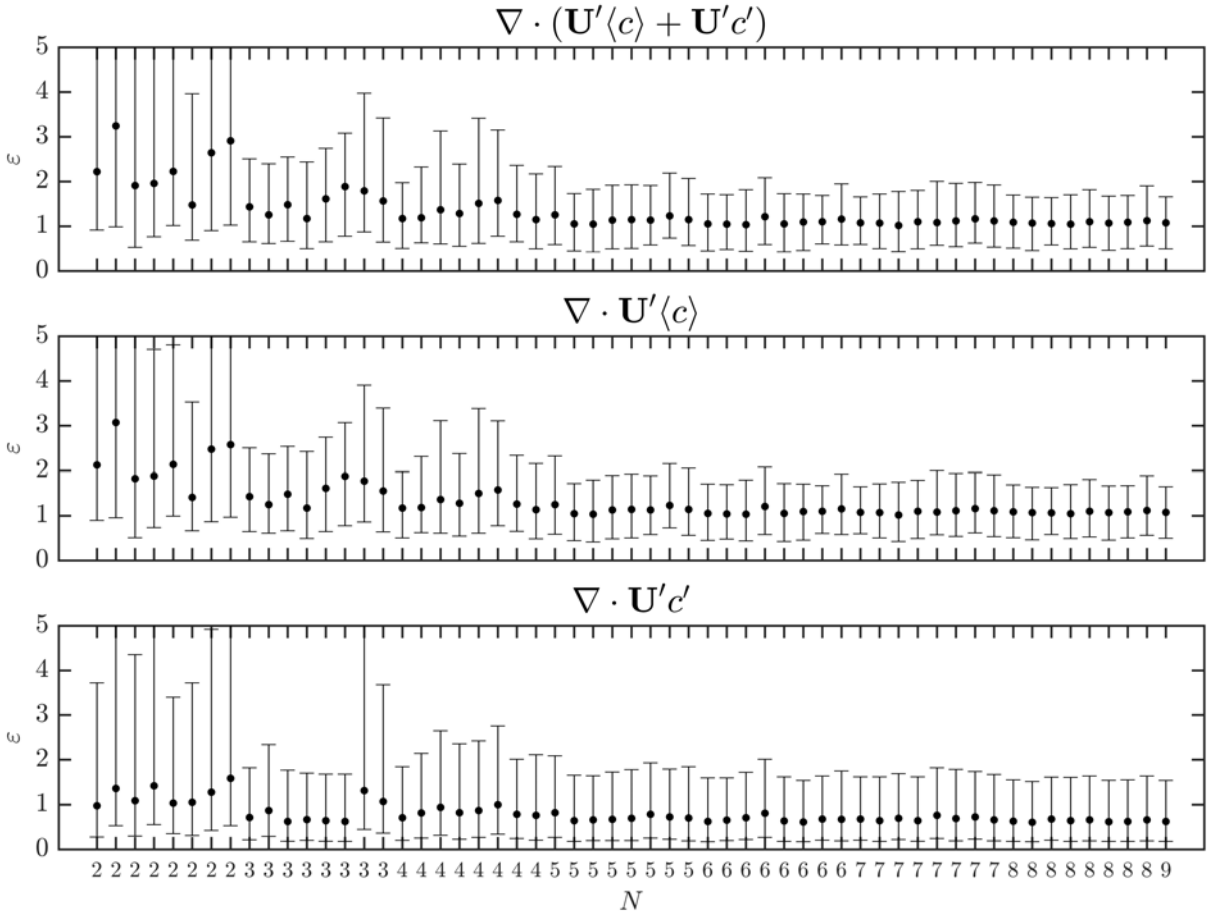


FIG. 6. Effects of the tracer dependence in the flux-gradient approach on the accuracy of the represented eddy forcing. Dots are median of the relative error ε of the eddy flux divergence at day 100, layer 24. Bars denote the 20th-80th percentile range. Different components of the eddy flux are reconstructed by $\mathbf{K}$ for tracer c_6. N is the number of tracers in the set $\{c_n\}$ ($c_6 \notin \{c_n\}$) that was used to calculate $\mathbf{K}$. Similar results were found for different times, layers and tracers.

study. We will, however, show in section 6 that the lateral eddy advection effectively captures the bulk effect of the eddy forcing.

The diffusivity κ and GEIA flux χ are calculated by inverting (14) from the lateral eddy advection using the same set of idealized tracers. We use either three independent tracers to get an exact solution or more than three tracers to get a least-squares solution of the three unknowns. Figure 7 shows snapshots of κ and χ overdetermined with five tracers. Similar to the diffusivities of $\mathbf{K}$ and $\mathbf{K}_{red}$, κ is spatially inhomogeneous and exhibits both positive and negative values. The

inhomogeneity and rms values of κ (Table 2a) have similar magnitude with K_{iso} of $\mathbf{K}_{red}$, although the spatial pattern is very different (Figure 3). This is not unexpected because the two isotropic diffusivities are formulated differently.

The new divergent GEIA flux $\boldsymbol{\chi}$ is of particular interest. Its magnitude (divided by $\langle h \rangle$) is about two times less than $\mathbf{u}_c^*$ from both $\mathbf{K}$ and $\mathbf{K}_{red}$ (Table 2b). A key question is on the importance of the new term $\mathbf{U}^{\chi}$ and whether this part is really divergent, i.e., cannot be represented by the antisymmetric tensor. Although $\mathbf{U}^{\chi}$ and $\mathbf{u}_c^* \langle h \rangle$ cannot be calculated individually in this framework, we can extract the divergent component of their sum, which equals $\boldsymbol{\chi} + \nabla (\langle h \rangle \kappa)$ by definition. The divergent component then measures the importance of $\mathbf{U}^{\chi}$ because $\mathbf{u}_c^* \langle h \rangle$ is purely rotational. Using the same technique as for the eddy tracer flux to calculate the divergent and rotational components, we found that their rms values, divided by $\langle h \rangle$, are 0.57 m s^{-1} and 0.04 m s^{-1}, respectively. This indicates that there is indeed a significant divergent component in the GEIA flux that cannot be possibly described by a streamfunction and thus by the eddy transport tensor.

We next evaluate the tracer dependence in κ and $\boldsymbol{\chi}$ using the same method as for $\mathbf{K}$ and $\mathbf{K}_{red}$. We remind the reader that this dependence contradicts the assumption that the parameters of $\widehat{\mathcal{D}}$ are functions of the flow only and increases uncertainty in estimating these parameters for practical applications. Figure 5 shows the spread of κ and $\boldsymbol{\chi}$ derived from different eddy advection terms among the tracer triplets. We observe modest reduction in the tracer dependence relative to the diffusive and advective parts of $\mathbf{K}$ and $\mathbf{K}_{red}$. The only exception is the spread of $\boldsymbol{\chi}$ for the term $\mathbf{U}' \cdot \nabla \langle c \rangle$ which is zero; this is expected because $\boldsymbol{\chi} = \mathbf{U}'$ in this case.

To examine how the tracer dependence affects the accuracy of representing the eddy forcing, we recalculated the relative error ε from (19), with $\mathcal{D} = \mathbf{U}' \cdot \nabla c$ and $\widehat{\mathcal{D}}$ given by (14). In this case, κ and $\boldsymbol{\chi}$ are calculated for a subset of tracers triplets and the error is evaluated for the rest of the tracer ensemble. Figure 8 shows ε for different eddy advection terms. We see that ε is significantly reduced compared to the flux-gradient model (Figure 6). We conclude that the new approach leads to reduced uncertainty in estimating the eddy forcing, despite the remaining sensitivity of its parameters to tracer distributions.

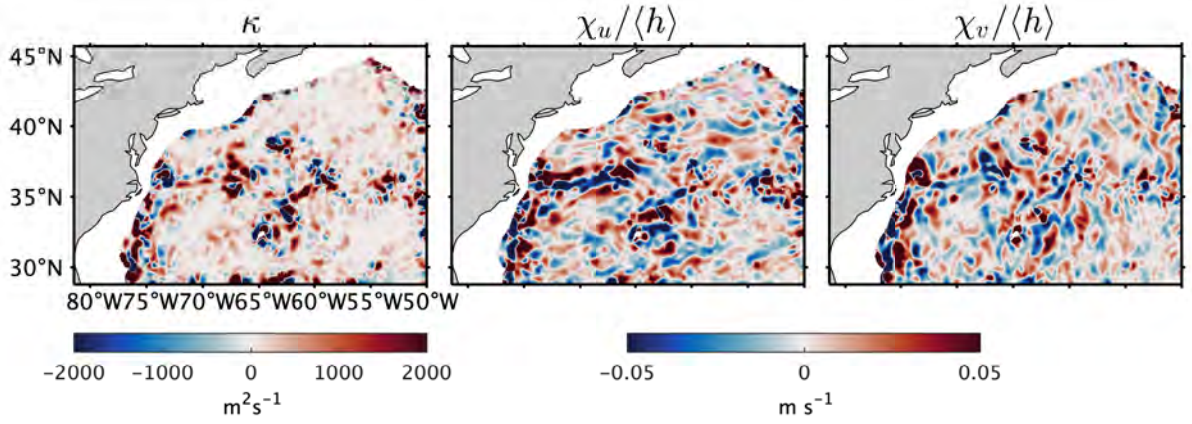
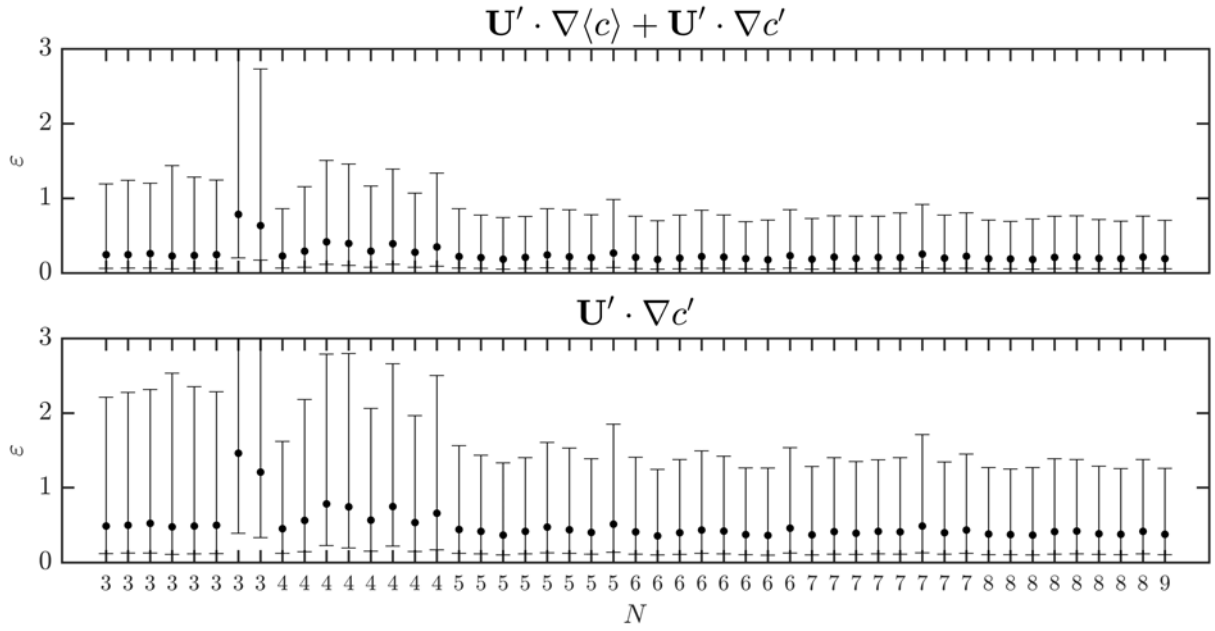


FIG. 7. Snapshots of κ and $\boldsymbol{\chi}/\langle h\rangle$ in the new approach, estimated from lateral eddy advection $\mathbf{U}' \cdot \nabla c$. They are averaged over days 71-80 in layer 24 and are overdetermined from tracers c_1-c_5. κ is on the same order of magnitude with $\lambda_{1,2}$ and K_{iso}, but has different spatial structure. Parameters are spatially smoothed by the 0.4°×0.4° boxcar filter for presentation clarity.

6. Application in the high-resolution model

This section aims to evaluate the skill of the flux-gradient relation and the generalized-advective-diffusive model in reproducing the eddy-induced stirring, as well as the effects of truncating the eddy forcing $\mathcal{D}_e$ to its lateral part. This is done by performing a series of tracer experiments (15) with various forms of eddy forcing $\mathcal{D}$, and by quantifying how close the approximate solution $\tilde{c}$ is to the reference solution. To assess the robustness of conclusions, our analysis on the tracers is carried out in two distinct parts of the domain: in the mixed layer and below the isopycnal layer 15 (~500 m depth) of the model. We remind the reader, that the tracer in the mixed is stirred



586 FIG. 8. Effects of tracer dependence on the accuracy of modeling of the eddy forcing in the generalized-
587 advective-diffusive approach. Median of the relative errors ε between the eddy advection reproduced by κ and $\boldsymbol{\chi}$
588 and the original advection term, for the tracer concentration c_6 in layer 24. Bars denote the 20th-80th percentile
589 range. The relative error for $\mathbf{U}' \cdot \nabla\langle c \rangle$ is nearly zero (not shown).

598 horizontally by a combination of mesoscale and partially-resolved submesoscale currents and is
599 homogenized vertically at each time step.

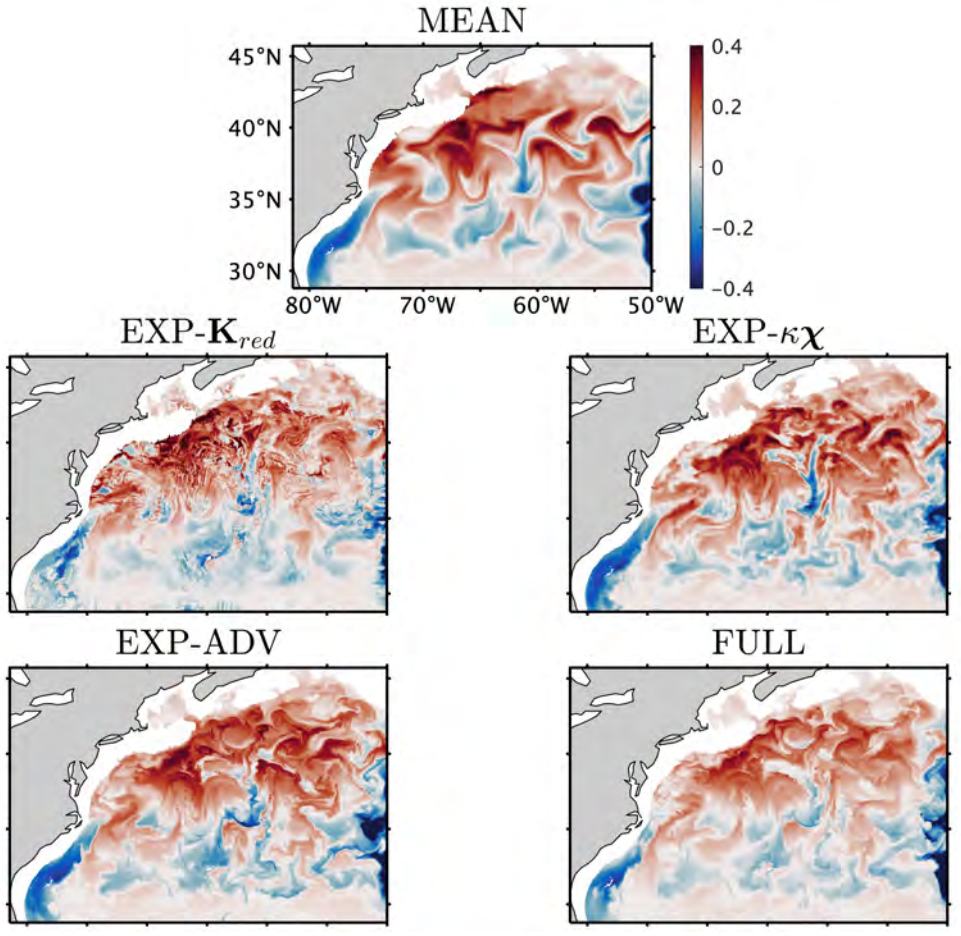
600 *a. Tracer experiments*

601 The total of five experiments are listed in Table 1. The MEAN run has $\mathcal{D} = 0$ with tracers
602 transported by the large-scale volume flux, and the FULL run is the reference (control) simulation
603 with the full flow. The EXP-ADV run illustrates the importance of the lateral eddy advection
604 ($\mathcal{D} = \mathbf{U}' \cdot \nabla c$) in the eddy forcing and serves to quantify the effect of omitting the non-lateral terms
605 in eddy forcing (7). It also represents the "best scenario" when this advection term is represented
606 accurately.

607 The other two experiments examine how well the flux-gradient approach (EXP-$\mathbf{K}_{red}$) and the new
608 approach (EXP-$\kappa\boldsymbol{\chi}$) represent the lateral eddy forcing. In both experiments, $\widehat{\mathcal{D}}(x, y, z, t)$ is applied
609 to simulate tracers that were not used to calculate the parameters of this $\widehat{\mathcal{D}}$. The tracer dependence in
610 these parameters will inevitably cause bias in $\widehat{\mathcal{D}}$, which leads to the difference between $\tilde{c}$ and c from

the FULL simulation. The difference between solutions thus mainly quantifies the impact of tracer dependence and the "goodness" of the approaches. Our previous analysis on the relative errors in eddy forcing is a complementary measure in a diagnostic way. The other source of the difference is the omitted eddy tendency and vertical/diapycnal terms, whose effect is measured by EXP-ADV. In the EXP-$\mathbf{K}_{red}$ run, we use the reduced transport tensor (11) to represent the lateral eddy flux. The new approach (14) is applied in the EXP-$\kappa\boldsymbol{\chi}$ run. Both representations are the combination of an isotropic diffusion and an eddy-induced advection. As discussed in section 3b, we did not employ the anisotropic diffusion because it is numerically unstable as a result of negative diffusivities. The comparison serves to evaluate the implications of the generalized advection. We used five tracers (c_1-c_5) to estimate $\mathbf{K}_{red}$, κ and $\boldsymbol{\chi}$, which ensures that the same amount of information is provided for both approaches. Note that these experiments are not eddy parameterizations because the parameters are diagnosed from the reference simulation, and the experiments are carried out in the same high-resolution model.

Figure 9 shows the snapshots of tracer solutions in the five experiments. Comparison of MEAN and FULL illustrates that eddies influence the tracer distribution mainly in two ways: (i) they smooth out sharp large-scale tracer concentration fronts (gradients) and the peak values on each side of a front; and (ii) they induce small-scale concentration structures all over the domain, and especially along the large-scale fronts. The tracer field in EXP-ADV and EXP-$\kappa\boldsymbol{\chi}$ is visually similar to FULL. We then quantify the difference between the solutions of FULL and the other runs by the Frobenius norm (Figure 10a) and the correlation coefficients (Figure 10b-e). The results show that the EXP-ADV solution (green line and Figure 10e) is significantly closer to the FULL than is the MEAN (black line and Figure 10b), indicating an improvement because of using the lateral eddy advection. The remaining bias results from the truncated terms $h'\partial_t c + (\partial_s(w'c) - c\partial_s w')$ in (7), and are confirmed small. This justifies our choice to model the lateral term by the new approach. In contrast, the same bias for the flux-gradient approach cannot be quantified because an additional experiment with $\mathcal{D} = \nabla\cdot(\mathbf{U}'c)$ (not shown) exhibits unrealistic tracer concentration patterns, which is a possible consequence of excessively large biases caused by the absent terms $\partial_t(ch') + \partial_s(w'c)$ in (6). The solution of EXP-$\kappa\boldsymbol{\chi}$ (red line and Figure 10d) is closer to FULL than EXP-$\mathbf{K}_{red}$ (blue line and Figure 10c), indicating a statistical improvement brought by the new approach.

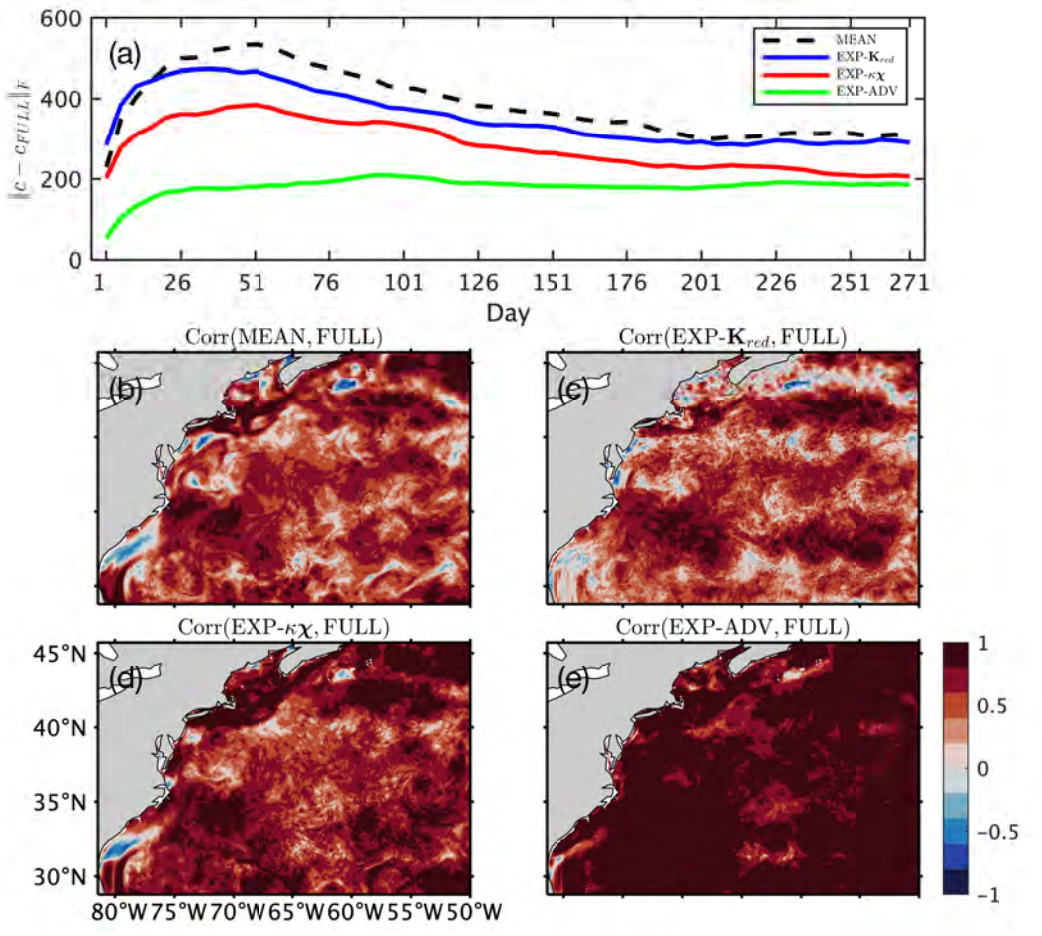


640 FIG. 9. Tracer concentration anomalies from the five experiments (Table 1) in layer 15 (~500 m depth) after
641 200 days of advection. The initial distribution (c_{t1}) is subtracted out.

647 Mesoscale eddies can have cumulative effect on the large-scale tracer distribution through stir-
648 ring and dispersion. We next consider two physically meaningful processes for the sensitivity
649 experiments: evolution of a large-scale tracer gradient and dispersion of a tracer patch along the
650 Gulf Stream path.

651 *b. Evolution of a large-scale tracer gradient*

652 The analysis in this section is focused on the efficiency of eddy stirring in reducing the meridional
653 gradients. Simulations are initialized with the tracer c_6 which has a sinusoidal structure in the
654 meridional direction and is uniform zonally and vertically. Over time, the eddy stirring reduces the
655 meridional gradients and the tracer slowly approaches a uniform distribution.



642 FIG. 10. The skill of the truncated and the approximated eddy forcings in reproducing the eddy effects. (a)
643 The Frobenius norm of the vertically-averaged tracer concentration differences between the four experiments
644 and FULL as a function of time. The Pearson correlation of tracer fields between (b) MEAN, (c) EXP-$\mathbf{K}_{red}$, (d)
645 EXP-$\kappa\boldsymbol{\chi}$ and (e) EXP-ADV and FULL, over 270 days. The initial distribution is c_{t1}. Tracers are averaged within
646 the mixed layer. Similar results are observed for tracers averaged below layer 15.

656 Figure 11 shows the relative difference of tracer fields ("tracer bias") between the FULL and three
657 sensitivity experiments, MEAN, EXP-$\mathbf{K}_{red}$ and EXP-$\kappa\boldsymbol{\chi}$, as a function of time and latitude. In both
658 the mixed layer and deeper layers, EXP-$\kappa\boldsymbol{\chi}$ is closer to FULL than EXP-$\mathbf{K}_{red}$, showing that the new
659 approach captures the eddy stirring more accurately than the flux-gradient approach. An intriguing
660 feature in the mixed-layer tracer concentrations (Figure 11, top) is the large relative difference at
661 about 42°N. We attribute this difference to the presence of a transport barrier at the Gulf Stream
662 core (Rypina et al. 2011) in the FULL run. Since velocities at this core in MEAN is smeared out

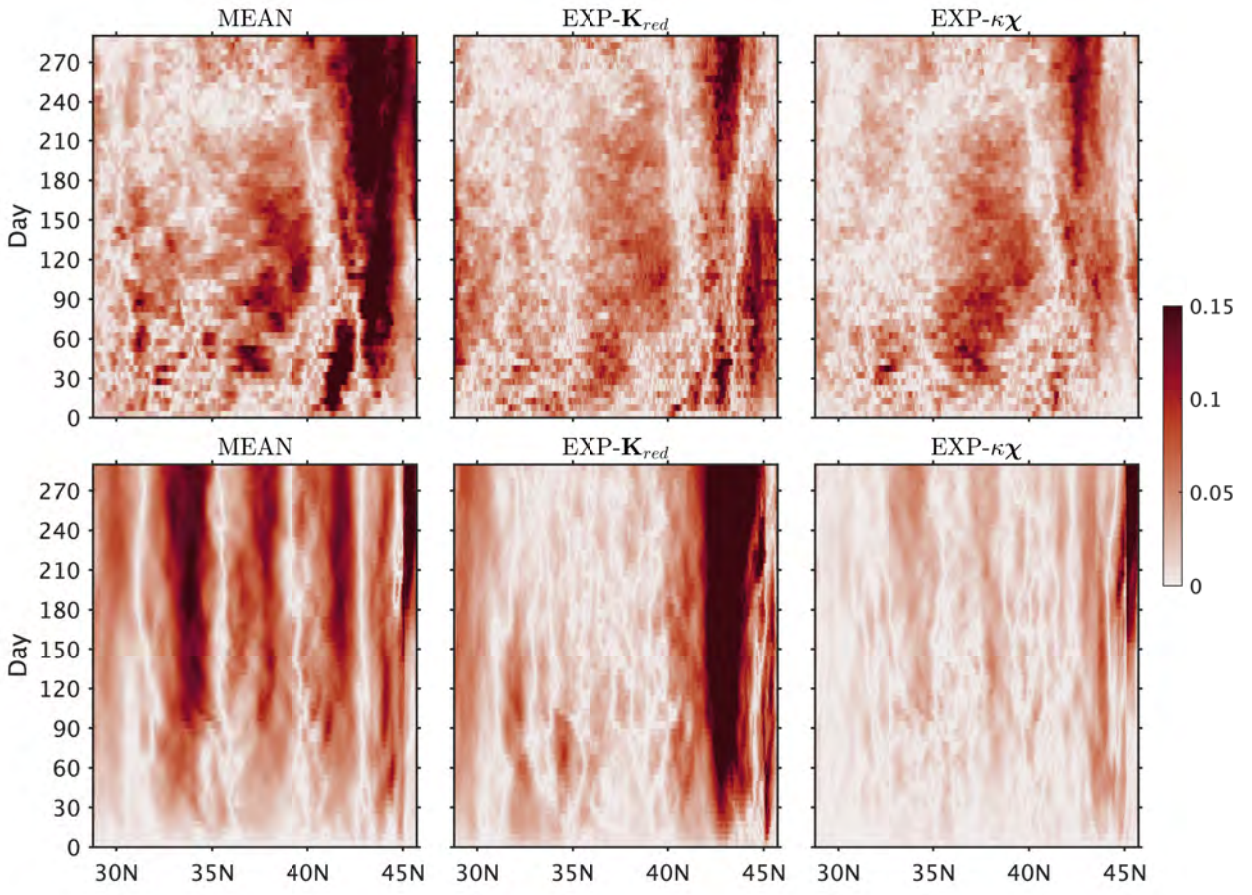


FIG. 11. The skill of different $\widehat{\mathcal{D}}$ in reproducing eddy stirring effect on a large-scale tracer profile. The figures show the temporal evolution of the relative difference in the zonally- and vertically-averaged tracer concentrations between the three sensitivity experiments and FULL. Tracer concentrations were averaged over the mixed layer (top row) and over the isopycnal layers below layer 15 (bottom row). The initial distribution is c_6.

by the spatial filter, the associated transport barrier breaks down, leading to a significant bias in tracer distributions.

c. Dispersion of a tracer patch

Mesoscale eddies disperse tracers anisotropically. The efficiency and preferential direction of the dispersion can be measured by the time rate of change of the second moment for a tracer patch

(Rypina et al. 2012). The second moment is defined as a covariance matrix (Wagner et al. 2019):

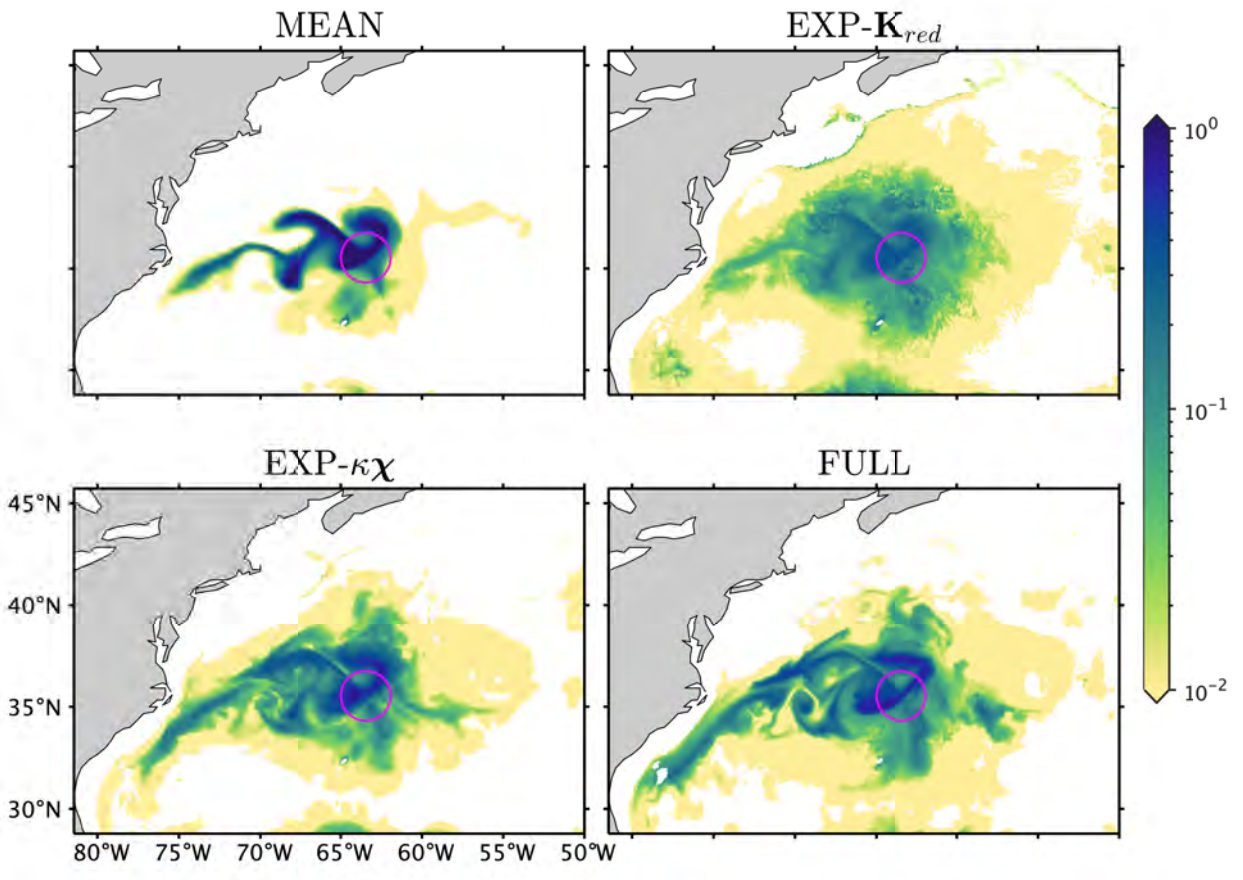
$$\sigma_{ij}^2 = \frac{\iint (x_i - x_i^c)(x_j - x_j^c) c(x,y)\, dx\, dy}{\iint c(x,y)\, dx\, dy}, \quad i,j = 1,2 \tag{20}$$

where $x_i^c = \iint x_i c(x,y)\, dx\, dy / \iint c(x,y)\, dx\, dy$ is the position of center of mass.

We released a tracer anomaly near the Gulf Stream axis at 35.5°N, 63.5°W. The initial tracer concentration distribution is vertically uniform and has a shape of a round patch with the diameter of 3 degrees. It is larger than the typical length scale of mesoscale eddies to prevent the tracer from being trapped in a single eddy like a Gulf Stream ring. Figure 12 shows the evolution of the vertically-averaged tracer patch in the sensitivity experiments. In all simulations, the patch propagates along the Gulf Stream axis due to the mean advection and disperses away due to the eddy stirring. Unsurprisingly, the tracer in FULL spreads much farther away from its initial position than in MEAN. The shape of the tracer patch in EXP-$\kappa\boldsymbol{\chi}$ is similar to that in the FULL run, which demonstrates the efficiency of GEIA in capturing anisotropic dispersion. In contrast, the spreading in EXP-$\mathbf{K}_{red}$ is more isotropic than in FULL.

A quantitative comparison is done through the dispersion covariance matrix. Similar to the symmetric tensor $\mathbf{S}$, the covariance matrix can be rotated onto a new coordinate, where the major axis is along the maximum dispersion direction, and the minor axis is perpendicular to it. Figure 13 compares the dispersion of the vertically-averaged tracers within the mixed layer. Along the major axis, all results are similar, because the large-scale sheared flow in the Gulf Stream dominates over the eddy effects. In contrast, the differences between the experiments are significant in the transverse direction, along the minor axis (e.g., Oh et al. 2000). We see that the FULL dispersion increases much faster and asymptotes to a larger value than in MEAN (Figure 13, bottom). Across the Gulf Stream, the dispersion in both the EXP-$\mathbf{K}_{red}$ and EXP-$\kappa\boldsymbol{\chi}$ cases is close to FULL, demonstrating that they can capture the eddy dispersive effects.

The improvement due to the generalized-advective-diffusive approach can be seen most clearly in the deeper layers (Figure 14). Along each axis, the dispersion in EXP-$\kappa\boldsymbol{\chi}$ is close to FULL. In contrast, the EXP-$\mathbf{K}_{red}$ dispersion along the minor axis is highly overestimated. The eddy-induced spreading in the Gulf Stream region is known to be strongly anisotropic (Rypina et al. 2012; Kamenkovich et al. 2015). Our results from EXP-$\mathbf{K}_{red}$ show that the instantaneous isotropic



694 FIG. 12. The skill of different $\widehat{\mathcal{D}}$ in reproducing eddy-induced dispersion of an isolated tracer patch. Shown is
695 the vertically-averaged tracer concentration below layer 15 at day 120 after release. Circle in magenta represents
696 the initial tracer patch (c_{t2}). Tracer concentration less than 1×10^{-3} (0.1% of the initial concentration) is set to
697 zero.

705 diffusivity with an anisotropic antisymmetric part mixes excessively along the minor axis. In
706 contrast, the anisotropic dispersion is reproduced effectively by using the generalized advection
707 (EXP-$\kappa\boldsymbol{\chi}$).

708 7. Conclusion and discussion

709 This study aims to make progress in understanding effects of large-scale and mesoscale currents
710 on passive tracer distribution. In light of the complexity of representing the eddy-induced transport

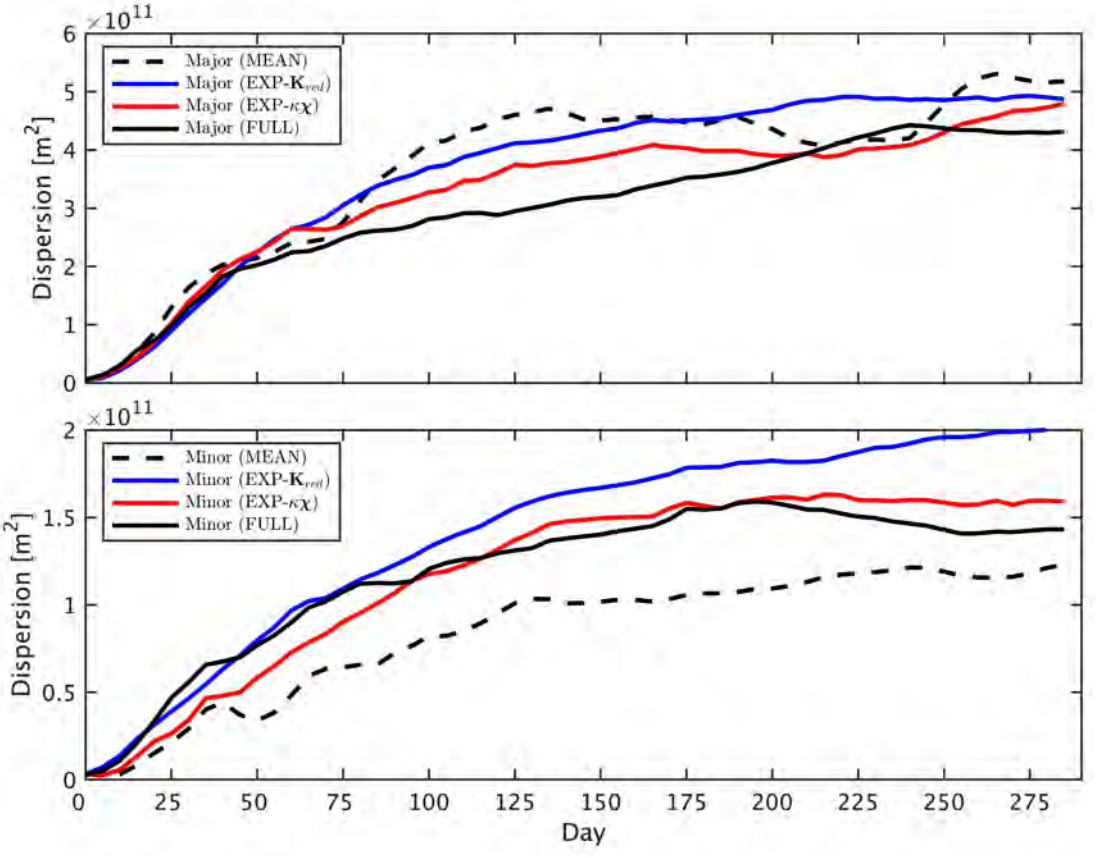


FIG. 13. Dispersion along the major (top panel) and minor (bottom panel) axes of the isolated tracer patch averaged within the mixed layer.

by the traditional flux-gradient relation, the study explores a new representation approach and evaluates the two approaches in a high-resolution model of the Gulf Stream region.

We focus on the lateral mesoscale eddy tracer transport, with "lateral" being defined as horizontal in the surface mixed layer and along-isopycnal in the ocean interior. Mesoscale eddies are defined broadly via a high-pass spatial filter, which is motivated by the need to study transient eddy-induced processes missing in coarse-resolution models. In this definition, the eddies include submesoscale currents that are contained mainly in the surface mixed layer and are partially resolved in our numerical simulations. This scale-based filter is different from the conventional Reynolds decomposition in that it introduces new eddy-mean "cross terms" that represent inter-scale interactions. Our results showed that these terms have distinct implications on tracer transport.

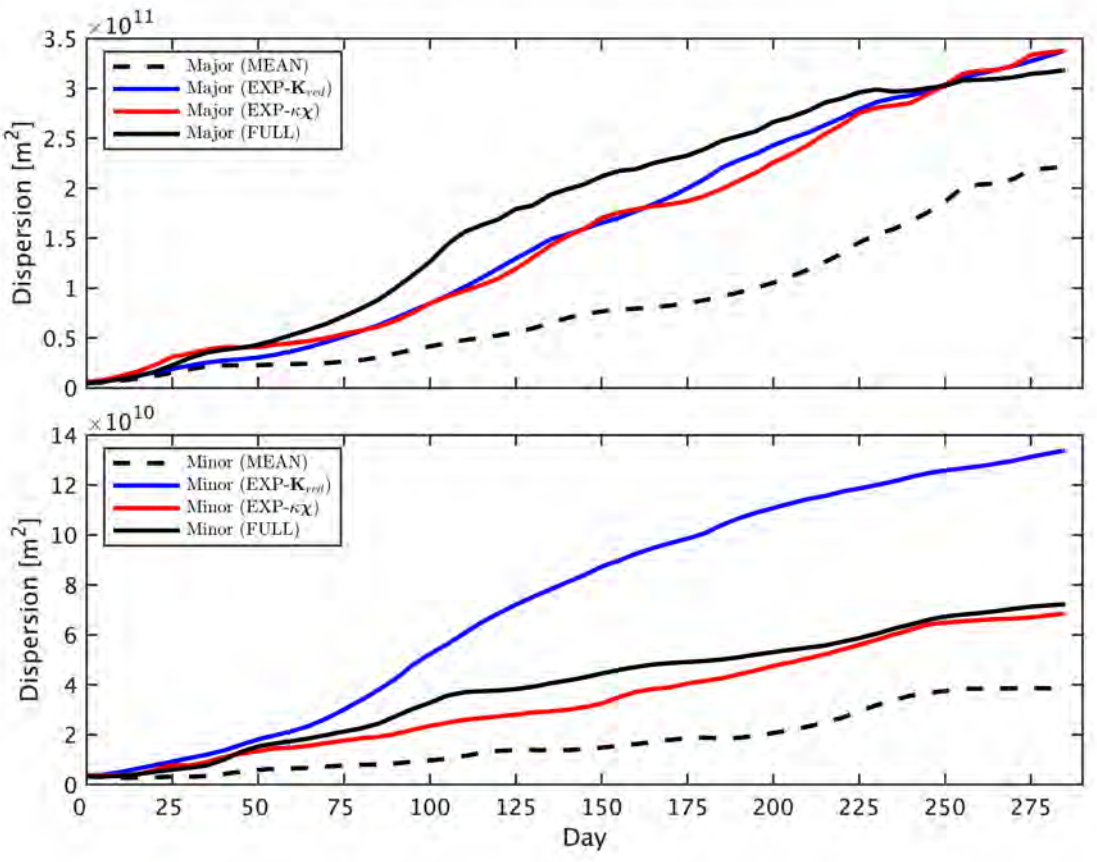


FIG. 14. Same with Figure 13 but for tracer patches averaged below layer 15.

For example, the divergence of $\mathbf{U}'\langle c\rangle$ dominates the lateral transport due to the divergence of the lateral eddy volume flux amplified by the large-scale tracer, $\langle c\rangle\nabla\cdot\mathbf{U}'$. This flux term not only contributes the most to the spatial variability of the eddy transport tensor, but is also a major source of biases in representing the eddy forcing by the flux-gradient relation. Using the mixing length theory, we also showed that $\mathbf{U}'\langle c\rangle$ cannot be properly approximated by the turbulent diffusion. These properties compromise the utility of the flux-gradient relation in isopycnal layers.

Besides the complications from the cross terms, the flux-gradient approach is also affected by the ambiguity due to the rotational (non-divergent) component in eddy tracer flux. The definition of this component and, thus, of the transport tensor, is not unique and is highly sensitive to the ill-defined boundary conditions. At the same time, errors in parameterization of the rotational component will be significant for tracer evolution, because the component tends to dominate the eddy tracer flux.

In addition, several non-flux and w-flux terms in the eddy forcing, i.e., $\partial_t ch' + \partial_s w'c$, cannot be readily represented by the flux-gradient relation in isopycnal layers and in a general computational domain. These challenges motivated us to consider an alternative to the flux-gradient model.

The proposed generalized-advective-diffusive approach is able to resolve the aforementioned issues by including a divergent generalized eddy-induced advective (GEIA) flux $\boldsymbol{\chi}$ and modeling the local eddy forcing. This study considered a simplified form of the approach that consists of isotropic diffusion and generalized advection. Although the GEIA flux is advective and introduces a well-defined direction in the eddy-induced stirring, a full diffusivity tensor would still be needed to capture anisotropic diffusive effects. We did not, however, use such a tensor in our tracer simulations because of persistent numerical instability that plausibly results from the opposite-signed diffusivities. Our results showed that the new model leads to reduced biases in representing the tracer eddy forcing. The biases cannot, however, be completely eliminated, because the parameters in both the flux-gradient and the new formulations still depend on tracers, and thus cannot be uniquely defined.

This study is one of the first attempts to use the eddy-induced advection and eddy-induced diffusion with their full spatio-temporal variability to represent the eddy effects. By performing a series of targeted tracer simulations in a high-resolution model, we showed the importance of the eddy-induced lateral advection. We also demonstrated that the new approach is more accurate in reproducing the stirring and dispersing effect of eddies than the flux-gradient approach.

One of the notable advantages of the generalized-advective-diffusive approach is that it can represent any form of the local eddy forcing. In this study, we focused on the lateral stirring for a direct comparison with the flux-gradient model. We used the new approach to represent the lateral eddy advection $\mathbf{U}' \cdot \nabla c$, and our analysis showed that it captures the full eddy effects more effectively than does the lateral eddy flux divergence $\nabla \cdot (\mathbf{U}'c)$. This could be one of the main reasons why the new approach outperforms the flux-gradient model besides the reduced tracer dependence.

The proposed approach could eventually lead to new eddy parameterizations. It should be re-emphasized, however, that this study is not a direct attempt at eddy parameterization for several reasons. Firstly, we do not consider eddy-induced mass fluxes. The isopycnal layer thickness eddy-induced velocity (EIV) $\mathbf{u}^*$ that can be parameterized by the Gent and McWilliams (1990)

closure in not studied here. In future developments, the GEIA flux $\boldsymbol{\chi}$ can be combined with $\mathbf{u}^*$ to advect tracers, uniting these two distinct eddy-induced advective effects: the former represents the advective eddy transport of tracers whereas the latter is the eddy advection of water mass in which the tracer is embedded. However, our calculation of $\mathbf{u}^*\langle h\rangle$, defined as $\langle \mathbf{U}\rangle - \langle \mathbf{u}\rangle\langle h\rangle$, showed that it is at least one order of magnitude smaller than $\boldsymbol{\chi}$. This result is in agreement with the conclusions from an idealized quasi-geostrophic study (Haigh et al. 2021b) that systematically compared the tracer EIV $\mathbf{u}_c^*$ and $\mathbf{u}^*$.

Secondly, the parameters $\mathbf{K}_{red}$, κ and $\boldsymbol{\chi}$ were diagnosed from the high-resolution reference simulation and applied to the same model. The use of a high-resolution model avoids additional errors associated with extrapolation and re-discretization of relevant physical terms to a coarser grid. In contrast, a parameterization scheme would need to be developed and tested on a coarse numerical grid and must involve a closure, that is, calculation of all the parameters from large-scale quantities. Development of closures is a difficult and unsettled task. For example, the physical expressions of Redi and GM coefficients are still challenging to derive (Marshall and Speer 2012; Mak et al. 2017), although their application has been fairly successful.

An important source of ambiguity in modelling eddy effects lies in the definition of eddies. This study used a spatial coarsening that may explain the strong variability in the eddy-induced diffusion and eddy-induced advection. Studies that use a long-term time- and/or zonal averaging alone (e.g., Bachman and Fox-Kemper 2013; Klocker and Abernathey 2014) or combine the temporal and spatial averaging (e.g., Bachman et al. 2020; Zhang and Wolfe 2022) to separate mesoscale and large-scale fields could find reduced variability of $\mathbf{K}$. For example, Zhang and Wolfe (2022) reported only occasional negative eigenvalues and a modest tracer dependence. This could be a result of the intensive smoothing they applied to eddy fluxes, which reduces variations in the tensor. Dynamically meaningful definition of eddies as the type of currents missing in coarse-resolution models is needed. A recently novel method is to define "dynamically unresolved eddies" (Agarwal et al. 2021; Berloff et al. 2021; Ryzhov and Berloff 2022), but it requires additional computations for passive tracers. The sensitivity of properties of the modeled eddy-induced transport to the definition of eddies and eddy forcing is an important research topic.

It is also an intriguing and practical question of how the (Eulerian) eddy diffusivities of the approaches in this study ($\mathbf{S}$ and κ) could be estimated by the Lagrangian method using drifters

(e.g., Lumpkin et al. 2002; LaCasce 2008) or a dye release experiment (e.g., Ledwell et al. 1998) in the real ocean. The task would be difficult in practice, because the required dense spatio-temporal coverage of the concurrent tracer and velocity fields is not available on a global scale. Besides the limitation of sparse observations, this open question roots in the unsettled physical connection between the Eulerian and Lagrangian (particle- or tracer-based) estimates in the complex oceanic flows. Even if a sufficient amount of Lagrangian observations can be obtained in a numerical model, the consistency between the two estimates has only been evaluated in a strongly zonal flow (Abernathey et al. 2013). This question is out of the scope of our study. We refer readers to Klocker et al. (2012b) and Qian et al. (2019) who used theoretical considerations to reconcile the different estimates of eddy diffusivities.

This study is motivated by the recently uncovered issues with the flux-gradient relation. However, we emphasize that these issues do yet not prove that the relation is fundamentally unsuitable for parameterization of eddies. The new approach suffers from some of the same issues: e.g., the tracer dependence in parameters is still large. Future works should be extended to include the non-flux eddy terms such as the time tendency of tracer mass (ch), and the anisotropic diffusion, which we leave in a companion study.

Acknowledgments. We thank the editor and three anonymous reviewers for their comments that led to substantial improvements to the manuscript. Y. Lu and I. Kamenkovich acknowledge the support of the NOAA grant NA16OAR4310165 and NSF grant 1849990. P. Berloff gratefully acknowledges the support by NERC grant NE/T002220/1 and Leverhulme grant RPG-2019-024.

Data availability statement. Processed data used to generate the results and figures are available at: https://doi.org/10.5281/zenodo.6946085. The analysis codes are available at: https://doi.org/10.5281/zenodo.6946093. Original model data and tracer model source code are available upon request from the author.

APPENDIX A

Initial distributions of tracers

Ten tracers were used in total for diagnosing the parameters in the flux-gradient and generalized-advective-diffusive approaches. Tracer concentrations were initialized with different horizontal

distributions and were taken constant in the vertical direction (Figure A1):

$$
\begin{aligned}
c_1 &= \cos 2\pi \frac{n_y}{N_y} + 2, \\
c_2 &= \cos 2\pi \frac{n_x}{N_x} + 2, \\
c_3 &= \cos 2\pi \left(\frac{n_x}{N_x} + \frac{n_y}{N_y} \right) + 2, \\
c_4 &= \cos 2\pi \left(\frac{N_x - n_x}{N_x} + \frac{n_y}{N_y} \right) + 2, \\
c_5 &= 0.5 * \left(\cos 4\pi \frac{n_x}{N_x} + \sin 4\pi \frac{n_y}{N_y} \right) + 2, \\
c_6 &= \cos 4\pi \frac{n_y}{N_y} + 2, \\
c_7 &= \cos 4\pi \frac{n_x}{N_x} + 2, \\
c_8 &= \cos 4\pi \left(\frac{n_x}{N_x} + \frac{n_y}{N_y} \right) + 2, \\
c_9 &= \cos 4\pi \left(\frac{N_x - n_x}{N_x} + \frac{n_y}{N_y} \right) + 2, \\
c_{10} &= 25 * \left(\frac{\left(0.4 \frac{n_x}{N_x} - 1 \right)^2}{9} + \frac{\left(0.4 \frac{n_y}{N_y} - 1 \right)^2}{16} \right),
\end{aligned}
\tag{A1}
$$

where n_x and n_y are grid indices in the zonal and meridional directions, respectively, and $N_x = 1573$ and $N_y = 1073$ are the total numbers of grid points in the corresponding directions. Tracers used for the experiments in section 6 have different initial distributions:

$$
\begin{aligned}
c_{t1} &= 0.5 * \cos \pi \frac{n_y}{N_y} + 0.5, \\
c_{t2} &= \begin{cases} 3 - 2 * \frac{d^2}{75^2}, & \text{if } d^2 \equiv (n_x - 900)^2 + (n_y - 400)^2 \le 75^2 \\ 0, & \text{otherwise} \end{cases}.
\end{aligned}
\tag{A2}
$$

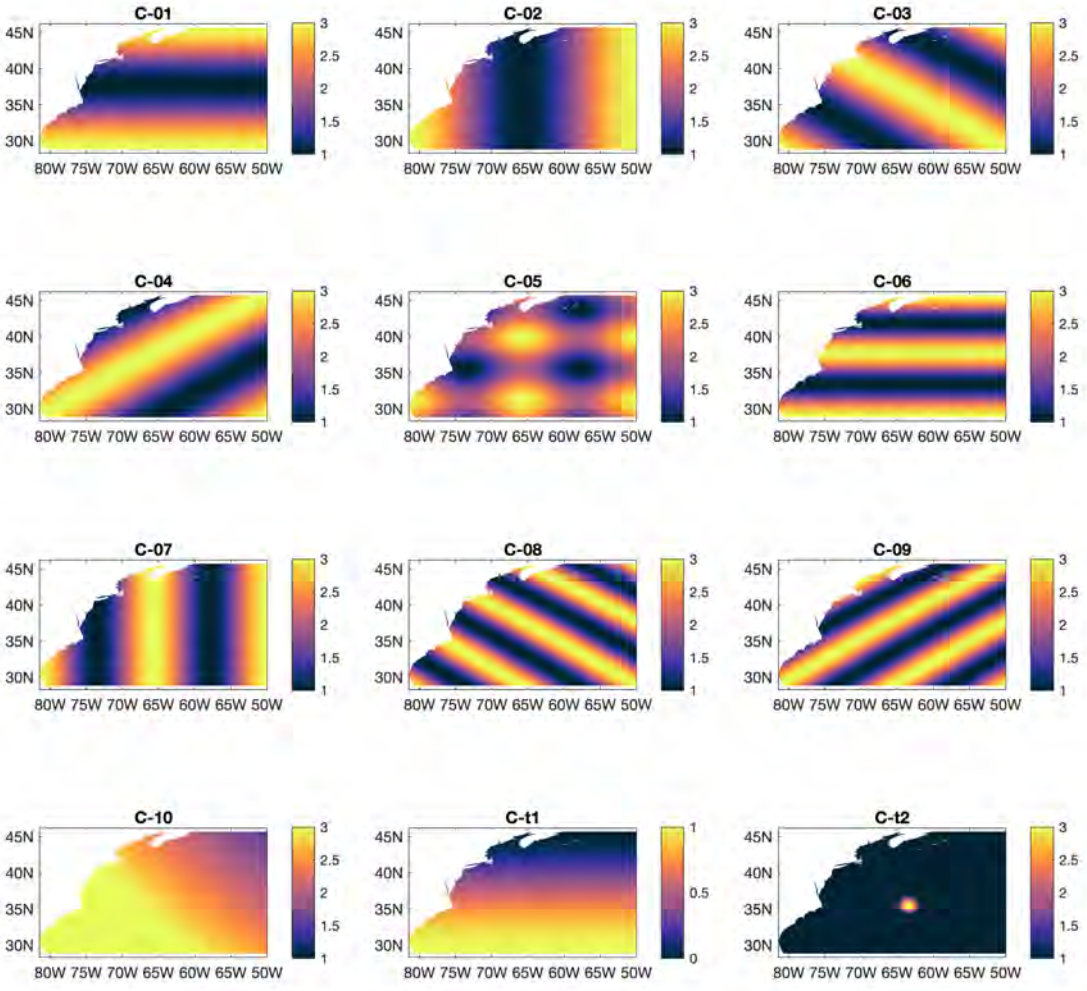


FIG. A1. Initial distributions of tracer fields. Tracers are constant in the vertical direction.

APPENDIX B

Derivation of the generalize-advective-diffusive approach

We start by modeling the eddy forcing field with combining the flux-gradient framework and a generalized advection:

$$\widehat{\mathcal{D}} = -\nabla \cdot (\langle h \rangle \mathbf{K} \nabla \langle c \rangle) + \mathbf{U}^{\chi} \cdot \nabla \langle c \rangle \,, \tag{B1}$$

where $\mathbf{U}^{\chi}$ is an independent (vector) parameter that can be divergent. Realizing that the spatial gradient of $\mathbf{K}$ also corresponds to advection, we expand (B1) as

$$\widehat{\mathcal{D}} = -\langle h \rangle K_{ij} \partial_i \partial_j \langle c \rangle + \{-\partial_i (\langle h \rangle K_{ij}) + U^{\chi}_j\} \partial_j \langle c \rangle \,, \tag{B2}$$

where i and j are used to index the spatial coordinates, and summation over the repeated indices is assumed. By splitting K_{ij}, in the first term on the right hand side, into its symmetric S_{ij} and antisymmetric A_{ij} parts, and by taking $A_{ij} = -A_{ji}$, (B2) becomes

$$\widehat{\mathcal{D}} = -\langle h \rangle \{S_{11} \partial_x \partial_x \langle c \rangle + 2 S_{12} \partial_x \partial_y \langle c \rangle + S_{22} \partial_y \partial_y \langle c \rangle\} + \boldsymbol{\chi} \cdot \nabla \langle c \rangle \,, \tag{B3}$$

where $\boldsymbol{\chi}(x, y, z, t)$ [m^2 s^{-1}] incorporating all advective terms, $-\partial_i (\langle h \rangle S_{ij}) + \hat{\mathbf{z}} \times \nabla (\langle h \rangle A_{21}) + \mathbf{U}^{\chi}$, is generally divergent, and is referred to as generalized eddy-induced advective (GEIA) flux. Both S_{ij} and $\boldsymbol{\chi}$ need to be diagnosed from tracer fields as independent parameters. Thus, the proposed approach involves a set of point-wise local problems (on each grid node).

This study considers a simplified form of (B3) by taking the diffusive part to be isotropic ($S_{12} = 0$ and $S_{11} = S_{22} = \kappa$) and by incorporating the antisymmetric part of the tensor into the generalized advection:

$$\widehat{\mathcal{D}} = -\langle h \rangle \kappa \nabla^2 \langle c \rangle + \boldsymbol{\chi} \cdot \nabla \langle c \rangle \,, \tag{B4}$$

where $\kappa(x, y, z, t)$ [m^2 s^{-1}] is an isotropic diffusivity. Here, we assumed that the anisotropy can be partly captured by the eddy-induced advection, since the advection has a well-defined direction (anisotropic).

References

Abernathey, R., D. Ferreira, and A. Klocker, 2013: Diagnostics of isopycnal mixing in a circumpolar channel. *Ocean Modell.*, **72**, 1–16, https://doi.org/10.1016/j.ocemod.2013.07.004.

Agarwal, N., E. Ryzhov, D. Kondrashov, and P. Berloff, 2021: Correlation-based flow decomposition and statistical analysis of the eddy forcing. *J. Fluid Mech.*, **924**, A5, https://doi.org/10.1017/jfm.2021.604.

Aluie, H., 2019: Convolutions on the sphere: commutation with differential operators. *Int. J. Geomath.*, **10 (1)**, 9, https://doi.org/10.1007/s13137-019-0123-9.

Aluie, H., M. Hecht, and G. K. Vallis, 2018: Mapping the energy cascade in the north atlantic ocean: The coarse-graining approach. *J. Phys. Oceanogr.*, **48 (2)**, 225–244.

Bachman, S., and B. Fox-Kemper, 2013: Eddy parameterization challenge suite I: Eady spindown. *Ocean Modell.*, **64**, 12–28, https://doi.org/10.1016/j.ocemod.2012.12.003.

Bachman, S., B. Fox-Kemper, and F. Bryan, 2015: A tracer-based inversion method for diagnosing eddy-induced diffusivity and advection. *Ocean Modell.*, **86**, 1–14, https://doi.org/10.1016/j.ocemod.2014.11.006.

Bachman, S. D., B. Fox-Kemper, and F. O. Bryan, 2020: A Diagnosis of Anisotropic Eddy Diffusion From a High-Resolution Global Ocean Model. *J. Adv. Model. Earth Syst.*, **12 (2)**, https://doi.org/10.1029/2019MS001904.

Bachman, S. D., B. Fox-Kemper, and B. Pearson, 2017: A scale-aware subgrid model for quasi-geostrophic turbulence. *J. Geophys. Res. Oceans*, **122 (2)**, 1529–1554, https://doi.org/10.1002/2016jc012265.

Berloff, P., E. Ryzhov, and I. Shevchenko, 2021: On dynamically unresolved oceanic mesoscale motions. *J. Fluid Mech.*, **920**, A41, https://doi.org/10.1017/jfm.2021.477.

Bleck, R., 2002: An oceanic general circulation model framed in hybrid isopycnic-Cartesian coordinates. *Ocean Modell.*, **4 (1)**, 55–88, https://doi.org/10.1016/S1463-5003(01)00012-9.

Capet, X., J. C. McWilliams, M. J. Molemaker, and A. F. Shchepetkin, 2008: Mesoscale to Submesoscale Transition in the California Current System. Part I: Flow Structure, Eddy Flux, and Observational Tests. *J. Phys. Oceanogr.*, **38 (1)**, 29–43, https://doi.org/10.1175/2007jpo3671.1.

Danabasoglu, G., and J. McWilliams, 1995: Sensitivity of the Global Ocean Circulation to Parameterizations of Mesoscale Tracer Transports. *J. Climate*, **8 (12)**, 2967–2987, https://doi.org/10.1175/1520-0442(1995)008<2967:SOTGOC>2.0.CO;2.

Eden, C., and R. J. Greatbatch, 2008: Towards a mesoscale eddy closure. *Ocean Modell.*, **20 (3)**, 223–239, https://doi.org/10.1016/j.ocemod.2007.09.002.

Ferrari, R., J. C. McWilliams, V. M. Canuto, and M. Dubovikov, 2008: Parameterization of Eddy Fluxes near Oceanic Boundaries. *J. Climate*, **21 (12)**, 2770–2789, https://doi.org/10.1175/2007jcli1510.1.

Ferrari, R., and M. Nikurashin, 2010: Suppression of Eddy Diffusivity across Jets in the Southern Ocean. *J. Phys. Oceanogr.*, **40 (7)**, 1501–1519, https://doi.org/10.1175/2010jpo4278.1.

Fox-Kemper, B., R. Ferrari, and J. Pedlosky, 2003: On the Indeterminacy of Rotational and Divergent Eddy Fluxes. *J. Phys. Oceanogr.*, **33**, 6.

Garabato, A. C. N., X. Yu, J. Callies, R. Barkan, K. L. Polzin, E. E. Frajka-Williams, C. E. Buckingham, and S. M. Griffies, 2022: Kinetic Energy Transfers between Mesoscale and Submesoscale Motions in the Open Ocean's Upper Layers. *J. Phys. Oceanogr.*, **52 (1)**, 75–97, https://doi.org/10.1175/JPO-D-21-0099.1.

Gent, P., and J. McWilliams, 1990: Isopycnal mixing in ocean circulation models. *Ocean Modell.*, **20 (1)**, 150–155, https://doi.org/10.1175/1520-0485(1990)020<0150:IMIOCM>2.0.CO;2.

Gent, P., J. Willebrand, T. McDougall, and J. McWilliams, 1995: Parameterizing Eddy-Induced Tracer Transports in Ocean Circulation Models. *Ocean Modell.*, **25 (4)**, 463–474, https://doi.org/10.1175/1520-0485(1995)025<0463:PEITTI>2.0.CO;2.

Gent, P. R., A. P. Craig, C. M. Bitz, and J. W. Weatherly, 2002: Parameterization Improvements in an Eddy-Permitting Ocean Model for Climate. *J. Climate*, **15**, 1447–1459, https://doi.org/10.1175/1520-0442(2002)015<1447:PIIAEP>2.0.CO;2.

George, T. M., G. E. Manucharyan, and A. F. Thompson, 2021: Deep learning to infer eddy heat fluxes from sea surface height patterns of mesoscale turbulence. *Nat. Commun.*, **12 (1)**, 800, https://doi.org/10.1038/s41467-020-20779-9.

Gnanadesikan, A., M. Pradal, and R. Abernathey, 2015: Isopycnal mixing by mesoscale eddies significantly impacts oceanic anthropogenic carbon uptake. *Geophys. Res. Lett.*, **42 (11)**, 4249–4255, https://doi.org/10.1002/2015GL064100.

Griesel, A., S. T. Gille, J. Sprintall, J. L. McClean, and M. E. Maltrud, 2009: Assessing eddy heat flux and its parameterization: A wavenumber perspective from a 1/10° ocean simulation. *Ocean Modell.*, **29 (4)**, 248–260, https://doi.org/10.1016/j.ocemod.2009.05.004.

Griffies, S. M., 1998: The Gent–McWilliams Skew Flux. *J. Phys. Oceanogr.*, **28**, 831–841, https://doi.org/10.1175/1520-0485(1998)028<0831:TGMSF>2.0.CO;2.

Haigh, M., and P. Berloff, 2021: On co-existing diffusive and anti-diffusive tracer transport by oceanic mesoscale eddies. *Ocean Modell.*, **168**, 101 909, https://doi.org/10.1016/j.ocemod.2021.101909.

Haigh, M., L. Sun, J. C. McWilliams, and P. Berloff, 2021a: On eddy transport in the ocean. Part I: The diffusion tensor. *Ocean Modell.*, **164**, 101 831, https://doi.org/10.1016/j.ocemod.2021.101831.

Haigh, M., L. Sun, J. C. McWilliams, and P. Berloff, 2021b: On eddy transport in the ocean. Part II: The advection tensor. *Ocean Modell.*, **165**, 101 845, https://doi.org/10.1016/j.ocemod.2021.101845.

Haigh, M., L. Sun, I. Shevchenko, and P. Berloff, 2020: Tracer-based estimates of eddy-induced diffusivities. *Deep Sea Res. I: Oceanogr. Res. Pap.*, **160**, 103 264, https://doi.org/10.1016/j.dsr.2020.103264.

Hewitt, H. T., and Coauthors, 2020: Resolving and Parameterising the Ocean Mesoscale in Earth System Models. *Curr. Clim. Change Rep.*, **6 (4)**, 137–152, https://doi.org/10.1007/s40641-020-00164-w.

Jayne, S. R., and J. Marotzke, 2002: The Oceanic Eddy Heat Transport. *J. Phys. Oceanogr.*, **32**, 18, https://doi.org/10.1175/1520-0485(2002)032<3328:TOEHT>2.0.CO;2.

Kamenkovich, I., P. Berloff, M. Haigh, L. Sun, and Y. Lu, 2021: Complexity of Mesoscale Eddy Diffusivity in the Ocean. *Geophys. Res. Lett.*, **48 (5)**, https://doi.org/10.1029/2020GL091719.

Kamenkovich, I., and Z. Garraffo, 2022: Importance of mesoscale currents in amoc pathways and timescales. *J. Phys. Oceanogr.*, https://doi.org/https://doi.org/10.1175/JPO-D-21-0244.1.

Kamenkovich, I., Z. Garraffo, R. Pennel, and R. A. Fine, 2017: Importance of mesoscale eddies and mean circulation in ventilation of the Southern Ocean. *J. Geophys. Res. Oceans*, **122 (4)**, 2724–2741, https://doi.org/10.1002/2016JC012292.

Kamenkovich, I., I. I. Rypina, and P. Berloff, 2015: Properties and Origins of the Anisotropic Eddy-Induced Transport in the North Atlantic. *J. Phys. Oceanogr.*, **45 (3)**, 778–791, https://doi.org/10.1175/JPO-D-14-0164.1.

Klocker, A., and R. Abernathey, 2014: Global Patterns of Mesoscale Eddy Properties and Diffusivities. *Ocean Modell.*, **44 (3)**, 1030–1046, https://doi.org/10.1175/JPO-D-13-0159.1.

Klocker, A., R. Ferrari, and J. H. LaCasce, 2012a: Estimating Suppression of Eddy Mixing by Mean Flows. *J. Phys. Oceanogr.*, **42**, 1566–1576, https://doi.org/10.1175/JPO-D-11-0205.1.

Klocker, A., R. Ferrari, J. H. LaCasce, and S. T. Merrifield, 2012b: Reconciling float-based and tracer-based estimates of lateral diffusivities. *J. Mar. Res.*, **70**, 569–602, https://doi.org/10.1357/002224012805262743.

Kuhlbrodt, T., R. Smith, Z. Wang, and J. Gregory, 2012: The influence of eddy parameterizations on the transport of the Antarctic Circumpolar Current in coupled climate models. *Ocean Modell.*, **52-53**, 1–8, https://doi.org/10.1016/j.ocemod.2012.04.006.

LaCasce, J., 2008: Statistics from Lagrangian observations. *Prog. Oceanogr.*, **77 (1)**, 1–29, https://doi.org/10.1016/j.pocean.2008.02.002.

Ledwell, J. R., A. J. Watson, and C. S. Law, 1998: Mixing of a tracer in the pycnocline. *J. Geophys. Res. Oceans*, **103 (C10)**, 21 499–21 529, https://doi.org/10.1029/98jc01738.

Leonard, A., 1997: Large-eddy simulation of chaotic convection and beyond. *35th Aerospace Sciences Meeting and Exhibit*, https://doi.org/10.2514/6.1997-204.

Li, Z., Y. Chao, and J. C. McWilliams, 2006: Computation of the Streamfunction and Velocity Potential for Limited and Irregular Domains. *Mon. Wea. Rev.*, **134 (11)**, 3384–3394, https://doi.org/10.1175/MWR3249.1.

Lumpkin, R., A.-M. Treguier, and K. Speer, 2002: Lagrangian Eddy Scales in the Northern Atlantic Ocean. *J. Phys. Oceanogr.*, **32 (9)**, 2425–2440, https://doi.org/10.1175/1520-0485-32.9.2425.

Maddison, J., D. Marshall, and J. Shipton, 2015: On the dynamical influence of ocean eddy potential vorticity fluxes. *Ocean Modell.*, **92**, 169–182, https://doi.org/10.1016/j.ocemod.2015.06.003.

Mak, J., D. Marshall, J. Maddison, and S. Bachman, 2017: Emergent eddy saturation from an energy constrained eddy parameterisation. *Ocean Modell.*, **112**, 125–138, https://doi.org/10.1016/j.ocemod.2017.02.007.

Mana, P. P., and L. Zanna, 2014: Toward a stochastic parameterization of ocean mesoscale eddies. *Ocean Modell.*, **79**, 1–20, https://doi.org/10.1016/j.ocemod.2014.04.002.

Marshall, D. P., and A. J. Adcroft, 2010: Parameterization of ocean eddies: Potential vorticity mixing, energetics and arnold's first stability theorem. *Ocean Modell.*, **32 (3)**, 188–204, https://doi.org/10.1016/j.ocemod.2010.02.001.

Marshall, J., and G. Shutts, 1981: A note on rotational and divergent eddy fluxes. *J. Phys. Oceanogr.*, **11 (12)**, 1677–503, https://doi.org/10.1175/1520-0485(1981)011<1677:ANORAD>2.0.CO;2.

Marshall, J., and K. Speer, 2012: Closure of the meridional overturning circulation through Southern Ocean upwelling. *Nat. Geosci.*, **5 (3)**, 171–180, https://doi.org/10.1038/ngeo1391.

Meijers, A. J. S., 2014: The Southern Ocean in the Coupled Model Intercomparison Project phase 5. *Phil. Trans. R. Soc. A.*, **372 (2019)**, 20130 296, https://doi.org/10.1098/rsta.2013.0296.

Mensa, J. A., Z. Garraffo, A. Griffa, T. M. Özgökmen, A. Haza, and M. Veneziani, 2013: Seasonality of the submesoscale dynamics in the Gulf Stream region. *Ocean Dyn.*, **63 (8)**, 923–941, https://doi.org/10.1007/s10236-013-0633-1.

Oh, I. S., V. Zhurbas, and W. Park, 2000: Estimating horizontal diffusivity in the East Sea (Sea of Japan) and the northwest Pacific from satellite-tracked drifter data. *J. Geophys. Res. Oceans*, **105 (C3)**, 6483–6492, https://doi.org/https://doi.org/10.1029/2000JC900002.

Patching, S., 2022: On divergence- and gradient-preserving coarse-graining for finite volume primitive equation ocean models. *Ocean Modell.*, **170**, 101 941, https://doi.org/10.1016/j.ocemod.2021.101941.

Prandtl, L., 1925: Bericht über untersuchungen zur ausgebildeten turbulenz. *Z. Angew. Math. Mech.*, **5 (2)**, 136–139, https://doi.org/10.1002/zamm.19250050212.

Pratt, L., R. Barkan, and I. Rypina, 2016: Scalar Flux Kinematics. *Fluids*, **1 (27)**, https://doi.org/doi:10.3390/fluids1030027.

Qian, Y.-K., S. Peng, and C.-X. Liang, 2019: Reconciling lagrangian diffusivity and effective diffusivity in contour-based coordinates. *J. Phys. Oceanogr.*, **49**, 1521–1539, https://doi.org/10.1175/JPO-D-18-0251.1.

Redi, M., 1982: Oceanic Isopycnal Mixing by Coordinate Rotation. *Ocean Modell.*, **12 (10)**, 1154–1158, https://doi.org/https://doi.org/10.1175/1520-0485(1982)012<1154:OIMBCR>2.0.CO;2.

Riha, S., and C. Eden, 2011: Lagrangian and Eulerian lateral diffusivities in zonal jets. *Ocean Modell.*, **39 (1-2)**, 114–124, https://doi.org/10.1016/j.ocemod.2011.02.002.

Rypina, I. I., I. Kamenkovich, P. Berloff, and L. J. Pratt, 2012: Eddy-Induced Particle Dispersion in the Near-Surface North Atlantic. *J. Phys. Oceanogr.*, **42 (12)**, 2206–2228, https://doi.org/10.1175/JPO-D-11-0191.1.

Rypina, I. I., L. J. Pratt, and M. S. Lozier, 2011: Near-Surface Transport Pathways in the North Atlantic Ocean: Looking for Throughput from the Subtropical to the Subpolar Gyre. *Ocean Modell.*, **41 (5)**, 911–925, https://doi.org/10.1175/2011JPO4498.1.

Ryzhov, E., and P. Berloff, 2022: On transport tensor of dynamically unresolved oceanic mesoscale eddies. *JFM*, **939**, A7, https://doi.org/10.1017/jfm.2022.169.

Sallée, J.-B., R. J. Matear, S. R. Rintoul, and A. Lenton, 2012: Localized subduction of anthropogenic carbon dioxide in the Southern Hemisphere oceans. *Nat. Geosci.*, **5 (8)**, 579–584, https://doi.org/10.1038/ngeo1523.

Sallée, J.-B., K. Speer, R. Morrow, and R. Lumpkin, 2008: An estimate of lagrangian eddy statistics and diffusion in the mixed layer of the Southern Ocean. *Nat. Geosci.*, **66 (4)**, 441–463(23), https://doi.org/10.1357/002224008787157458.

Sun, L., M. Haigh, I. Shevchenko, P. Berloff, and I. Kamenkovich, 2021: On non-uniqueness of the mesoscale eddy diffusivity. *J. Fluid Mech.*, **920**, A32, https://doi.org/10.1017/jfm.2021.472.

Taylor, G. I., 1921: Diffusion by Continuous Movements. *Proc. London Math. Soc.*, **s2-20 (1)**, 196–212, https://doi.org/10.1112/plms/s2-20.1.196.

Trias, F. X., F. Dabbagh, A. Gorobets, and C. Oliet, 2020: On a Proper Tensor-Diffusivity Model for Large-Eddy Simulation of Buoyancy-Driven Turbulence. *Flow Turbulence Combust*, **105 (2)**, 393–414, https://doi.org/10.1007/s10494-020-00123-3.

Wagner, P., S. Rühs, F. U. Schwarzkopf, I. M. Koszalka, and A. Biastoch, 2019: Can Lagrangian Tracking Simulate Tracer Spreading in a High-Resolution Ocean General Circulation Model? *Ocean Modell.*, **49 (5)**, 1141–1157, https://doi.org/10.1175/JPO-D-18-0152.1.

Waterman, S., N. G. Hogg, and S. R. Jayne, 2011: Eddy–mean flow interaction in the Kuroshio Extension Region. *J. Phys. Oceanogr.*, **41 (6)**, 1182–1208, https://doi.org/10.1175/2010JPO4564.1.

Zhang, W., and C. Wolfe, 2022: On the vertical structure of oceanic mesoscale tracer diffusivities. *J. Adv. Model. Earth Syst.*, https://doi.org/10.1029/2021MS002891.