CrossMark

JM PAPERS

On non-uniqueness of the mesoscale eddy diffusivity

Luolin Sun[1],†, Michael Haigh[1], Igor Shevchenko[1], Pavel Berloff[1,2] and Igor Kamenkovich[3]

[1]Department of Mathematics, Imperial College London, London SW7 2AZ, UK

[2]Institute of Numerical Mathematics of the Russian Academy of Sciences, Moscow, 119333, Russia

[3]RSMAS, University of Miami, Miami, FL 33149, USA

Oceanic mesoscale currents ('eddies') can have significant effects on the distributions of passive tracers. The associated inhomogeneous and anisotropic eddy fluxes are traditionally parametrised using a transport tensor (K-tensor), which contains both diffusive and advective components. In this study, we analyse the eddy transport tensor in a quasigeostrophic double-gyre flow. First, the flow and passive tracer fields are decomposed into large- and small-scale (eddy) components by spatial filtering, and the resulting eddy forcing includes an eddy tracer flux representing advection by eddies and non-advective terms. Second, we use the flux-gradient relation between the eddy fluxes and the large-scale tracer gradient to estimate the associated K-tensors in their entire structural, spatial and temporal complexity, without making any additional assumptions or simplifications. The divergent components of the eddy tracer fluxes are extracted via the Helmholtz decomposition, which yields a divergent tensor. The remaining rotational flux does not affect the tracer evolution, but dominates the total tracer flux, affecting both its magnitude and spatial structure. However, in terms of estimating the eddy forcing, the transport tensor prevails over its divergent counterpart because of the significant numerical errors induced by the Helmholtz decomposition. Our analyses demonstrate that, in general, the K-tensor for the eddy forcing is not unique, that is, it is tracer-dependent. Our study raises serious questions on how to interpret and use various estimates of K-tensors obtained from either observations or eddy-resolving model solutions.

Key words: turbulent mixing, mixing and dispersion, quasi-geostrophic flows

† Email address for correspondence: luolin.sun14@imperial.ac.uk

L. Sun, M. Haigh, I. Shevchenko, P. Berloff and I. Kamenkovich

1. Introduction

The focus of this study is on material transport by mesoscale eddies, which are defined here as geostrophic currents with length scales shorter than approximately 200–500 km. These currents have a significant influence on the distributions of various oceanic tracers. Eddy tracer transport has been conventionally quantified by turbulent eddy diffusion – this study describes the lateral eddy-induced tracer fluxes using an inhomogeneous and time-dependent, two-dimensional K-tensor. We call it the 'transport tensor' because it parametrises both diffusive and advective effects. From a practical point of view, the eddy transport tensor can be used to represent eddy-induced fluxes in ocean models that either completely miss the mesoscales or only partially resolve them. The approach is based on the assumption of the flux-gradient relation: the eddy tracer flux is proportional to (minus) the large-scale tracer gradient. It is implicitly assumed that the transport tensor is uniquely determined by the turbulent flow and stratification, and is independent of the tracer field itself. Relying on this assumption, one can then obtain meaningful transport tensor estimates using a variety of Eulerian and Lagrangian methods. If, however, the transport tensor is not unique for each physical flow field and is instead a function of the tracer field, the utility of the model becomes questionable. This non-uniqueness is the focus of the present study.

Previous studies have accumulated significant evidence of the inherent complexity of the eddy transport tensor. Observation-based estimates exhibit strong dependence on depth and geographical location (Lumpkin, Treguier & Speer 2002; Marshall *et al.* 2006; Zhai & Greatbatch 2006; Abernathey & Marshall 2013; Klocker & Abernathey 2014; Canuto *et al.* 2019). The transport tensor is also anisotropic, which means that all components of the tensor are significantly different and the magnitude of the eddy flux depends on the orientation of the tracer gradient. For example, Lagrangian statistics for real-ocean and model-simulated floats and drifters predict a direction of maximal dispersion (Freeland, Rhines & Rossby 1975; O'Dwyer *et al.* 2000; McClean *et al.* 2002; Sallée *et al.* 2008; Kamenkovich, Berloff & Pedlosky 2009; Rypina *et al.* 2012; Kamenkovich, Rypina & Berloff 2015). Recent tracer-based Eulerian estimates also exhibit significant inhomogeneity and anisotropy of the transport tensor (Berloff 2016; Bachman, Fox-Kemper & Bryan 2020). Note, despite the similar findings, differences exist in the transport tensors from Lagrangian and Eulerian methods and such differences are poorly understood. Our study focuses only on the Eulerian eddy transport tensor, and any connections to the Lagrangian counterpart are beyond the scope of our study.

The origins of transport anisotropy remain unclear, but several mechanisms can be at play, including eddy propagation along mean currents (Killworth 1997; Abernathey *et al.* 2010; Ferrari & Nikurashin 2010; Klocker, Ferrari & LaCasce 2012*a*), shear dispersion (Smith 2005) and partial transport barriers in strong currents (Samelson 1992; Dritschel & McIntyre 2008; Berloff, Kamenkovich & Pedlosky 2009; Rypina, Pratt & Lozier 2011). In addition to the time-mean currents, the transport anisotropy can be explained by propagating elongated transients (Kamenkovich *et al.* 2015; Rudko *et al.* 2018). Eddy-induced transport along mean currents can be dwarfed by strong mean advection, which can justify the focus on the cross-flow ('effective') diffusivities (Nakamura 1996; Marshall *et al.* 2006; Shuckburgh *et al.* 2009; Abernathey *et al.* 2010; Klocker *et al.* 2012*a*; Abernathey & Marshall 2013). However, weaker mean flows, typical of most of the ocean, will render the two-dimensionality (and anisotropy) of the transport important. In particular, Kamenkovich *et al.* (2015) demonstrated that the eddy-induced transport in the along-stream direction is as important as the mean advection in most of the North Atlantic and, therefore, must be taken into account in eddy parametrisations.

In recent studies, a broad agreement on the isopycnal diffusivity (a scalar transfer coefficient quantifying down-gradient fluxes), its magnitude and its spatial structure, has been demonstrated by the comparisons among the scalar estimates from the Lagrangian and the Eulerian approaches (e.g. the effective diffusivity, Lagrangian dispersion approaches and multiple tracer inversion) (Klocker *et al.* 2012*b*; Abernathey, Ferreira & Klocker 2013; Wolfram & Ringler 2017). Moreover, a relative equivalence of the Eulerian diffusivities was also approved among different tracers, such as potential vorticity and passive tracers, even though Eulerian methods are sensitive to the tracer choices (Abernathey *et al.* 2013). These studies strongly indicate that the eddy diffusivity is a real object and thus unique.

However, the transport tensor estimates from a direct tracer-based method expose its inherent non-uniqueness, that is, dependence on the tracer fields. This direct tracer-based method was initially introduced by Plumb & Mahlman (1987) for the atmosphere and then applied to oceanic models by Bachman & Fox-Kemper (2013), Bachman, Fox-Kemper & Bryan (2015), Bachman *et al.* (2017) and Bachman *et al.* (2020). For a channel flow, Bachman & Fox-Kemper (2013), Bachman *et al.* (2015) and Bachman *et al.* (2017) computed a two-dimensional transport tensor which relates the zonal-mean eddy fluxes and tracer gradients via the inversion method. This approach has been recently extended to the lateral diffusivity in a realistic three-dimensional ocean flow (Bachman *et al.* 2020). These studies reported a significant sensitivity of their tensor estimates to the tracer groups of multiple sizes, and this non-uniqueness was attributed to the inaccuracy of the inversion method. Constraints on the choice of the tracers were discussed by Bachman *et al.* (2015, 2020) with the aim to ensure the accuracy of the inversion method and obtain a more universal fit. Considerations included the number of tracers used in the linear system (2.18), the initialisation of the tracer profiles and the strength of relaxation of the tracers. Our study adapts the inversion method for passive tracers in a two-dimensional framework and analyses these options.

The importance of rotational fluxes in eddy transport requires further investigation. Plumb (1979) suggested that the non-divergent part of an eddy potential vorticity (PV) flux should be removed because it does not influence the dynamics. This flux, which is associated with the spatial growth or decay of eddies, is also considered dynamically unimportant because it balances the advection of the mean flow (Marshall & Shutts 1981). Furthermore, Marshall & Shutts (1981) argued that negative diffusivities of heat – which are often undesirable owing to model stability issues – arise from the rotational component of the heat flux. However, more recent results from Haigh *et al.* (2020) and Mak, Maddison & Marshall (2016) using active and passive tracers showed that the negative diffusivities persist after removing the rotational flux. In contrast, Bachman *et al.* (2015) posed that any flux decomposition should not be combined with the tracer-independent inversion method because the rotational flux is tracer-dependent. The separation of the flux into the rotational and divergent components is also not unique and depends on the choice of boundary conditions for each of these flux components. Furthermore, additional boundary conditions applied to (2.21) and the large numerical errors caused by the decomposition will reduce the reliability of the results. However, if the large rotational component is not removed from the total tracer flux (as by Bachman *et al.* 2020), multi-tracer transport tensor fits can lead to even larger biases in the estimates of the eddy-flux divergence. In our study, the transport tensor and its divergent counterpart are compared to enable an interpretation of the effects of the rotational flux.

A precursor to the present study is that of Haigh *et al.* (2020), hereafter HSSB20, in which essential statistics pertaining to the eddy transport tensor were presented.

HSSB20 used a quasigeostrophic (QG) double-gyre model with a spatial filter – as opposed to a Reynolds average – to separate the flow and tracer field scales. The study considered the eigenvalues of the symmetric (i.e. diffusive) part of the transport tensor and found polar, i.e. opposite-signed, eigenvalues to be a dominant feature at all depths. The amplitudes of the eigenvalues were typically two orders of magnitude smaller after removal of the rotational part of the eddy tracer flux, owing to the dominance of such fluxes (Marshall & Shutts 1981). In general, the transport tensor considered by HSSB20 was tracer-dependent. The aim of the present study is to discuss this dependence in depth.

In addition to the eddy tracer flux, non-advective eddy terms, which are filtered out in a Reynolds decomposition, remain in the large-scale tracer equation owing to our spatial filtering, and thus need to be parametrised. One way of doing this is to absorb the terms into the flux-gradient relation so that the resulting K-tensor corresponds to the total eddy term. We diagnose this new tensor and complete the discussion of the tracer dependence initiated in HSSB20.

The paper is organised as follows. In §2, we describe the tracer experiments, which are set up in a double-gyre QG model. Section 3 discusses the numerical accuracy of the methodology, the effects of the rotational flux and the tracer-dependence of K-tensors under multiple circumstances. Finally, we conclude and discuss the direction of future studies in §4.

2. Problem formulation

In this section, we outline the ocean model and the K-tensors on which this study focuses. We describe the three-layer QG model in which passive tracers are advected in §2.1. In §2.2, we introduce the scale-aware flow decomposition, the K-tensor inversion method and K-tensor ($\mathbf{K}_f$) for the eddy tracer flux. In §2.3, we describe the Helmholtz decomposition used to remove the rotational flux and introduce the K-tensor ($\mathbf{K}_{div}$) for the divergent part of the eddy tracer flux. In §2.4, we apply the flux-gradient relation to non-advective eddy terms and introduce the K-tensor ($\mathbf{K}_g$) for these terms. Finally, we present the details of the experiments in §2.5.

2.1. *Ocean model*

A three-layer QG model for double-gyre, mid-latitude circulation, driven by wind forcing, is configured in a square basin with rigid lateral boundaries and flat bottom topography. The governing equations for the potential vorticity (PV) anomalies q_l are

$$\frac{\partial q_l}{\partial t} + \frac{\partial (u_l q_l)}{\partial x} + \frac{\partial (v_l q_l)}{\partial y} = \delta_{1l} f_{wind} - \beta v_l - \delta_{3l} \mu_{bot} \nabla^2 \psi_l + \mu_{eddy} \nabla^4 \psi_l, \qquad (2.1)$$

where $l = 1, 2, 3$ denotes the index for the top, intermediate and bottom layers which have depths $H_1 = 0.25$ km, $H_2 = 0.75$ km and $H_3 = 3$ km, respectively, and δ_{kl} is the Kronecker symbol. The PV anomalies are related to the velocity streamfunctions ψ_l through the elliptic equations:

$$\left.\begin{aligned}
q_1 &= \nabla^2 \psi_1 - s_1(\psi_1 - \psi_2), \\
q_2 &= \nabla^2 \psi_2 - s_{21}(\psi_1 - \psi_2) - s_{22}(\psi_2 - \psi_3), \\
q_3 &= \nabla^2 \psi_3 - s_3(\psi_3 - \psi_2).
\end{aligned}\right\} \qquad (2.2)$$

Parameter symbol	Parameter name	Dimensional values	Units
L	Basin length	3840	km
β	Meridional Coriolis gradient	2×10^{-11}	$m^{-1} s^{-1}$
μ_{bot}	Bottom friction	4×10^{-8}	s
μ_{eddy}	Eddy viscosity	20	$m^2 s^{-1}$
$\{s_1, s_{21}, s_{22}, s_3\}$	Stratification parameters	$\{22.3, 7.42, 9.83, 2.46\} \times 10^{-7}$	m^{-2}
τ	Wind stress amplitude	0.4	$N m^2$
ρ_1	Top-layer density	10^3	$kg m^3$

Table 1. Parameters of the QG model.

The wind forcing is chosen to be asymmetric with respect to the middle latitude of the basin to avoid artificial symmetry of the gyres:

$$f_{wind}(x, y) = \left\{ \begin{array}{ll} A \sin\left(\dfrac{\pi y}{y_0}\right), & 0 \le y < y_0, \\[3mm] -A \sin\left(\dfrac{\pi(y - y_0)}{L - y_0}\right), & y_0 \le y < L, \\[3mm] y_0 = \dfrac{L}{2} + 0.2 \cdot \left(x - \dfrac{L}{2}\right), \\[3mm] A = -2\dfrac{\pi\tau}{(0.9L\rho_1)}, & 0 \le x, y \le L. \end{array} \right\} \tag{2.3}$$

The notation and the values of other parameters are listed in table 1, and readers can refer to HSSB20 for the model details.

The velocity $\boldsymbol{u}_l = (u_l, v_l)^T$ is given by

$$u_l = -\frac{\partial \psi_l}{\partial y}, \quad v_l = \frac{\partial \psi_l}{\partial x}. \tag{2.4a,b}$$

On the lateral boundaries, we use a partial-slip boundary condition given by

$$\alpha \frac{\partial^2 \psi_l}{\partial n^2} - \frac{\partial \psi_l}{\partial n} = 0, \tag{2.5}$$

where $\alpha = 120$ km.

Equations (2.1)–(2.4a,b) are solved using the CABARET scheme (Karabasov, Berloff & Goloviznin 2009) on a uniform Cartesian grid of size $G = 1025 \times 1025$ with grid spacing $\Delta_x = \Delta_y = 3.75$ km. The potential vorticities are saved at each grid point every day for 183 days.

2.2. *Tracer inversion and the $\boldsymbol{K}_f$ tensor*

In this section, we introduce the scale-aware spatial filtering to decompose the tracer equation and describe the inversion method by which we can obtain a K-tensor for quantifying the eddy tracer flux. Consider the advection–diffusion equation on a fine grid:

$$\frac{\partial C}{\partial t} + \boldsymbol{\nabla} \cdot (\boldsymbol{u}C) = \nu\nabla^2 C + F, \tag{2.6}$$

where C is the tracer advected by the non-divergent flow $\boldsymbol{u}$, F represents tracer sources or sinks and ν is the small-scale diffusivity. The tracer evolution equation has the same

form in all three layers, so we have omitted the layer subscript for brevity. A common way to obtain a governing equation for the large-scale tracer is to decompose the velocity and tracer fields into large-scale/small-scale components by a localised filtering:

$$\boldsymbol{u} = \bar{\boldsymbol{u}} + \boldsymbol{u}', \tag{2.7}$$

$$C = \bar{C} + C'. \tag{2.8}$$

Substituting the decomposed quantities into (2.6) and filtering the equation on the large scales in space yields:

$$\frac{\partial \bar{C}}{\partial t} + \boldsymbol{\nabla} \cdot (\bar{\boldsymbol{u}}\bar{C}) + \boldsymbol{\nabla} \cdot (\overline{\bar{\boldsymbol{u}}\bar{C} + \bar{\boldsymbol{u}}C' + \boldsymbol{u}'\bar{C} + \boldsymbol{u}'C'} - \bar{\boldsymbol{u}}\bar{C}) = \nu\nabla^2\bar{C} + \bar{F}, \tag{2.9}$$

where the eddy terms $\partial C'/\partial t$, $\nu\nabla^2 C'$ and F' are filtered out. Here the filtering is assumed to commute with the divergence operator. To model (2.9) on a coarse grid, the small-scale variations under the second divergence operator need to be parametrised by some large-scale quantities. We denote the divergent term as the eddy forcing of C,

$$E^{(1)} := \boldsymbol{\nabla} \cdot (\overline{\bar{\boldsymbol{u}}\bar{C} + \bar{\boldsymbol{u}}C' + \boldsymbol{u}'\bar{C} + \boldsymbol{u}'C'} - \bar{\boldsymbol{u}}\bar{C}), \tag{2.10}$$

and label (2.10) as the first expression of the eddy forcing. Here, $E^{(1)}$ can not be directly parametrised by the flux-gradient relation because it is not in the form of the divergence of an eddy flux. Additionally, the physical meaning of this expression is unclear. The issue can be solved by choosing a suitable filtering such as the Reynolds averaging. However, the Reynolds averaging is not suitable for our study because we want to consider the full spatio-temporal evolution of the eddy tracer flux, and the interactions between the large-scale and eddy flows. Therefore, instead of the mean-fluctuation decomposition in time, we determine the large/small scales of a variable by applying a spatial decomposition. This method of eddy filtering is also motivated by the practicality of eddy parametrisations in ocean models.

Consider a variable $A(t, \boldsymbol{x})$ on a Cartesian grid G. A running-average spatial filter G^s is applied to its instantaneous fields. The large-scale component $\bar{A}_{ij}$ at a grid point G_{ij} is defined as the mean of A evaluated over the region covered by the filter G^s centred at this point. That is,

$$\bar{A}_{ij}(t) := \frac{1}{d^2} \sum_{k=i-r}^{i+r} \sum_{p=j-r}^{j+r} A_{kp}(t), \quad r = \frac{d-1}{2}, \tag{2.11}$$

where d is the discrete side length of the filter, which is set to be odd to avoid any interpolation. Dropping the grid point indices, the small-scale part of $A(t, \boldsymbol{x})$ is $A'(t, \boldsymbol{x}) = A(t, \boldsymbol{x}) - \bar{A}(t, \boldsymbol{x})$. For grid points near the boundaries, where the distance between the grid point and the nearest boundary is $d_b < d$, we set $d = d_b$. In this case, the square shape of the filter is maintained, and A' is zero on the boundaries. Note, in general for a spatial filter, $\bar{\bar{A}} \neq \bar{A}$ and $\overline{A'} \neq 0$, which is not the case with Reynolds averaging.

Using the spatial filter, decomposing the terms in the tracer equation (2.6) yields

$$\frac{\partial \bar{C}}{\partial t} - \nu\nabla^2\bar{C} - \bar{F} = -E^{(2)}, \tag{2.12}$$

where $E^{(2)}$ is the equivalent eddy forcing

$$E^{(2)} := \boldsymbol{\nabla} \cdot (\bar{\boldsymbol{u}}C' + \boldsymbol{u}'\bar{C} + \boldsymbol{u}'C') + \frac{\partial C'}{\partial t} - \nu\nabla^2 C' - F'. \tag{2.13}$$

Unlike in the Reynolds decomposition approach, we do not filter/average the tracer equation. The physical meaning of each eddy term is clear in $E^{(2)}$. The advection operators for $\bar{u}$ and u' are preserved in the divergence term, so we define the corresponding eddy tracer flux as

$$\boldsymbol{f} := \bar{\boldsymbol{u}}C' + \boldsymbol{u}'\bar{C} + \boldsymbol{u}'C'. \tag{2.14}$$

The eddy tendency $\partial C'/\partial t$, the explicit diffusion $\nu\nabla^2 C'$ and the forcing F' are each present in $E^{(2)}$, but are filtered out when deriving $E^{(1)}$. These non-advective terms contribute to the eddy forcing, and thus need to be parametrised. We will discuss this in more detail in § 2.4.

We now present the inversion method for obtaining a transport tensor for $\boldsymbol{f}$. Because $\boldsymbol{u}'$ and C' are zero on the boundaries, there is no flux through and along the boundaries,

$$\boldsymbol{f}_{\partial\Omega} = 0. \tag{2.15}$$

The transport tensor $\boldsymbol{K}_f$ can then be locally estimated from the flux-gradient relation

$$\boldsymbol{f} = -\boldsymbol{K}_f \cdot \nabla\bar{C}, \tag{2.16}$$

where

$$\boldsymbol{K}_f(t, \boldsymbol{x}) = \begin{bmatrix} K_{11}(t, \boldsymbol{x}) & K_{12}(t, \boldsymbol{x}) \\ K_{21}(t, \boldsymbol{x}) & K_{22}(t, \boldsymbol{x}) \end{bmatrix}. \tag{2.17}$$

Because the system (2.16), with four unknowns and two equations, is under-determined for a single tracer, we use two tracers C^p and C^q. This gives

$$\left. \begin{cases} \boldsymbol{f}^p = -\boldsymbol{K}_f \cdot \nabla\bar{C}^p, \\ \boldsymbol{f}^q = -\boldsymbol{K}_f \cdot \nabla\bar{C}^q \end{cases} \right\}, \tag{2.18}$$

where $\boldsymbol{f}^{p,q} = (f_u^{p,q}, f_v^{p,q})^T$ is the eddy flux for tracer $C^{p,q}$. Then, the tensor is obtained by inverting (2.18):

$$\boldsymbol{K}_f(C^p, C^q) = \frac{1}{\partial_x\bar{C}^p\partial_y\bar{C}^q - \partial_y\bar{C}^p\partial_x\bar{C}^q} \begin{bmatrix} f_u^p & f_u^q \\ f_v^p & f_v^q \end{bmatrix} \begin{bmatrix} -\partial_y\bar{C}^q & \partial_x\bar{C}^q \\ \partial_y\bar{C}^p & -\partial_x\bar{C}^p \end{bmatrix}. \tag{2.19}$$

To summarise, here we have derived the eddy forcing $E^{(2)}$ for a spatial filtering decomposition, which we show is distinct from the eddy forcing obtained with the Reynolds averaging. The eddy forcing $E^{(2)}$ is contributed to by an eddy tracer flux divergence and non-advective terms. We have presented a method for computing the transport tensor for eddy tracer fluxes for a pair of tracers.

2.3. *Helmholtz flux decomposition and the $\boldsymbol{K}_{div}$ tensor*

In the previous section, we presented the method for computing the K-tensor ($\boldsymbol{K}_f$) for a pair of eddy tracer fluxes. In this section, we describe the Helmholtz decomposition which is necessary to compute the K-tensor ($\boldsymbol{K}_{div}$) for the divergent part of two eddy tracer fluxes. We are only interested in these purely divergent fluxes as the rotational part does not contribute to the tracer dynamics, as shown in HSSB20.

For an eddy tracer flux $\boldsymbol{f}$, its Helmholtz decomposition is

$$\boldsymbol{f} = \nabla\Phi_f + \nabla \times \Psi + \boldsymbol{H}. \tag{2.20}$$

Here, $\nabla\Phi_f$ is the divergent flux, $\nabla \times \Psi$ is the rotational flux and $\boldsymbol{H}$ is the non-divergent and irrotational gauge term (Maddison, Marshall & Shipton 2015). The decomposition can

be achieved by solving the Poisson equations of Φ and Ψ (Lau & Wallace 1979; Roberts & Marshall 2000; Maddison *et al.* 2015; Mak *et al.* 2016), i.e.

$$\nabla \cdot f = \nabla^2 \Phi_f, \tag{2.21}$$

$$\nabla \times f = \nabla^2 \Psi. \tag{2.22}$$

The decomposition is not, however, unique in a bounded or singly periodic domain owing to a dependence on boundary conditions (Fox-Kemper, Ferrari & Pedlosky 2003; Maddison *et al.* 2015). For example, although the total flux through all solid boundaries must be zero, the rotational and divergent components separately need not be zero on the boundaries. Maddison *et al.* (2015) provided a decomposition by using the example of a three-dimensional eddy PV flux: the horizontal divergent component was minimised by introducing an eddy force function Φ_e with a zero tangential boundary condition. We adopt this boundary condition so that the rotational component can be removed as much as possible. We let the boundary condition for Φ_f be

$$\Phi_{f \, \partial\Omega} = 0, \tag{2.23}$$

because there is no flux on the boundaries, as in (2.15). Then, the Poisson (2.21) can be solved uniquely, and $\nabla\Phi$ is minimised (Maddison *et al.* 2015). We define the divergent component of the eddy flux as

$$f_{div} = \nabla\Phi_f. \tag{2.24}$$

Then, the total eddy flux is conveniently split into the divergent f_{div} and the non-divergent f_{rot} parts:

$$f = f_{div} + f_{rot}. \tag{2.25}$$

The non-divergent flux is obtained by subtracting f_{div} from f, and it contains the rotational flux $\nabla \times \Psi$ and the gauge term H. We denote it as f_{rot} and do not extract the gauge term because the gauge term does not affect the interpretation of the rotational component. We will call it the rotational flux for simplicity.

The divergent tensor $\boldsymbol{K}_{div}$ is associated with the divergent flux via the flux-gradient relation, i.e.

$$f_{div} = -\boldsymbol{K}_{div} \cdot \nabla\bar{C}. \tag{2.26}$$

We consider it as part of the transport tensor $\boldsymbol{K}_f$, denoting the part corresponding to f_{rot} as $\boldsymbol{K}_{rot}$.

The divergent tensor, and other K-tensors, can be separated into a symmetric component,

$$\boldsymbol{S} = \begin{bmatrix} K_{div,11} & S_{12} \\ S_{12} & K_{div,22} \end{bmatrix}, \quad S_{12} = (K_{div,12} + K_{div,21})/2, \tag{2.27a,b}$$

and an antisymmetric component,

$$\boldsymbol{A} = \begin{bmatrix} 0 & A_{12} \\ -A_{12} & 0 \end{bmatrix}, \quad A_{12} = (K_{div,12} - K_{div,21})/2. \tag{2.28a,b}$$

We refer to $\boldsymbol{S}$ as the *diffusion tensor* because its eigenvalues λ_1, λ_2 are the diffusivities in the directions of their eigenvectors (Plumb & Mahlman 1987). Here, λ_1 is assigned to the eigenvalue with the most positive value such that $\lambda_1 \geq \lambda_2$. The method of computing the eigenvalues is specifically described in HSSB20. We refer to $\boldsymbol{A}$ as the advection

tensor, where the component A_{12} quantifies the advection of the large-scale tracer field in the direction perpendicular to the tracer gradient. Our study only focuses on the tracer-dependence of the eigenvalues and does not discuss the interpretation of $\boldsymbol{S}$ and $\boldsymbol{A}$.

2.4. *Non-advective eddy terms and the $\boldsymbol{K}_g$ tensor*

In the Reynolds decomposition approach, non-advective eddy terms do not contribute to the mean eddy transport. On the non-advective terms in the spatial filter approach (2.13), the explicit eddy diffusion is negligible, but the eddy tendency and the external forcing terms make leading-order contributions to the eddy forcing. This motivates incorporating their effects into a K-tensor ($\boldsymbol{K}$) that parametrises all eddy effects. To do this, we use the fact that the non-advective eddy terms can be quantified using a purely divergent eddy flux $\boldsymbol{g}$, where

$$\nabla \cdot \boldsymbol{g} := \frac{\partial C'}{\partial t} - \nu \nabla^2 C' - F'. \tag{2.29}$$

As with the divergent flux $\boldsymbol{f}_{div}$, we obtain $\boldsymbol{g}$ by inverting its corresponding Poisson equation with zero normal and tangent flux boundary conditions. (This requires choosing a suitable tracer forcing with $F'_{\partial\Omega} = 0$.)

We apply the flux-gradient relation to $\boldsymbol{g}$ and obtain a K-tensor $\boldsymbol{K}_g$ for the non-advective eddy terms. Then the eddy forcing can be expressed in terms of the full tensor $\boldsymbol{K}$,

$$E^{(2)} = \nabla \cdot (-\boldsymbol{K} \cdot \nabla \bar{C}), \tag{2.30}$$

where

$$\boldsymbol{K} := \boldsymbol{K}_f + \boldsymbol{K}_g. \tag{2.31}$$

The tensors $\boldsymbol{K}_f$ and $\boldsymbol{K}_g$ both vary on the small scales, because they are determined by eddy terms. Their sum $\boldsymbol{K}$, however, will not contain such small-scale variability because in (2.30), we relate it to strictly large-scale terms. The $E^{(2)}$ only varies on the large scale as it is equal to the left-hand side of (2.12), which includes only large-scale quantities. As a result, any tracer dependence in the full tensor $\boldsymbol{K}$ cannot be caused by small-scale variability.

2.5. *Numerical experiment set-up*

In the previous subsections, we outlined the method for computing four K-tensors, $\boldsymbol{K}_f$, $\boldsymbol{K}_{div}$, $\boldsymbol{K}_g$ and $\boldsymbol{K} = \boldsymbol{K}_f + \boldsymbol{K}_g$. In this section, we outline the experiments in which we test the non-uniqueness of these tensors and their ability to reproduce their corresponding fluxes.

We prepare two sets of passive tracers with different forms of initialisation for the experiments. For each set, six tracers are chosen to provide a sufficient variety to the initial profiles, and are simulated for half a year. The first set is initialised with linear profiles:

$$C_B^p = \frac{A_c}{\sqrt{a^2 + b^2}}(ax + by + \gamma), \quad p = 1, \ldots, 6, \tag{2.32}$$

where the gradient $(a, b) \in (\mathbb{Z}, \mathbb{Z})$ and the amplitude $A_c \in \mathbb{Z}^+$. The constant γ is set to be positive, so that $C_B^p \geq 0$. The second set is initialised with sinusoidal wave profiles:

$$C_B^p = \frac{1}{n_p} \sum_{n=1}^{n_p} A_c \cos(k_n x + l_n y + \varphi_n), \quad p = 7, \ldots, 12, \tag{2.33}$$

where the number of waves $n_p \in \mathbb{Z}^+$ and the phase angle $\varphi_n \in [0, 2\pi]$. The wavenumber for each sinusoidal wave is set to be $k = |(k_n, l_n)| = 6$, where k_n, l_n are randomly determined. Then the wavelength of the nonlinear C_B^p is approximately the basin size.

For every pair of tracers in a set (15 pairs per set of 6), we carry out a test that computes the transport tensor $\boldsymbol{K}_f$, its divergent counterpart $\boldsymbol{K}_{div}$, the corresponding $\boldsymbol{K}_g$ and consequently the full tensor $\boldsymbol{K}$. To avoid the singularity in (2.19) when diagnosing the transport tensors, it is required that the large-scale gradients of each tracer pair remain misaligned. We maintain this misalignment by setting F to be a restoring force,

$$F = \frac{C - C_B}{T}, \qquad (2.34)$$

with a relaxation time scale of $T = 10$ days. The velocity and the tracer fields are decomposed into large scales and small scales by a spatial filter with $d = 31$, the distance given by which is approximately 112.5 km.

For the set of tracers initialised linearly, we refer to this experiment as RL, and for the set of those initialised with sinusoidal waves, we refer to it as RN. The two experiments RL and RN, as described above, are repeated but with the tracer relaxation switched off, and we denote these experiments as UL and UN. For reference, the letters R, U, L, N stand for relaxed, unrelaxed, linear and nonlinear, respectively. In summary, we have presented the methods to obtain four K-tensors in four different experiments.

3. Results

In this section, we present the results from the four experiments described in §2.5 and focus on the conditions for obtaining unique tensors. We begin by diagnosing the accuracy of K-tensors in §3.1. Then, in §3.2, we present the K-tensors for linearly initialised tracers and discuss the effects of the rotational flux. Finally, in §3.3, we examine the tracer dependence of K-tensors for different tracer sets.

3.1. *Numerical experiment errors*

In this section, we analyse the accuracy of K-tensors in terms of reproducing the eddy fluxes and eddy forcing. First, we measure the accuracy of the transport tensor $\boldsymbol{K}_f$ by reconstructing the tracer flux $\boldsymbol{f}$. The relative error $Err(\boldsymbol{f})$ of the inversion is given by

$$Err(\boldsymbol{f}) := \frac{\|\boldsymbol{f} + \boldsymbol{K}_f \cdot \nabla \bar{C}\|_2}{\|\boldsymbol{f}\|_2} \times 100\,\%. \qquad (3.1)$$

The ensemble-averaged $Err(\boldsymbol{f})$ for the experiments RL, RN, UL and UN are presented in figure 1. Bachman *et al.* (2015) suggested to over-determine the tracer flux-gradient relation to balance the removal of a large number of degrees of freedom owing to Reynolds decomposition. However, our linear system (2.18) is successfully solved with two tracers, with relative errors for all experiments almost always less than 10^{-4}. Therefore, we deduce that using more tracers is not necessary, because the scale-aware decomposition can preserve sufficient spatial and temporal information of the flow. Panel (*a*) compares the spatial averages of $Err(\boldsymbol{f})$. Overall, the errors do not increase in time after 25 days when the total amount of fluxes have reached equilibrium, and the error amplitudes are likely determined by the tracer initial conditions. Comparing the results from experiments RL and RN, we find that $\boldsymbol{K}_f$ for linearly initialised tracers gives an order of magnitude better reconstruction, compared with the case of nonlinear tracers. The effect of relaxation on

the accuracy of the method is investigated using the nonlinear tracers set – the error for the unrelaxed tracers is larger than that for the relaxed tracers. The large peak in panel (*a*) arises from an instantaneous near-alignment of the large-scale tracer gradients. Our results confirm that such alignment can be restrained if tracers are weakly relaxed to their initial profiles.

The spatial distributions of $Err(f)$ are shown in panels (*b*), (*c*), (*d*) and (*e*) of figure 1. When the tracers are relaxed in experiments RL and RN, the relative errors are more pronounced in the jet region because the energetic flow accelerates alignment of the tracer gradients. Some additional curves of large errors are found for tracers with nonlinear initial profiles. As their locations correlate with the distribution of the gradient fields $\nabla \bar{C}$, we argue a steep change in the tracer gradient arising from the random combination of cosine waves can cause a large numerical error. As shown in panels (*d*) and (*e*), for the case of unrelaxed tracer profiles, the error increases globally (except in the south-east corner), thus, in agreement with results shown in panel (*a*).

We now measure the accuracy of the divergent tensor $\boldsymbol{K}_{div}$, noting that the Helmholtz decomposition can introduce large numerical errors (Bachman *et al.* 2015). To estimate these errors, we reconstruct the flux divergence by using $\boldsymbol{K}_f$ and $\boldsymbol{K}_{div}$, and then compare the relative errors:

$$Err(\nabla \cdot f) := \frac{|\nabla \cdot f + \nabla \cdot (\boldsymbol{K}^* \cdot \nabla \bar{C})|}{|\nabla \cdot f|} \times 100\,\%, \qquad (3.2)$$

where $\boldsymbol{K}^* \in \{\boldsymbol{K}_f, \boldsymbol{K}_{div}\}$. Figure 2 shows the fields of $\nabla \cdot f$ reconstructed by $\boldsymbol{K}_f$ and $\boldsymbol{K}_{div}$ for a single test in experiment RN. It also shows histograms of the range of the relative errors. From the panels in the first two columns, we see that both tensors are able to reproduce the flux divergence in all layers. The error distributions are roughly Gaussian, as illustrated in panels (*c*), (*f*) and (*i*). Even though the divergence fields obtained via the two tensors are visually indistinguishable, the errors associated with $\boldsymbol{K}_{div}$ are, on average, three orders of magnitude larger than those for $\boldsymbol{K}_f$ (10^{-1} versus 10^2). These increased numerical errors are injected by the divergence operators during the calculation of $\nabla \cdot f$ in (2.21) and $\nabla \Phi$ in (2.25), similar to what was observed by Bachman *et al.* (2015). Specifically, we hypothesise that the second-order finite-difference method used in the Helmholtz decomposition can shift 'correct' values to neighbouring grid points, which leads to large point-wise errors. This can explain why the two reconstructed flux divergences are visually indistinguishable, but have average errors three orders of magnitude apart. Because any tensor needs to be coarse-grained before being applied to a coarse-resolution model, it is unclear whether or not use of $\boldsymbol{K}_f$ over $\boldsymbol{K}_{div}$ would lead to a more accurate parametrisation, despite the large errors associated with $\boldsymbol{K}_{div}$.

Finally, we analyse the accuracy of $\boldsymbol{K}$ and $\boldsymbol{K}_g$ by measuring the reconstruction error of the eddy forcing E. (Hereafter, we drop the superscript of $E^{(2)}$ and denote the eddy forcing as just E.) figure 3 shows the temporally averaged E and the relative errors $Err(E)$ for $\boldsymbol{K}_f$ and $\boldsymbol{K}$. Even though $\nabla \cdot f$ dominates E in the top layer by accounting for 73 % of the total, in the intermediate and bottom layers, the non-advective eddy terms become comparably large by accounting for 45 % of the total. This implies that the non-advective eddy terms are not negligible and the associated $\boldsymbol{K}_g$ should be included in the reconstruction of the eddy forcing. We illustrate this by comparing $Err(E)$ given by $\boldsymbol{K}_f$ and $\boldsymbol{K}$. The latter, being of order 10^1, lies between $Err(\nabla \cdot f)$ for $\boldsymbol{K}_f$ and $\boldsymbol{K}_{div}$, so it is purely numerical. This is because the divergence operator is used once during the computation of g, unlike zero times or twice for the other two cases. However, the former is on average one order of magnitude larger owing to the absence of the non-advective eddy terms. Hence, we deduce
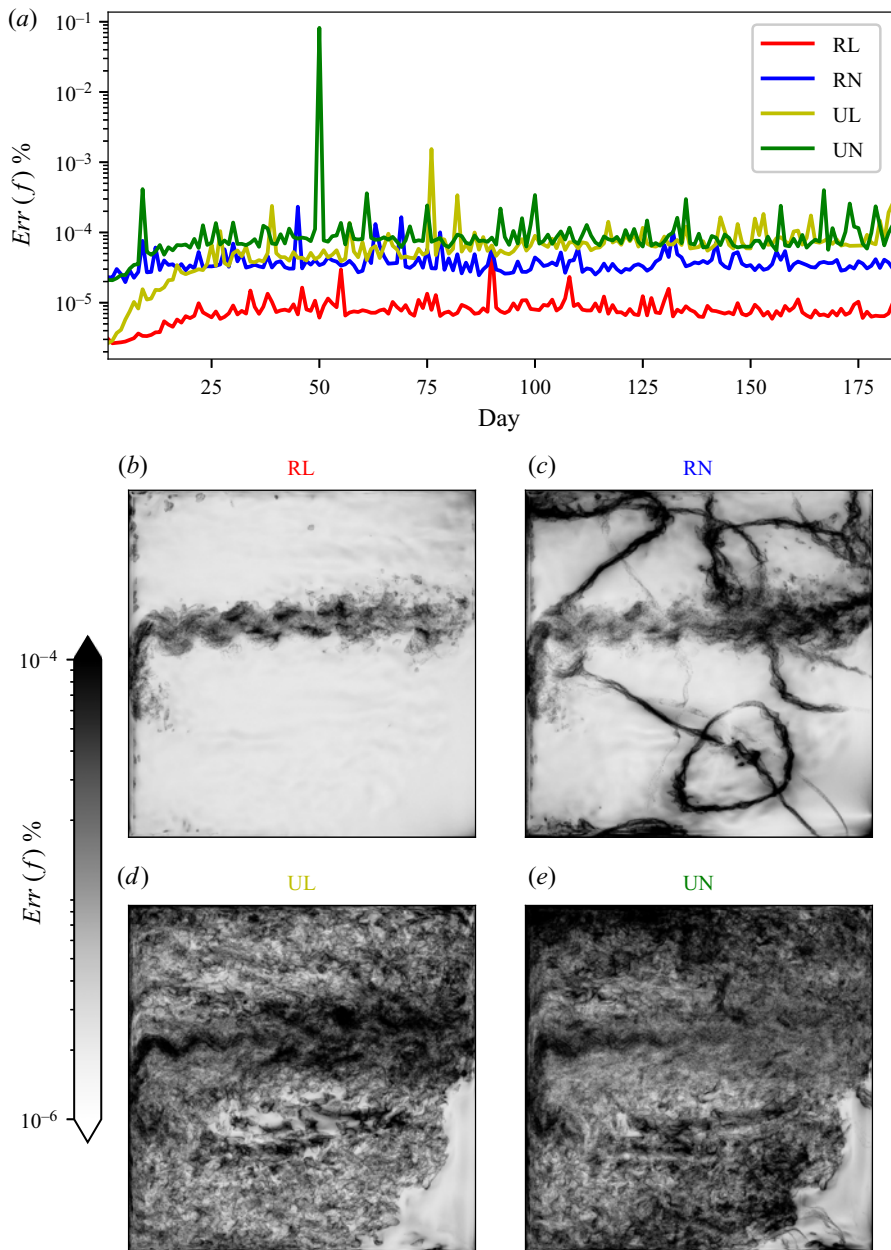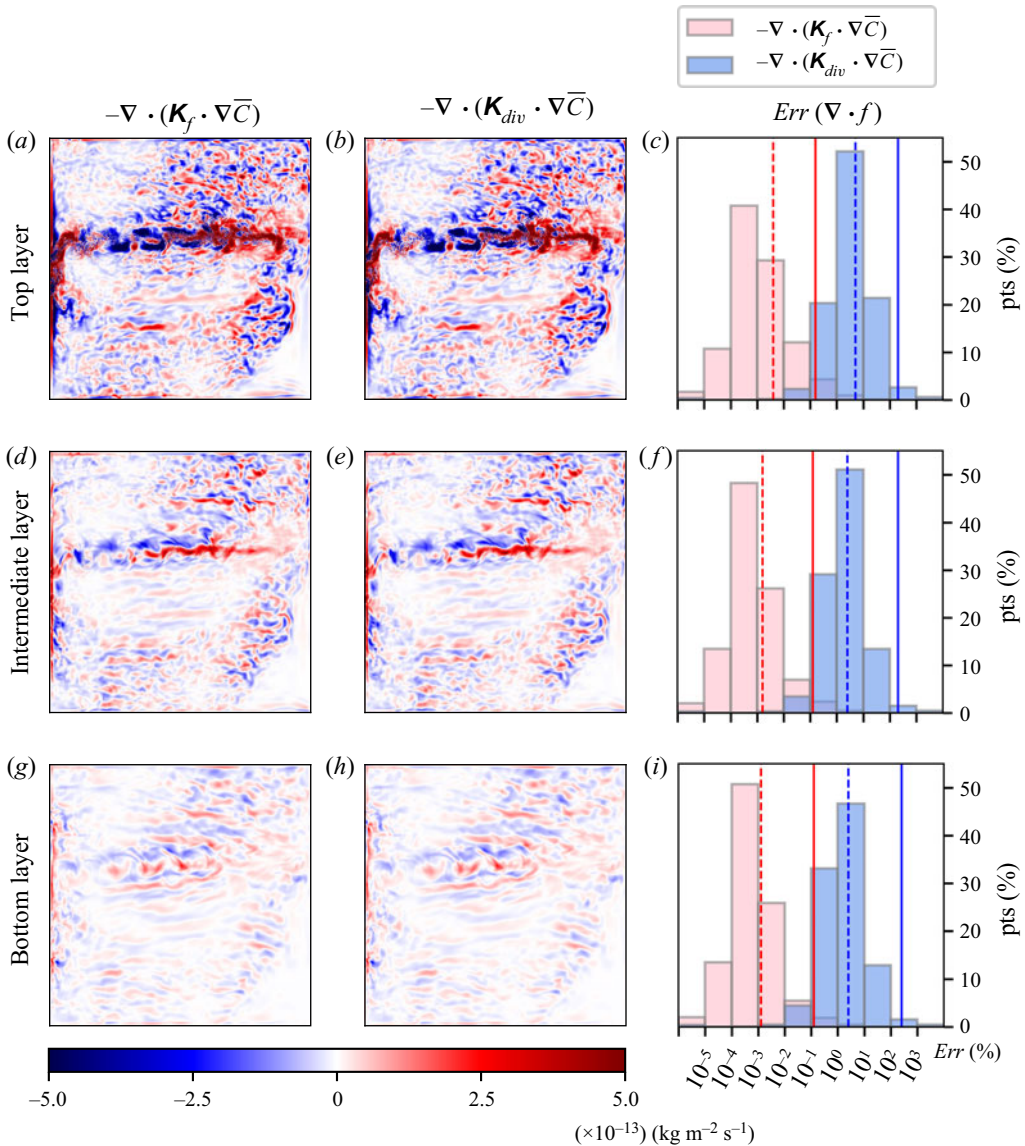
Figure 1. Ensemble averages of the relative error $Err(f)$ for experiments RL, RN, UL and UN in the top layer. The spatially averaged errors are shown in panel (a), and the temporally averaged over half a year errors are in panels (b,c,d,e). Logarithmic scales are used for all panels and the colour bar is shared among (b,c,d,e). The figure shows that the fluxes are successfully reconstructed by the transport tensor $\boldsymbol{K}_f$, for tracers with both linear and nonlinear initial profiles, with and without the relaxation.

that it is not sufficient to reproduce the eddy forcing by $\boldsymbol{K}_f$ but is sufficient by the full tensor $\boldsymbol{K}$.

Overall, the results show that K-tensors computed by the inversion method can successfully reproduce the eddy tracer flux, its divergence and then the eddy forcing.

Figure 2. Temporal averages of the flux divergence associated with $\boldsymbol{K}_f$ (a,d,g) and $\boldsymbol{K}_{div}$ (b,e,h) in the three layers. (c,f,i) The histograms of the relative errors $Err(\nabla \cdot \boldsymbol{f})$ together with the spatial averages marked by the vertical lines. The solid vertical lines represent the domain-mean errors, while the dashed lines only include the 95 % of grid points with the smallest errors. Transparent red and blue bars represent $\boldsymbol{K}_f$ and $\boldsymbol{K}_{div}$, respectively. These results are from experiment RN. It is shown that both tensors reconstruct the flux divergence well, but the latter introduces much larger numerical errors. The same observations are made for linear cases.

Although the transport and the divergent tensors lead to visually indistinguishable flux convergences, the Helmholtz decomposition introduces non-negligible errors in the latter that need to be addressed. As the non-advective eddy terms make considerable contributions to the eddy forcing, the corresponding K-tensor needs to be included in the reconstruction.
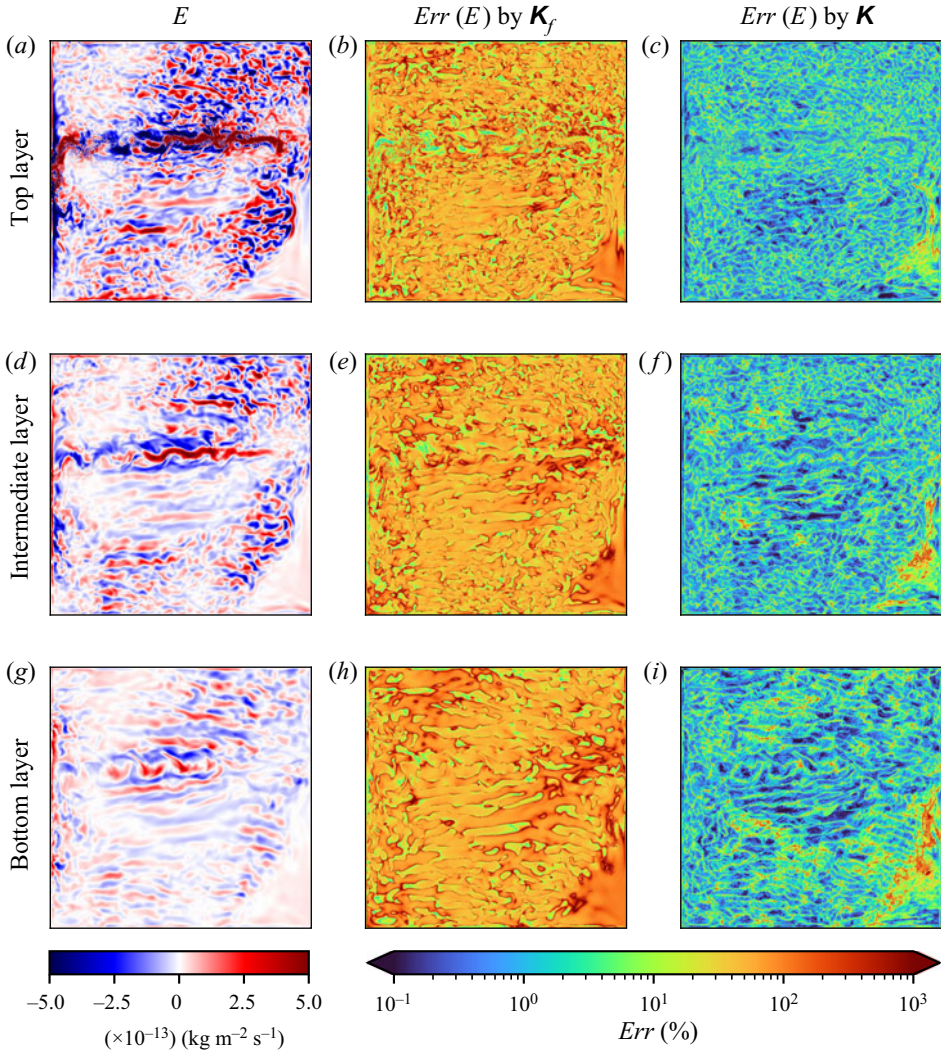
Figure 3. (*a,d,g*) Time-mean fields of the eddy forcing E in the three layers. The relative errors of their reconstructions by $\boldsymbol{K}_f$ (*b,e,h*) and $\boldsymbol{K}$ (*c,f,i*). They are in the range of 10–300 % and 0.1–10 % at 90 % grid points, respectively.

3.2. *K-tensors*

In this section, we diagnose the representative properties of K-tensors and discuss the influence of the rotational eddy flux on tensor components. Because $\boldsymbol{K}_g$ is given by a purely divergent flux $\boldsymbol{g}$, it is sufficient to discuss the influence of the rotational flux by using $\boldsymbol{K}$ instead of $\boldsymbol{K}_f$.

Figure 4 presents the time-mean entries of $\boldsymbol{K}$, $\boldsymbol{K}_{div}$ and $\boldsymbol{K}_g$ in the top layer for a tracer pair in experiment RL. The magnitudes of their components for the three layers are listed in tables 2–4. The inclusion/exclusion of the rotational flux leads to notable distinctions between $\boldsymbol{K}$ and $\boldsymbol{K}_{div}$. A comparison of them between tables 2 and 3 shows that the former is approximately two orders of magnitude larger than the latter in all layers, which means that the eddy transport given by the dynamically active flux $\boldsymbol{f}_{div}$ is only 1 % of the total.
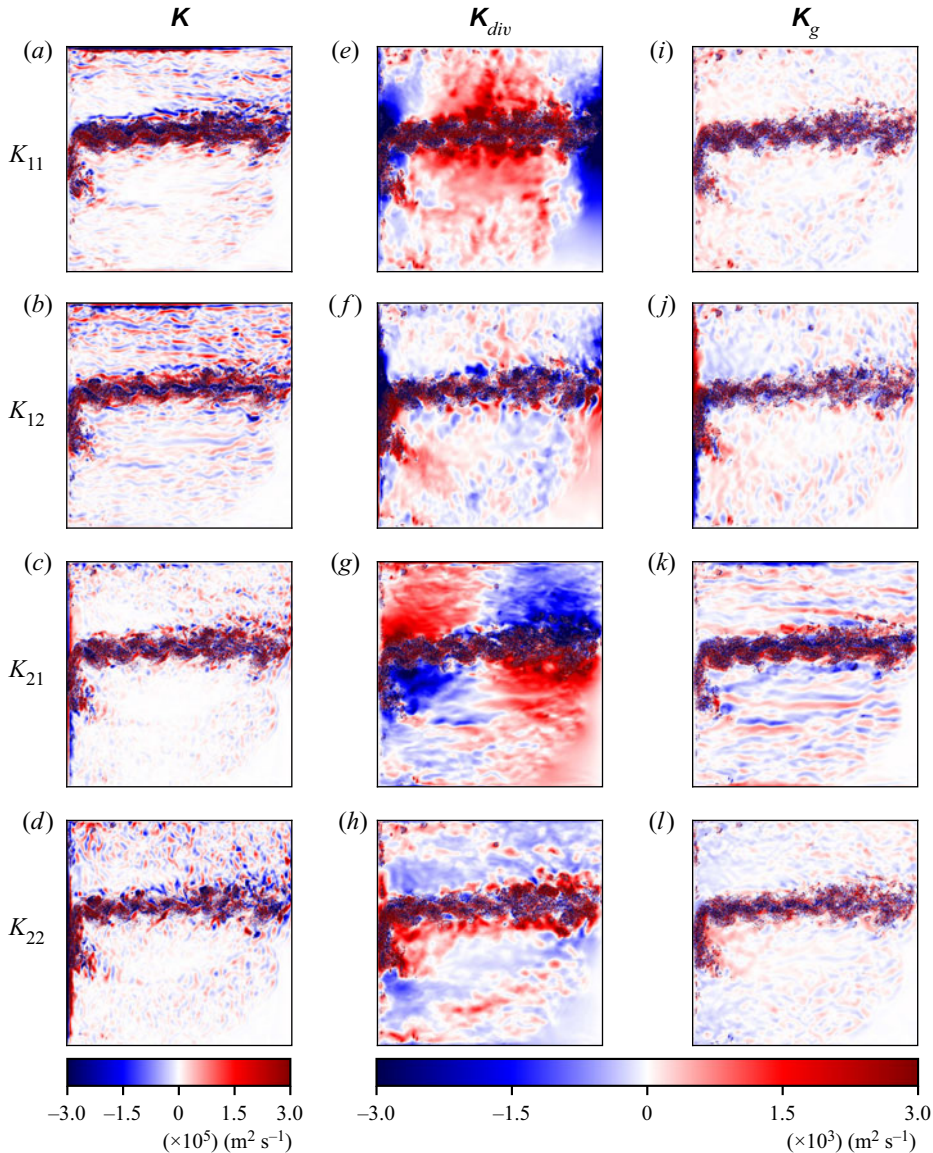
Figure 4. Temporally averaged $\boldsymbol{K}$ (*a–d*), $\boldsymbol{K}_{div}$ (*e–h*) and $\boldsymbol{K}_g$ (*j–l*), all evaluated over half a year in the top layer for a test in experiment RL. The values of $\boldsymbol{K}$ (as well as the indistinguishable $\boldsymbol{K}_f$) are found to be generally two orders of magnitudes larger than those for $\boldsymbol{K}_{div}$ and $\boldsymbol{K}_g$. The magnitudes of the tensor components in the three layers are listed in tables 2–4.

Additionally, from $\boldsymbol{K}_{div}$, we see that diffusion outside the energetic jet region is relatively strong, that is, in comparison to the disparity inside and outside the jet region with $\boldsymbol{K}_f$.

It is worth comparing $\boldsymbol{K}_{div}$ and $\boldsymbol{K}_g$, as both are computed for divergent fluxes. They have the same orders of magnitudes, but $\boldsymbol{K}_{div}$ is approximately three times larger than $\boldsymbol{K}_g$. Additionally, the spatial patterns of their components are very different. We further illustrate the differences by presenting the velocity potentials of $\boldsymbol{f}$ and $\boldsymbol{g}$ in the three layers in figure 5. As expected, ϕ_g shows more pronounced small-scale patterns in the subtropical regions. However, it also generates mesoscale components along the jet, similarly to Φ_f,

$|K|_{\bar{\Omega}}$ (m² s⁻¹)	K_{11}	K_{12}	K_{21}	K_{22}
Top layer	459799.8	233578.1	278369.1	262450.6
Middle layer	71764.2	29132.4	13097.3	21073.9
Bottom layer	4200.7	7831.1	2952.3	5470.1

Table 2. Magnitudes of **K** shown in figure 4 in three layers. The components K_{11}, K_{12}, K_{21}, K_{22} have the same order of magnitude in each layer, and their values are roughly reduced by one order of magnitude in the layer below.

$|K|_{\bar{\Omega}}$ (m² s⁻¹)	K_{11}	K_{12}	K_{21}	K_{22}
Top layer	5169.1	3592.1	4863.8	3019.4
Middle layer	317.6	282.6	425.7	329.8
Bottom layer	72.7	64.2	106.4	87.7

Table 3. Magnitudes of $\boldsymbol{K}_{div}$ shown in figure 4 in three layers. The components K_{11}, K_{12}, K_{21}, K_{22} have the same order of magnitude in each layer, and their values are roughly reduced by one order of magnitude in the layer below.

$|K|_{\bar{\Omega}}$ (m² s⁻¹)	K_{11}	K_{12}	K_{21}	K_{22}
Top layer	1954.1	1630.5	3040.8	2065.7
Middle layer	176.2	236.5	322.1	170.0
Bottom layer	35.4	52.5	94.3	36.1

Table 4. Magnitudes of $\boldsymbol{K}_g$ shown in figure 4 in three layers. The components K_{11}, K_{12}, K_{21}, K_{22} have the same order of magnitude in each layer, and their values are roughly reduced by one order of magnitude in the layer below.

thus, contradicting the assumption that the non-advective terms do not contribute as much to the mean flow and thus can be filtered out. The sign of Φ_g is distributed differently from that of Φ_f. Their distributions are roughly perpendicular in the top layer and are the opposite in the lower two layers. Overall, the results presented highlight the necessity of interpreting the velocity potential by depth, but this is beyond the scope of this study.

Despite the distinctions in the magnitudes and spatial patterns, negative values appear in the fields of $K_{11,12,21,22}$ for all tensors. They are more pronounced in $\boldsymbol{K}_{div}$, but also exist in $\boldsymbol{K}_g$, which naturally has no relation with the rotational flux. Therefore, the removal of the rotational flux does not eliminate the negative values in the tensor components. Furthermore, the second eigenvalues of the symmetric parts of $\boldsymbol{K}_f$ and $\boldsymbol{K}_{div}$ are most often negative (HSSB20).

To summarise, we have presented $\boldsymbol{K}$, $\boldsymbol{K}_{div}$ and $\boldsymbol{K}_g$, and the velocity potentials Φ_f and Φ_g. Owing to the dominance of the rotational flux $\boldsymbol{f}_{rot}$, there are major differences in the magnitudes and spatial structures of $\boldsymbol{K}$ when compared with the other two tensors. The inclusion of $\boldsymbol{f}_{rot}$ is not responsible for the negative values or negative diffusion eigenvalues (HSSB20), as these persist for all K-tensors.

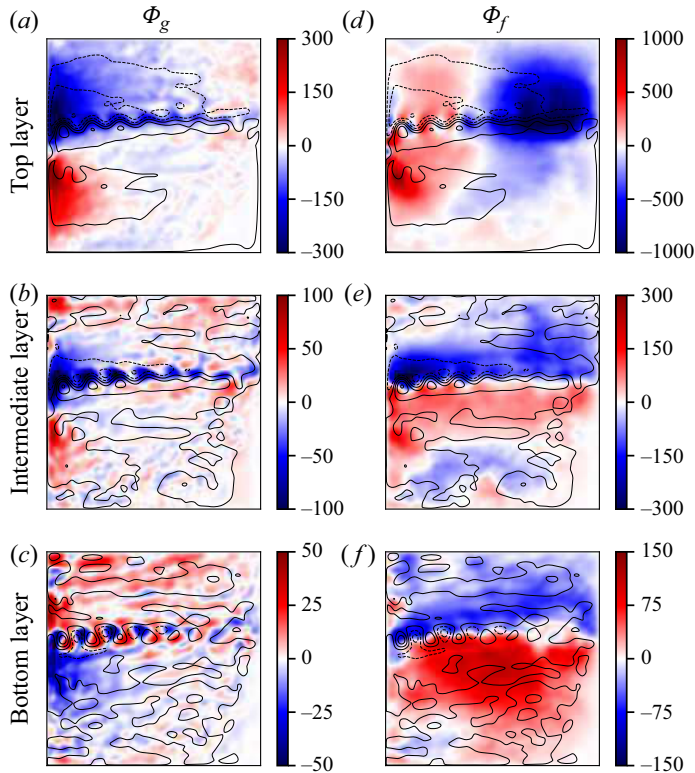Figure 5. Temporally averaged velocity potentials of divergent fluxes $\boldsymbol{g}$ $(a–c)$ and $\boldsymbol{f}_{div}$ $(d–f)$ in the three layers. Units for all panels are $\times 10^2$ m^2 s^{-1}. The black solid/dashed lines represent the time-mean streamlines of the large-scale $\bar{\psi}$ with positive/negative values.

3.3. *Non-uniqueness of diffusivity*

In this section, we investigate the tracer-dependence of K-tensors. We start by obtaining tracer-independent tensors for the linearly initialised tracers by exploiting the linearity of the flux-gradient relation. We then measure their relative distances to use as a reference for purely numerical discretisation errors. Next, we analyse the tracer-dependence by comparing this reference case with the relative distances of the K-tensors corresponding to the nonlinearly initialised tracers.

The linearity of the tracer equation and the flux-gradient relation have important consequences for the non-uniqueness of the K-tensor. That is, a tensor obtained for any two tracers, C^p and C^q, is the same as would be obtained for any other tracer that is a linear combination of C^p and C^q. Consider two solutions C^p, C^q of the tracer (2.6) with initial conditions $C^p(t_0)$, $C^q(t_0)$. Then, any tracer C initialised as $C(t_0) = mC^p(t_0) + nC^q(t_0)$ is a solution in the form

$$C(t) = mC^p(t) + nC^q(t), \tag{3.3}$$

where m and n are arbitrary constants. Because the spatial filter is a linear operator, the large- and small-scale components can also be written as the above linear combination.

920 A32-17

Then, the eddy tracer flux and the large-scale tracer gradient are given by

$$f = mf^p + nf^q, \tag{3.4}$$

$$\nabla \bar{C} = m\nabla \bar{C}^p + n\nabla \bar{C}^q. \tag{3.5}$$

Suppose $\boldsymbol{K}_f$ is the transport tensor obtained with C^p and C^q, then the linearity of the flux-gradient relation gives

$$mf^p + nf^q = -\boldsymbol{K}_f \cdot (m\nabla \bar{C}^p + n\nabla \bar{C}^q), \tag{3.6}$$

and implies that $\boldsymbol{K}_f$ is also a transport tensor for C. Because the same linear combination as in (3.4) can be applied to the non-advective eddy terms, the associated flux divergence can be written as

$$\nabla \cdot \boldsymbol{g} = m\nabla \cdot \boldsymbol{g}^p + n\nabla \cdot \boldsymbol{g}^q. \tag{3.7}$$

Therefore, the same conclusion can be inferred for $\boldsymbol{K}_g$. Overall, any tracer pairs that are linear combinations of C^p and C^q share the same $\boldsymbol{K}_f$ and $\boldsymbol{K}_g$, and, thus, the same full tensor $\boldsymbol{K}$.

It is not necessary to linearly initialise tracers to obtain the same tensor. In fact, the tracers with linear initial profiles do not always have the same tensor. Recall the initialisation of the linear profiles C_B in (2.32). Even though every linear profile has a constant gradient (a, b), it can only be written as a linear combination of other pairs of linear profiles with an additional constant owing to the scalar term γ. This γ is included in the large-scale tracer component and causes the tensors for linearly initialised tracers to be distinct. Even though it does not affect the large-scale gradient in (3.5), the eddy fluxes cannot, in general, be written as the linear combinations of those given by the linearly initialised tracers, as in (3.4), because of an additional flux component $\boldsymbol{u}'\gamma$. Thus, $\boldsymbol{K}_f$ can not be unique in our linear tracer experiments, as part of the tensor component $\boldsymbol{K}_\gamma$ depends on the scalar constant,

$$\boldsymbol{u}'\gamma = -\boldsymbol{K}_\gamma \cdot \nabla \bar{C}. \tag{3.8}$$

Note that $\boldsymbol{K}_\gamma$ is not, in general, equal to $\boldsymbol{K}_{rot}$, even though $\boldsymbol{u}'\gamma$ only contributes to $\boldsymbol{f}_{rot}$. Nonetheless, the transport tensor for initially linear tracers cannot be tracer-independent or unique, unless the rotational flux is removed. However, $\boldsymbol{K}_g$ is always unique, because it only depends on C'.

To verify this, we measure the relative distance between respective tensor components. That is, for a pair of components (denoted with superscripts p or q), the relative distance between them is

$$d_r(K^p, K^q) = \frac{1}{N_\Omega} \sum_\Omega \frac{|K^p - K^q|}{|\langle K \rangle|}, \tag{3.9}$$

where $K^{p,q} \in \{K_{11}, K_{12}, K_{21}, K_{22}\}$, and the ensemble-average $\langle K \rangle$ is the global reference. For each of $\boldsymbol{K}_f$, $\boldsymbol{K}_{div}$ and $\boldsymbol{K}_g$, we have 15 ensembles from RL, so in total we can measure the relative distances for 105 pairs. The standard deviation is not used as a measure owing to the large differences in magnitudes among K-tensors.

Figure 6 presents the relative distances for $\boldsymbol{K}_f$ and $\boldsymbol{K}_{div}$. We see that d_r for $\boldsymbol{K}_f$ typically varies between 0.7 and 3, while for $\boldsymbol{K}_{div}$ it is mainly between 0.8 and 1.25. Furthermore, more than 70 % of tensor pairs have larger distances for the former, despite the additional numerical errors induced in the latter. According to the linearity of the flux-gradient relation, $\boldsymbol{K}_{div}$ for the linearly initialised tracers should be unique, because the tracer-dependent components $\boldsymbol{K}_\gamma$ are excluded by removing the rotational fluxes.
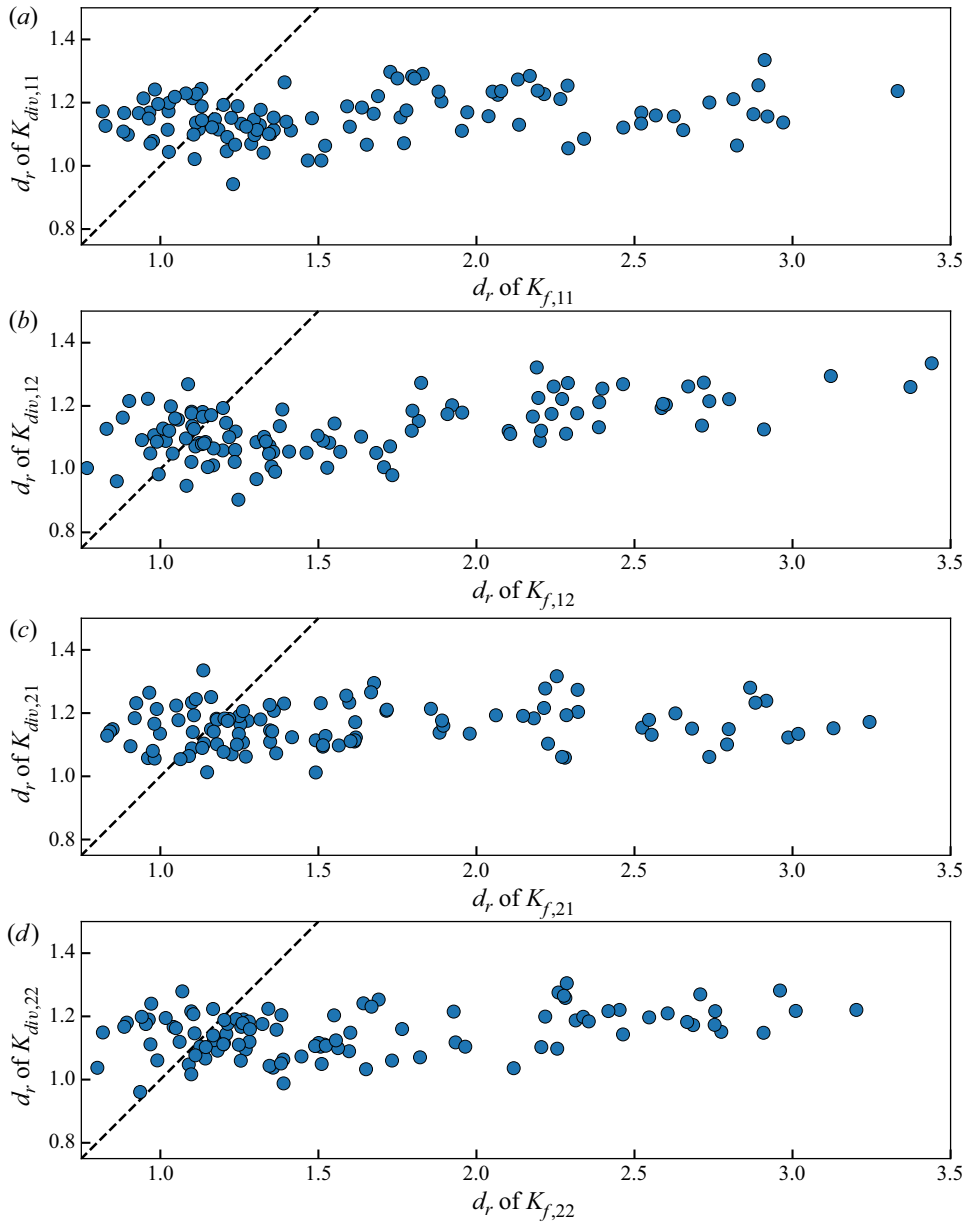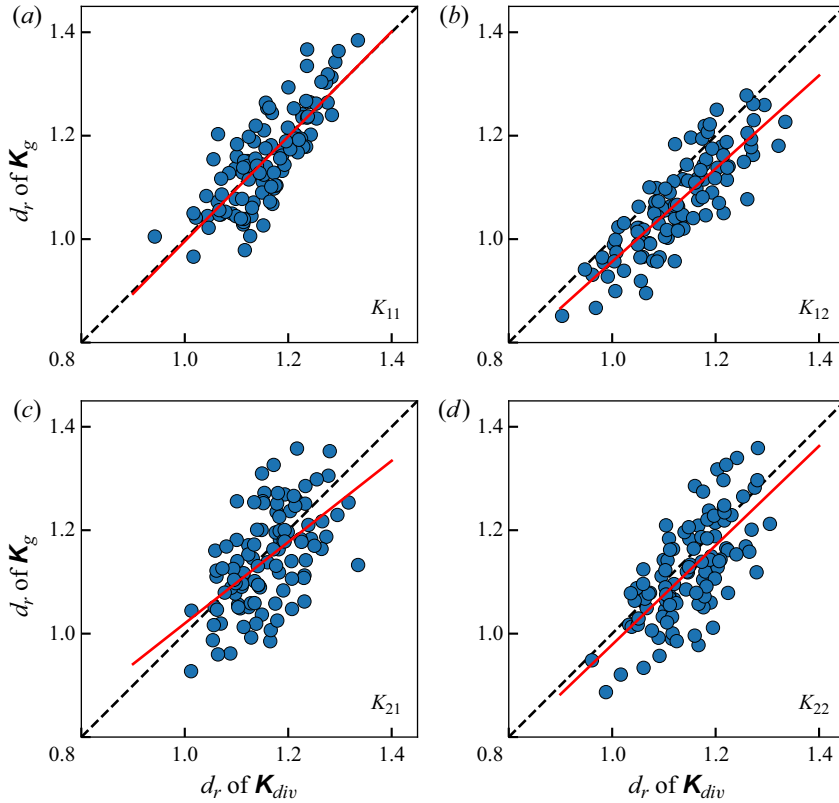
Figure 6. The relative distance d_r of components K_{11}, K_{12}, K_{21}, K_{22} (*a–d*) in the top layer in experiment RL. One dot represents one comparison between $\boldsymbol{K}_f$ and $\boldsymbol{K}_{div}$ given by the same tracer pairs. A total of 105 comparisons are carried out amongst the 15 tracer pairs. The *x*-values are the temporal averages of d_r for $\boldsymbol{K}_f$, while the *y*-values are for $\boldsymbol{K}_{div}$.

The non-zero d_r for $\boldsymbol{K}_{div}$ is attributed to numerical errors in the flux decomposition, as well as to some unavoidable alignments of gradients during the simulation. Moreover, d_r can be uncommonly large for some pairs, because the ensemble-average $\langle K \rangle$ is used as a global reference in (3.9), instead of considering only K^p and K^q. Overall, it is evident that the greater relative distances in $\boldsymbol{K}_f$ arise from the tracer-dependent $\boldsymbol{K}_{rot}$.

Figure 7. The relative distance d_r of components K_{11}, K_{12}, K_{21}, K_{22} (a–d) in the top layer in experiment RL. One dot represents one comparison between $\boldsymbol{K}_{div}$ and $\boldsymbol{K}_g$ given by the same tracer pairs. A total of 105 comparisons are carried out amongst the 15 tracer pairs. The x-values are the temporal averages of d_r for $\boldsymbol{K}_{div}$, while the y-values are for $\boldsymbol{K}_g$. The red solid lines are the linear regression lines, and the corresponding coefficients are 1.01, 0.9, 0.79, 0.99.

Because $\boldsymbol{K}_g$ is not associated with the tracer-dependent γ, it is unique for linearly initialised tracers. This is shown in figure 7, where the relative distances for $\boldsymbol{K}_g$ are in the range similar to those for $\boldsymbol{K}_{div}$. As both tensors are unique, we determine the relation between them by fitting linear regression models to the comparisons for each component (x for $\boldsymbol{K}_{div}$ and y for $\boldsymbol{K}_g$). We find that the regression lines are closely aligned with $x = y$, especially for the diagonal elements K_{11} and K_{22}. Therefore, d_r for the tensors $\boldsymbol{K}_g$ and $\boldsymbol{K}_{div}$ are very similar, with slight differences caused by numerical errors. This rough similarity shows that the non-uniqueness of the tensors is caused by the linearity of the flux-gradient relation and linear tracers, and may not be a property inherent to all transport tensors.

To investigate if the uniqueness holds for all the passive tracers, we repeat the analyses for $\boldsymbol{K}_{div}$ and $\boldsymbol{K}_g$ using the ensembles from RN. In figure 8, we scatter plot d_r for the RN ensemble against d_r for the RL ensemble. Most d_r values for $\boldsymbol{K}_{div}$ from RN are greater than 1.5, and all of them are larger than those from RL. The same is observed for the sum of $\boldsymbol{K}_{div}$ and $\boldsymbol{K}_g$. Because we have shown (analytically) that all linearly initialised tracers generate the same $\boldsymbol{K}_{div}$ and $\boldsymbol{K}_g$, we deduce that they are different and, thus, tracer-dependent, when the tracers are initialised with sinusoidal – or more generally, nonlinear – profiles.

Until now, our conclusions are based on tracers that are relaxed towards their initial profiles. To test these conclusions for freely evolving tracers, in figure 9, we present time
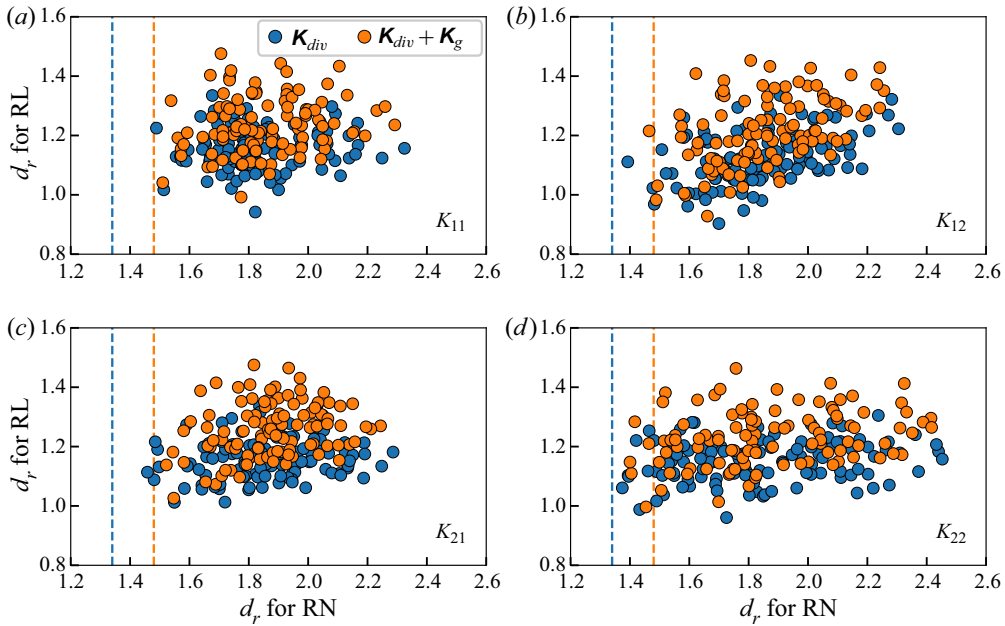
Figure 8. The relative distance d_r of components $K_{11}, K_{12}, K_{21}, K_{22}$ (*a–d*) in the top layer for comparisons between experiments RL and RN. The tensors are for the divergent fluxes: blue dots represent $\boldsymbol{K}_{div}$ and orange dots represent $\boldsymbol{K}_{div} + \boldsymbol{K}_g$. The *x*- and *y*-values are the relative distances for the nonlinear and linear tracer profiles, respectively. The vertical dashed lines are located at the maximum d_r for RL on *x*-axes. If a dot is to the right of its corresponding line, then d_r of this tensor in RN is larger than all ensembles of d_r in RL. The clouds of dots clearly show the non-uniqueness of both $\boldsymbol{K}_{div}$ and $\boldsymbol{K}_{div} + \boldsymbol{K}_g$ for nonlinearly initialised tracers.

series of the ensemble-averaged d_r from all four experiments. For initially linear tracers, the relative differences d_r increase, when the tracers are not relaxed, and this agrees with our findings regarding the numerical errors, as shown in figure 1. For nonlinear tracers, though, d_r does not notably depend on the use of relaxation. Therefore, even though the relaxation effect can affect the numerical accuracy, it is irrelevant for the non-uniqueness of $\boldsymbol{K}_{div}$. Similar observations are made for the other K-tensors, so we opt to not show them for brevity.

Because the diffusion and advection tensors have different physical interpretations, they are often separately parametrised. Therefore, it is necessary to investigate whether they are also individually tracer-dependent. To examine this, we compare the eigenvalues λ_1, λ_2 of the diffusion tensor $\boldsymbol{S}$ and the off-diagonal component A_{12} of the advection tensor $\boldsymbol{A}$ between the experiments RL and RN. We measure the differences by the relative ensemble standard deviation $\delta_r(\langle \cdot \rangle)$, which divides the ensemble standard deviation by the ensemble mean $\langle \cdot \rangle$. For example, $\delta_r(\langle \lambda_1 \rangle)$ is given by

$$\delta_r(\langle \lambda_1 \rangle) = \frac{1}{\langle \lambda_1 \rangle} \sqrt{\frac{1}{14} \sum_{p=1}^{15} (\lambda_1^p - \langle \lambda_1 \rangle)^2}, \tag{3.10}$$

where $\lambda_1^p, p = 1, \ldots, 15$, are the eigenvalues for 15 cases in one experiment.

Figure 10 shows the instantaneous fields of the relative ensemble standard deviations of the eigenvalues λ_1, λ_2 ($\lambda_1 \geq \lambda_2$) and of the off-diagonal element A_{12}. The first and second columns of panels are for RL and RN, respectively. Because the denominator of δ_r is the
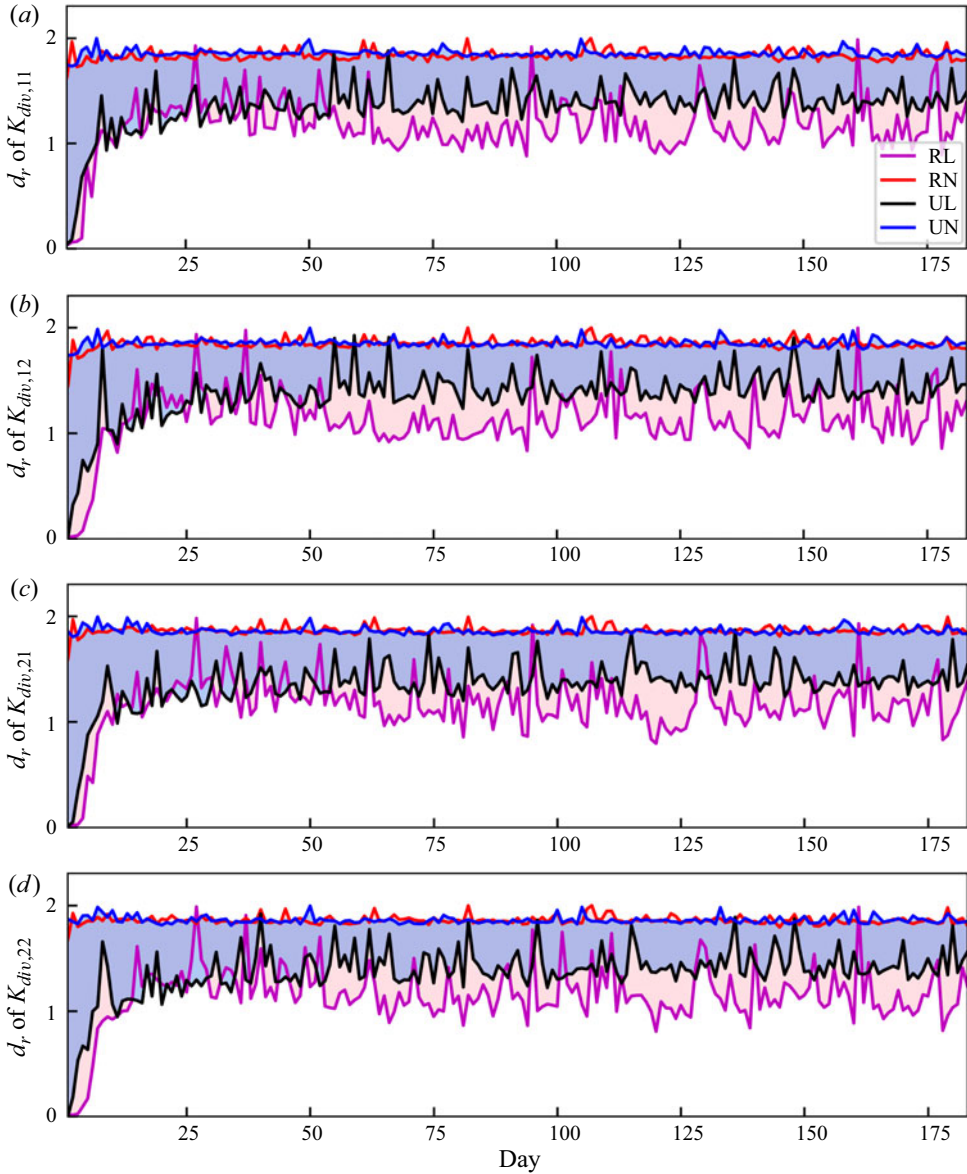
Figure 9. Evolution of d_r in the top layer for all experiments. See the legend in the upper panel for details. The panels correspond to K_{11} (*a*), K_{12} (*b*), K_{21} (*c*) and K_{22} (*d*). The difference between RL and RN, and the difference between UL and UN, are shaded in pink and purple, respectively. The relative distance increases rapidly from zero for RL and UL owing to the initial linear dependence, until the total fluxes in the domain reach equilibrium after 25 days. The black and purple lines fluctuate more, because of the higher frequency of the gradients being nearly aligned.

ensemble-mean, rather than its absolute value, the signs of the components are preserved in the figure. For both experiments, λ_1 is mostly positive in the domain and λ_2 is mostly negative. This polarity of the eigenvalues is also observed for $\boldsymbol{K}_g$ (not shown). Therefore, we deduce that two eigenvalues are often of opposite signs in the domain even without the existence of the rotational flux, which is consistent with HSSB20. The first column shows
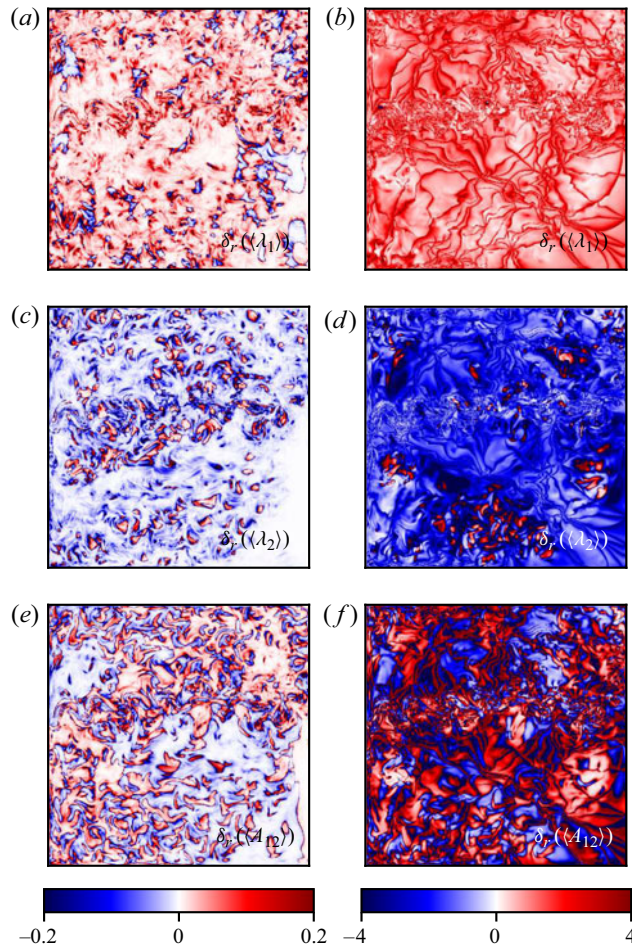
Figure 10. Snapshots of the relative ensemble standard deviation δ_r of the eigenvalues λ_1 (a,b), λ_2 (c,d) of $\boldsymbol{S}$ and of the off-diagonal element A_{12} (e,f) of $\boldsymbol{A}$, all at day 183. Panels (a,c,e) are for RL, where tracers are linearly initialised, while (b,d,f) are for RN (nonlinearly initialised tracers). The non-uniqueness of λ_1, λ_2 and A_{12} is observed, as the standard deviation for the latter is much larger than for the former. For both experiments, $\delta_r(\langle\lambda_1\rangle)$ is mainly positive, as shown in (a) and (b), but $\delta_r(\langle\lambda_2\rangle)$ is mainly negative, as shown in (c) and (d). This indicates that the eigenvalues tend to have opposite signs.

that the ensembles of three variables for initially linear tracers do not differ much from the ensemble-means, as δ_r is predominantly less than 0.2 in the domain. This observation shows that for both $\boldsymbol{S}$ and $\boldsymbol{A}$, the ensembles are close to each other in distances, and agrees with our argument that $\boldsymbol{K}_{div}$ is unique, up to numerical errors, if tracers are initially linear. For the case when tracers are initialised with nonlinear profiles, the relative ensemble standard deviations are much larger, by at least an order of magnitude. Therefore, the diffusion and the advection tensors are also tracer-dependent, that is, they are non-unique for nonlinearly initialised tracers.

Overall, $\boldsymbol{K}_{div}$ and $\boldsymbol{K}_g$ are unique for linearly initialised tracers, and any tracer dependence can serve as a measure of the numerical errors in our approach. The comparison with more general, nonlinearly initialised tracers demonstrates the inherent dependence of the K-tensor on the tracer pairs, and, thus, its non-uniqueness. This conclusion is not affected by a relaxation forcing. The same argument also applies to the eigenvalues of the diffusion

tensor and to the off-diagonal element of the advection tensor. It is worth noting that the non-uniqueness of K-tensors is not limited to the idealised QG model but is also found in a more comprehensive general circulation model (GCM) of the entire Atlantic Ocean (Kamenkovich *et al.* 2021).

4. Conclusions and discussion

This work aimed to investigate the uniqueness of the transport K-tensor in an eddy-resolving model, with the long-term motivation of parametrising eddy tracer fluxes in coarse-resolution ocean models. Our objective here was to calculate K-tensors from high-resolution flow and tracer distributions, to investigate the resulting errors in the eddy flux reconstructions and to examine the K-tensors' dependencies on the passive tracers. The main conclusion is that the K-tensors depend on the tracer and are, therefore, non-unique. This non-uniqueness is strongly influenced by the rotational (non-divergent) component of the eddy tracer flux.

A double-gyre, mid-latitude QG model was used in the study. Tracers were initialised with either linear or nonlinear (sinusoidal wave) profiles, and were advected by the flow for half a year. The flow and tracer fields were decomposed into large- and small-scale components using a spatial filtering method. The resulting eddy forcing E consisted of the divergence of the eddy tracer flux $\boldsymbol{f}$ and non-advective eddy terms. We assumed a flux-gradient relation to link $\boldsymbol{f}$ to the large-scale tracer gradient $\nabla \bar{C}$ via a transport tensor $\boldsymbol{K}_f$. By quantifying the non-advective eddy terms using a divergent flux $\boldsymbol{g}$, we obtained the corresponding K-tensor $\boldsymbol{K}_g$. Then, the full tensor $\boldsymbol{K}$ corresponding to the eddy forcing was given by the sum of $\boldsymbol{K}_f$ and $\boldsymbol{K}_g$. The divergent tensor $\boldsymbol{K}_{div}$ – that is, the portion of $\boldsymbol{K}_f$ corresponding to the divergent eddy flux $\boldsymbol{f}_{div}$ – was obtained by applying the Helmholtz decomposition to $\boldsymbol{f}$. The non-advective eddy forcing terms were comparable in size to $\nabla \cdot \boldsymbol{f} = \nabla \cdot \boldsymbol{f}_{div}$, and similarly the K-tensor $\boldsymbol{K}_g$ was as large as $\boldsymbol{K}_{div}$. Therefore, we deduce that the non-advective terms should be parametrised along with the advective eddy terms.

The accuracy of the inversion method for computing the K-tensors was measured by reconstructing $\boldsymbol{f}$, $\nabla \cdot \boldsymbol{f}$ and E. Generally, $\boldsymbol{K}$ is two orders of magnitude larger than $\boldsymbol{K}_{div}$, owing to the dominant rotational fluxes (Marshall & Shutts 1981). Despite this difference, both tensors were shown to be capable of reconstructing the eddy tracer flux convergence, with their parametrised flux convergences being visually indistinguishable from each other. The largest errors were typically found in the eastward jet region where the eddies were most energetic, but for initially nonlinear tracers, they also occurred in regions where the tracer profiles exhibited sharp spatial gradients (a consequence of the randomly selected nonlinear profiles).

Large errors in the reconstructed flux divergence are introduced by the Helmholtz decomposition, more precisely by the divergence operator, and hence errors associated with $\boldsymbol{K}_{div}$ are much larger than those for $\boldsymbol{K}$. This is despite the fact that the corresponding flux convergences are visually indistinguishable from each other. As the transport tensor component associated with the rotational flux is eliminated automatically under the divergence operator in the tracer evolution equation, it seems more suitable to use $\boldsymbol{K}_f$ rather than $\boldsymbol{K}_{div}$ for a parametrisation, given the errors associated with the latter. However, the additional model errors induced by simplifications in parametrisations may be as large as the errors induced by the flux decomposition. Therefore, the question of whether the difference in errors remains significant after smoothing the fields for coarse-grid models needs to be answered.

For each K-tensor, the tracer-dependence was quantified by measuring the relative distances among its ensembles given by different tracer pairs. We analysed how the tracer-dependence of the K-tensors was affected by various factors such as the initial tracer profiles, the rotational flux and the relaxation forcing. We found that $\boldsymbol{K}_{div}$ and $\boldsymbol{K}_g$ were unique, up to numerical errors, if the tracers were initialised with linear profiles, owing to the linearity of the flux-gradient relation. However, $\boldsymbol{K}_f$ and, thus, the full tensor $\boldsymbol{K}$ were found to be tracer-dependent because the dominant rotational component (i.e. that corresponding to the rotational flux) was tracer-dependent. For nonlinearly initialised tracer profiles, the uniqueness of $\boldsymbol{K}_{div}$ and $\boldsymbol{K}_g$ was lost, and the non-uniqueness of $\boldsymbol{K}$ was further amplified by the presence of the tracer-dependent rotational flux. Therefore, in general, K-tensors are non-unique for passive tracers. The conclusion can be extended to diffusion tensor $\boldsymbol{S}$ and advection tensor $\boldsymbol{A}$, separately. Finally, switching off the relaxation forcing did not affect the non-uniqueness for nonlinear tracers.

The effect of the non-uniqueness of the eddy transport tensor on parametrisations needs to be investigated. Generally, the estimation of the eddy flux for a tracer is different when using different tensors. A question then arises: does it make a fundamental difference whether one uses the tensor given by a tracer pair including this tracer or any tensor from any tracer pair? If not, can these tensors be implemented for the stochastic estimations of the eddy flux patterns? Stochastic closures have been employed in the past (Berloff & McWilliams 2003; Grooms 2016), and our study suggests that treating the eddy transport tensor as a random process may be consistent with its inherent non-uniqueness. In situations when the large-scale flow varies much slower than the eddies, a deterministic solution can be obtained as the tensor given by the flux-gradient relation between $\overline{\boldsymbol{u}'C'}$ and the steady large-scale gradient $\bar{C}$, where $\bar{\cdot}$ denotes the Reynolds averaging. Whether the difference between such tensors and our non-unique transport tensor (which is supposed to be given by the variations of the large-scale terms in the eddy flux) can be treated as stochastic variations still needs to be investigated.

Author ORCIDs.

ⓘ Luolin Sun https://orcid.org/0000-0003-4887-2898;

ⓘ Michael Haigh https://orcid.org/0000-0001-9062-7613;

ⓘ Pavel Berloff https://orcid.org/0000-0001-9143-5955;

ⓘ Igor Kamenkovich https://orcid.org/0000-0001-6228-5599.

REFERENCES

ABERNATHEY, R., FERREIRA, D. & KLOCKER, A. 2013 Diagnostics of isopycnal mixing in a circumpolar channel. *Ocean Model.* **72**, 1–16.

ABERNATHEY, R.P. & MARSHALL, J. 2013 Global surface eddy diffusivities derived from satellite altimetry. *J. Geophys. Res. Oceans* **118** (2), 901–916.

ABERNATHEY, R., MARSHALL, J., MAZLOFF, M. & SHUCKBURGH, E. 2010 Enhancement of mesoscale eddy stirring at steering levels in the Southern Ocean. *J. Phys. Oceanogr.* **40** (1), 170–184.

BACHMAN, S. & FOX-KEMPER, B. 2013 Eddy parameterization challenge suite I: eady spindown. *Ocean Model.* **64**, 12–28.

BACHMAN, S.D., FOX-KEMPER, B. & BRYAN, F.O. 2015 A tracer-based inversion method for diagnosing eddy-induced diffusivity and advection. *Ocean Model.* **86**, 1–14.

BACHMAN, S.D., FOX-KEMPER, B. & BRYAN, F.P. 2020 A diagnosis of anisotropic eddy diffusion from a high-resolution global ocean model. *J. Adv. Model. Earth Syst.* **12** (2), e2019MS001904.

BACHMAN, S.D., MARSHALL, D.P., MADDISON, J.R. & MAK, J. 2017 Evaluation of a scalar eddy transport coefficient based on geometric constraints. *Ocean Model.* **109**, 44–54.

BERLOFF, P. 2016 Dynamically consistent parameterization of mesoscale eddies. Part II: eddy fluxes and diffusivity from transient impulses. *Fluids* **1** (3), 22.

BERLOFF, P., KAMENKOVICH, I. & PEDLOSKY, J. 2009 A model of multiple zonal jets in the oceans: dynamical and kinematical analysis. *J. Phys. Oceanogr.* **39** (11), 2711–2734.

BERLOFF, P. & MCWILLIAMS, J. 2003 Material transport in oceanic gyres. Part III: randomized stochastic models. *J. Phys. Oceanogr.* **33**, 1416–1445.

CANUTO, V.M., CHENG, Y., HOWARD, A.M. & DUBOVIKOV, M.S. 2019 Three-dimensional, space-dependent mesoscale diffusivity: derivation and implications. *J. Phys. Oceanogr.* **49** (4), 1055–1074.

DRITSCHEL, D.G. & MCINTYRE, M.E. 2008 Multiple jets as PV staircases: the Phillips effect and the resilience of eddy-transport barriers. *J. Atmos. Sci.* **65** (3), 855–874.

FERRARI, R. & NIKURASHIN, M. 2010 Suppression of eddy diffusivity across jets in the Southern Ocean. *J. Phys. Oceanogr.* **40** (7), 1501–1519.

FOX-KEMPER, B., FERRARI, R. & PEDLOSKY, J. 2003 On the indeterminacy of rotational and divergent eddy fluxes. *J. Phys. Oceanogr.* **33** (2), 478–483.

FREELAND, H.J., RHINES, P.B. & ROSSBY, T. 1975 Statistical observations of the trajectories of neutrally buoyant floats in the North Atlantic. *J. Mar. Res.* **33**, 383–404.

GROOMS, I. 2016 A Gaussian-product stochastic Gent–McWilliams parameterization. *Ocean Model.* **106**, 27–43.

HAIGH, M., SUN, L., SHEVCHENKO, I. & BERLOFF, P. 2020 Tracer-based estimates of eddy-induced diffusivities. *Deep Sea Res.* I 103264.

KAMENKOVICH, I., BERLOFF, P., HAIGH, M., SUN, L. & LU, Y. 2021 Complexity of mesoscale eddy diffusivity in the ocean. *Geophys. Res. Lett.* **48** (5), e2020GL091719.

KAMENKOVICH, I., BERLOFF, P. & PEDLOSKY, J. 2009 Anisotropic material transport by eddies and eddy-driven currents in a model of the North Atlantic. *J. Phys. Oceanogr.* **39** (12), 3162–3175.

KAMENKOVICH, I., RYPINA, I.I. & BERLOFF, P. 2015 Properties and origins of the anisotropic eddy-induced transport in the North Atlantic. *J. Phys. Oceanogr.* **45** (3), 778–791.

KARABASOV, S.A., BERLOFF, P.S. & GOLOVIZNIN, V.M. 2009 CABARET in the ocean gyres. *Ocean Model.* **30** (2-3), 155–168.

KILLWORTH, P.D. 1997 On the parameterization of eddy transfer. Part I: theory. *J. Mar. Res.* **55**, 1171–1197.

KLOCKER, A. & ABERNATHEY, R. 2014 Global patterns of mesoscale eddy properties and diffusivities. *J. Phys. Oceanogr.* **44** (3), 1030–1046.

KLOCKER, A., FERRARI, R. & LACASCE, J.H. 2012*a* Estimating suppression of eddy mixing by mean flows. *J. Phys. Oceanogr.* **42** (9), 1566–1576.

KLOCKER, A., FERRARI, R., LACASCE, J.H. & MERRIFIELD, S.T. 2012*b* Reconciling float-based and tracer-based estimates of lateral diffusivities. *J. Mar. Res.* **70** (4), 569–602.

LAU, N. & WALLACE, J.M. 1979 On the distribution of horizontal transports by transient eddies in the northern hemisphere wintertime circulation. *J. Atmos. Sci.* **36**, 1844–1861.

LUMPKIN, R., TREGUIER, A. & SPEER, K. 2002 Lagrangian eddy scales in the Northern Atlantic Ocean. *J. Phys. Oceanogr.* **32** (9), 2425–2440.

MADDISON, J.R., MARSHALL, D.P. & SHIPTON, J. 2015 On the dynamical influence of ocean eddy potential vorticity fluxes. *Ocean Model.* **92**, 169–182.

MAK, J., MADDISON, J.R. & MARSHALL, D.P. 2016 A new gauge-invariant method for diagnosing eddy diffusivities. *Ocean Model.* **104**, 252–268.

MARSHALL, J., SHUCKBURGH, E., JONES, H. & HILL, C. 2006 Estimates and implications of surface eddy diffusivity in the Southern Ocean derived from tracer transport. *J. Phys. Oceanogr.* **36** (9), 1806–1821.

MARSHALL, J. & SHUTTS, G. 1981 A note on rotational and divergent eddy fluxes. *J. Phys. Oceanogr.* **11** (12), 1677–1680.

MCCLEAN, J.L., POULAIN, P., PELTON, J.W. & MALTRUD, M.E. 2002 Eulerian and Lagrangian statistics from surface drifters and a high-resolution POP simulation in the North Atlantic. *J. Phys. Oceanogr.* **32**, 2472–2491.

NAKAMURA, N. 1996 Two-dimensional mixing, edge formation, and permeability diagnosed in an area coordinate. *J. Atmos. Sci.* **53** (11), 1524–1537.

O'DWYER, J., WILLIAMS, R.G., LACASCE, J.H. & SPEER, K.G. 2000 Does the potential vorticity distribution constrain the spreading of floats in the North Atlantic. *J. Phys. Oceanogr.* **30**, 721–732.

PLUMB, R.A. 1979 Eddy fluxes of conserved quantities by small-amplitude waves. *J. Atmos. Sci.* **36** (9), 1699–1704.

PLUMB, R.A. & MAHLMAN, J.D. 1987 The zonally averaged transport characteristics of the GFDL general circulation/transport model. *J. Atmos. Sci.* **44** (2), 298–327.

ROBERTS, M.J. & MARSHALL, D.P. 2000 On the validity of downgradient eddy closures in ocean models. *J. Geophys. Res. Oceans* **105** (C12), 28613–28627.

RUDKO, M.V., KAMENKOVICH, I.V., ISKADARANI, M. & MARIANO, A.J. 2018 Zonally elongated transient flows: phenomenology and sensitivity analysis. *J. Geophys. Res. Oceans* **123** (6), 3982–4002.

RYPINA, I.I., KAMENKOVICH, I., BERLOFF, P. & PRATT, L.J. 2012 Eddy-induced particle dispersion in the near-surface North Atlantic. *J. Phys. Oceanogr.* **42** (12), 2206–2228.

RYPINA, I.I., PRATT, L.J. & LOZIER, M.S. 2011 Near-surface transport pathways in the North Atlantic Ocean: looking for throughput from the subtropical to the subpolar gyre. *J. Phys. Oceanogr.* **41** (5), 911–925.

SALLÉE, J.B., SPEER, K., MORROW, R. & LUMPKIN, R. 2008 An estimate of Lagrangian eddy statistics and diffusion in the mixed layer of the Southern Ocean. *J. Mar. Res.* **66**, 441–463.

SAMELSON, R. 1992 Fluid exchange across a meandering jet. *J. Phys. Oceanogr.* **22**, 431–444.

SHUCKBURGH, E., JONES, H., MARSHALL, J. & HILL, C. 2009 Robustness of an effective diffusivity diagnostic in oceanic flows. *J. Phys. Oceanogr.* **39** (9), 1993–2009.

SMITH, K.S. 2005 Tracer transport along and across coherent jets in two-dimensional turbulent flow. *J. Fluid Mech.* **544**, 133–142.

WOLFRAM, P.J. & RINGLER, T.D. 2017 Computing eddy-driven effective diffusivity using Lagrangian particles. *Ocean Model.* **118**, 94–106.

ZHAI, X. & GREATBATCH, R.J. 2006 Inferring the eddy-induced diffusivity for heat in the surface mixed layer using satellite data. *Geophys. Res. Lett.* **33** (24), L24607.