



Estimating a Submesoscale Diffusivity Using a Roughness Measure Applied to a Tracer Release Experiment in the Southern Ocean

EMMA J. D. BOLAND AND EMILY SHUCKBURGH

British Antarctic Survey, Cambridge, United Kingdom

PETER H. HAYNES

Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge, United Kingdom

JAMES R. LEDWELL

Woods Hole Oceanographic Institution, Woods Hole, Massachusetts

MARIE-JOSÉ MESSIAS

School of Environmental Sciences, University of East Anglia, Norwich, United Kingdom

ANDREW J. WATSON

College of Life and Environmental Sciences, University of Exeter, Exeter, United Kingdom

(Manuscript received 15 February 2014, in final form 25 March 2015)

ABSTRACT

The use of a measure to diagnose submesoscale isopycnal diffusivity by determining the best match between observations of a tracer and simulations with varying small-scale diffusivities is tested. Specifically, the robustness of a “roughness” measure to discriminate between tracer fields experiencing different submesoscale isopycnal diffusivities and advected by scaled altimetric velocity fields is investigated. This measure is used to compare numerical simulations of the tracer released at a depth of about 1.5 km in the Pacific sector of the Southern Ocean during the Diapycnal and Isopycnal Mixing Experiment in the Southern Ocean (DIMES) field campaign with observations of the tracer taken on DIMES cruises. The authors find that simulations with an isopycnal diffusivity of $\sim 20 \text{ m}^2 \text{ s}^{-1}$ best match observations in the Pacific sector of the Antarctic Circumpolar Current (ACC), rising to $\sim 20\text{--}50 \text{ m}^2 \text{ s}^{-1}$ through Drake Passage, representing submesoscale processes and any mesoscale processes unresolved by the advecting altimetry fields. The roughness measure is demonstrated to be a statistically robust way to estimate a small-scale diffusivity when measurements are relatively sparse in space and time, although it does not work if there are too few measurements overall. The planning of tracer measurements during a cruise in order to maximize the robustness of the roughness measure is also considered. It is found that the robustness is increased if the spatial resolution of tracer measurements is increased with the time since tracer release.

1. Introduction

The large-scale effect of mixing across isopycnal surfaces potentially plays an important role in determining

ocean stratification and circulation. Ocean tracer release experiments have been much used to determine this effect, often parameterized as a diapycnal diffusivity K_V . However, information on other aspects of oceanic transport and mixing can be obtained from such experiments. For example, the North Atlantic Tracer Release Experiment (NATRE; Ledwell et al. 1998) not only estimated diapycnal diffusivity but also provoked a lively debate about other mixing processes that lead to

Corresponding author address: Emma Boland, British Antarctic Survey, High Cross, Madingley Road, Cambridge CB30ET, United Kingdom.
E-mail: emmomp@bas.ac.uk

horizontal spreading of tracer and, acting in opposition to quasi-horizontal deformation by the mesoscale eddy field, limit the thinning of tracer filaments. In this paper, we examine whether corresponding information on such mixing processes can be obtained from other tracer measurements, such as those made in the Diapycnal and Isopycnal Mixing Experiment in the Southern Ocean (DIMES; Ledwell et al. 2011; Watson et al. 2013) research cruises. The result is an isopycnal diffusivity, along a fixed neutral density surface, that represents the processes unresolved by the current generation of satellite altimetry, constrained by observations, acting at physical scales of 5–50 km and time scales on the order of days. More importantly, the study demonstrates a unique methodology for combining cruise observations, satellite altimetry, and numerical modeling to elucidate the submesoscale, and as such provides a benchmark for future tracer studies, next-generation altimetry, and more sophisticated numerical models.

Discussion of transport and mixing along isopycnal surfaces naturally leads to discussion of diffusivities but often to more than one diffusivity, each representing the effects of distinct physical processes. To clarify the discussion in the remainder of the paper, it is useful to refer explicitly to the description by Garrett (1983) who considered the evolution, in a horizontal plane (e.g., representing an isopycnal surface), of an initially localized patch of tracer under the effect of a mesoscale eddy field characterized by an rms strain rate $\gamma = (u_x^2 + v_y^2)^{1/2}$. The tracer is additionally subject to a submesoscale isopycnal diffusivity K_s , representing small-scale processes that lead directly to molecular mixing. In the model considered by Garrett (1983), this plays the role that would in a laminar flow be played by molecular diffusivity itself. The effect of quasi-random advection by the mesoscale eddy field is assumed to be represented by a horizontal eddy diffusivity K_h . This implies the “domain of occupation” of the tracer, that is, the area that contains most of the tracer, lies within a circle of radius $2(K_h t)^{1/2}$.

The three stages of evolution identified by Garrett (1983) are as follows:

- 1) The tracer patch expands isotropically under the influence of K_s as a circle of radius $2(K_s t)^{1/2}$ until it reaches a size of order $(K_s/\gamma)^{1/2}$, which occurs at a time of order $1/4\gamma^{-1}$.
- 2) The mesoscale eddies begin stirring the tracer into filaments, with widths $(K_s/\gamma)^{1/2}$ and lengths that grow exponentially at a rate proportional to γ . The total area of the filaments therefore also grows exponentially, but the filaments are well separated.
- 3) The total area occupied by the filaments approaches that of the circle of radius $2(K_h t)^{1/2}$. The filaments

merge and lead to approximately uniform tracer fields within the circle.

These three stages need to be taken into account in interpreting observations of a tracer release. Garrett (1983) uses indicative values of $K_s \approx 10^{-2} \text{m}^2 \text{s}^{-1}$, $K_h \approx 10^3 \text{m}^2 \text{s}^{-1}$, and $\gamma \approx 10^{-6} \text{s}^{-1}$ to estimate in particular that the transition between stages 2 and 3 would take place on a time scale of a year or so when the radius of the circle is about 400 km.

The diffusivity K_h is equivalent to that estimated from large-scale tracer observations, float separation measurements, or model calculations. Typical estimates for the Southern Ocean are in the range $100\text{--}1000 \text{m}^2 \text{s}^{-1}$, found using tracer observations, models, or a combination of both (McKeague et al. 2005; Garabato et al. 2007; Zika et al. 2009; Abernathey et al. 2010), broadly consistent with the indicative value $10^3 \text{m}^2 \text{s}^{-1}$ used by Garrett (1983).

Our primary concern in this paper is not K_h but the submesoscale diffusivity K_s . Uncertainty continues over the processes that control the magnitude of K_s . Young et al. (1982) suggested shear dispersion by inertial waves as an important process and correspondingly estimated $K_s \approx K_i = N^2 K_V / f^2$, where N^2/f^2 is the ratio of the buoyancy and Coriolis frequencies. Taking $K_V = 0.2 - 4 \times 10^{-4} \text{m}^2 \text{s}^{-1}$ (Watson et al. 2013) and $N^2/f^2 = 200$ (Smith and Marshall 2009) for the DIMES region implies $K_i = 0.004 - 0.08 \text{m}^2 \text{s}^{-1}$ [$0.01 \text{m}^2 \text{s}^{-1}$ being the value used by Garrett (1983) in his indicative calculations].

The NATRE experiment involved tracer release on a density surface ~ 300 m deep, and there were sufficient observations to track the evolution of the horizontal structure of the tracer through stages 1 and 2. Using these observations, Ledwell et al. (1998), taking account of numerical simulations of Sundermeyer and Price (1998), find a K_s of $0.07 \text{m}^2 \text{s}^{-1}$ at scales of 0.1 to 1 km, estimated from horizontal dispersion in stage 1, and a K_s of $2 \text{m}^2 \text{s}^{-1}$ at scales of 1 to 10 km, estimated from tracer filament widths and stretching rates in stage 2. [Note that the apparent scale dependence of K_s would be consistent with a broad range of scales of submesoscale flow structures, with the diffusivity at a given scale being determined by the flow structures with scale smaller than the given scale (Richardson 1926; Richardson and Stommel 1948; Okubo 1971). See, for example, Lumpkin and Elipot (2010) or Koszalka et al. (2009) for oceanic measurements showing this scale dependence.] The clear conclusion is that the K_s required to account for the filament width in stage 2 must be significantly larger, by two orders of magnitude, than K_i . The three diffusivities so far mentioned, including order of magnitude estimates for the Southern Ocean and the scales

at which they operate, are depicted in Fig. 1. (Note that if there was a strong scale dependence to, for example, K_s , then this might be indicated by a slope to the corresponding ellipse in the figure, but bearing in mind the significant uncertainty over the processes controlling K_s , we have chosen not to include this.)

At least two explanations have been suggested for the difference between K_i and the inferred K_s . One is that vortical modes (circulations arising from potential vorticity anomalies arising from localized vertical mixing events) can provide additional horizontal dispersion (Polzin and Ferrari 2004; Ferrari and Polzin 2005). Another is that the K_s that is inferred from filament widths can be explained only by considering three-dimensional processes. In particular, the tilting effect of vertical shear means that a tracer filament observed on a single horizontal surface is in fact a cross section through a sloping tracer sheet. The action of the vertical diffusion K_V can play an important role in setting the horizontal tracer structure (Haynes and Anglade 1997; Haynes 2001; Smith and Ferrari 2009).

In the NATRE experiment, tracer measurements were taken within a few weeks and a few months of the tracer release and were of sufficient spatial resolution to be able to map out the geometry of the filaments directly (see Fig. 1 of Ledwell et al. 1993). In DIMES, on the other hand, the combination of the resolution of the measurements and the fact that the first return measurements were made roughly a year after the release, when we might expect to be approaching Garrett's (1983) stage 3 with filament merger taking place, means that this direct approach is not feasible.

Here, we seek to exploit some of the techniques that have been used in atmospheric science to infer information on mixing processes from in situ measurements of atmospheric chemical distributions in the lower stratosphere and upper troposphere. When the measurements are taken from aircraft they are largely horizontal sections, and when they are taken from balloons they are vertical sections. Information on mixing processes has been extracted by a two-stage approach. The first stage has been to use numerical simulations based on the solution of the advection–diffusion equation to generate a set of model chemical distributions for different assumed diffusivities. The velocity fields have been taken from large-scale meteorological datasets. The chemical fields have in some cases been initialized from satellite observations and in some cases driven by a hypothesized large-scale forcing. The approach has sometimes been to try to simulate specific features in the observed chemical distributions and sometimes to simulate the generic spatial structure of the chemical fields. The second stage has been to make some quantitative

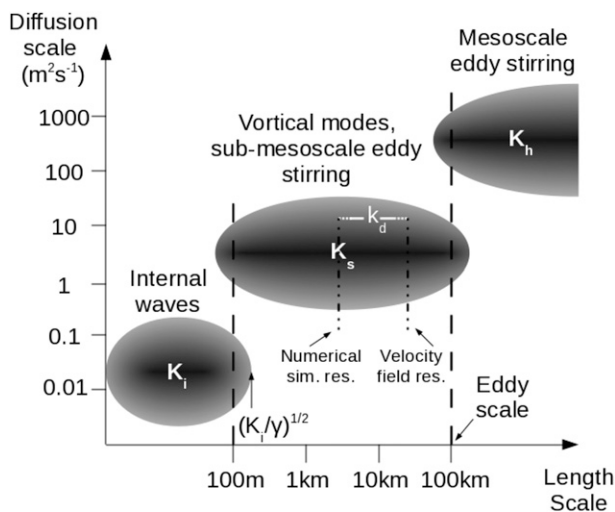


FIG. 1. Depiction of the isopycnal diffusivities active at various length scales, along with the physical processes they are thought to represent, as described in section 1, along with order of magnitude estimates relevant for the midlatitude ocean. The term K_d as labeled is the diffusivity we attempt to estimate for the DIMES tracer in this study and is representative of the value of K_s at a length scale determined by the resolutions of the numerical simulations and the advecting altimetric velocity fields.

comparison between the set of simulated fields (with each member of the set corresponding to a different diffusivity) and the observations and thereby to deduce a “best-estimate” diffusivity for the atmosphere. Previous studies that have taken this two-stage approach include Balluch and Haynes (1997), Waugh et al. (1997), Legras et al. (2003), Haynes and Vanneste (2004), and Legras et al. (2005).

In many of the studies described above, the vertical and horizontal structure of the flow (which is now routinely available from meteorological datasets) has been taken into account and a vertical diffusivity K_V has been inferred. The notion of an equivalent horizontal diffusivity, predicted by Haynes and Anglade (1997) to be $K_V\alpha^2$, where α is the aspect ratio of tracer structures, remains useful for some purposes, for example, in the oceanic case in order to compare the relative roles of vertical diffusivity acting on tilted sheets versus the effect of vortical modes. However, it needs to be kept in mind that such an equivalent horizontal diffusivity may be an imprecise quantification at best (Haynes and Vanneste 2004; Smith and Ferrari 2009), essentially because there is no single value of the aspect ratio α .

In applying a corresponding approach to DIMES, we have to accept the following: The best available information on velocity fields is that calculated from satellite altimetry. Such velocity fields have been used in many previous studies to estimate large-scale Southern

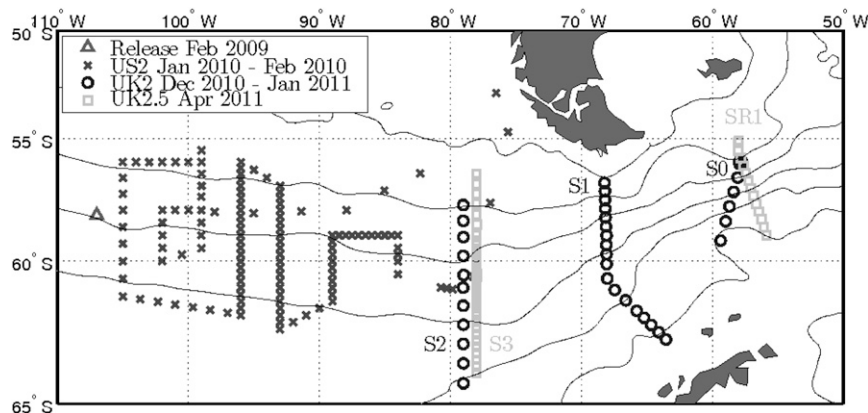


FIG. 2. Location of tracer measurements in the US2, UK2, and UK2.5 cruises as indicated. The triangle shows the location of the tracer release on the US1 cruise. The transects from the UK2 and UK2.5 cruises are labeled. The contours are mean streamlines, separation $2 \times 10^4 \text{ m}^2 \text{ s}^{-1}$.

Ocean diffusivities (see, e.g., Marshall et al. 2006). However, the altimetry gives only the surface geostrophic velocity and gives no useful information on vertical structure. Therefore, any advective–diffusive calculation has to be two-dimensional, and what can potentially be inferred is an equivalent horizontal diffusivity. In addition, the tracer information available from the DIMES measurements is relatively sparse in space and time. Direct comparison of individual spatial structures in the tracer field that are observed and those that are simulated is not feasible. Therefore, the comparison must be on the basis of some gross measure of the spatial structure. We choose here to use a roughness measure previously exploited in the atmospheric context by Legras et al. (2003, 2005). The roughness measure, described in detail in section 4, is a measure of the comparative streakiness of a set of data points. The relation to the ideas in Garrett (1983), discussed previously, is that in stage 2, when there are well separated streaks, the roughness will be relatively large, whereas in stage 3, when streaks are beginning to merge, the roughness will be smaller. The approach therefore is to compare the roughness measure obtained from the tracer measurements with the same roughness measure from a set of numerical simulations of tracer evolution with different imposed submesoscale diffusivities. The simulated roughness measure is a strong function of the imposed diffusivity and therefore, provided there are sufficient observations, a best-match value of the submesoscale diffusivity can be determined.

In the following, we first set out the particulars of the numerical simulations undertaken and show how they compare to observations at first glance. We then describe in detail the roughness measure, in particular investigating its robustness when applied to the relatively sparse observations of the DIMES research cruises.

Having determined the uncertainty in the roughness measure, we show it can be used successfully to compare observations and simulations in cases with a sufficient number of observations.

Additionally, we investigate using the roughness measure to design cruise sampling plans. When at sea, the observationalist faces resource constraints on time and distance and, depending on the study, may have to balance the desire for high-resolution measurements with the desire to cover a large area. We test this on both a representative transect of the DIMES tracer and by analyzing past cruise sampling patterns.

Finally, we discuss the interpretation of the value of K_s found in this study, in particular to what precisely it applies and where it fits in the context of previous studies. Having tested the robustness of the roughness measure, we discuss its value for future studies.

2. Observations and numerical simulations

a. The DIMES observations

DIMES (<http://dimes.ucsd.edu>) is a joint U.K. and U.S. program designed to measure interior mixing in the Southern Ocean. The DIMES observational campaign was designed to encompass the relatively smooth bathymetry of the east Pacific sector of the Southern Ocean and the relatively rough bathymetry of Drake Passage and the Scotia Sea and is expected to result in increased vertical diffusivities in the latter compared with the former. Results have already been reported in, for example, Ledwell et al. (2011) and Watson et al. (2013). Figure 2 shows a schematic of the first 2.5 yr of the experimental side of the project. This began in early 2009 with the release of the tracer, chosen for its

low background concentrations and the ability to measure very small concentrations accurately, in the east Pacific sector of the Antarctic Circumpolar Current (ACC; triangle) at the depth of the $\gamma_n = 27.9 \text{ kg m}^{-3}$ neutral density surface (approximately 1.5 km deep). There have been nine return research cruises to date to measure the distribution of the tracer. In this paper, we will exploit the information available from the US2, UK2, and UK2.5 cruises, in particular concentrating on the horizontal variation of the column-integrated tracer.

b. Numerical simulations

To provide a set of simulated fields for comparison against these observations, we used a model based on the solution of the advection–diffusion equation in a two-dimensional horizontal flow with imposed constant horizontal diffusivity K_d (with K_d taking different values in different simulations). The numerical code used was the MITgcm in offline mode, that is, with imposed velocity fields (see details below). The concentration field in this two-dimensional simulation is intended to correspond to the column-integrated tracer in the observations. As discussed in the introduction, this requires neglect of vertical structure in the flow and its effect, along with vertical mixing, on the tracer field. It might be noted that the approach here is different from that in previously mentioned atmospheric studies in that the simulations solve the full advection–diffusion partial–differential equation over a finite region rather than using a Lagrangian–stretching approach that allows the spatial structure of the tracer field to be deduced from integration along particle trajectories (Haynes and Vanneste 2004) or a stochastic Feynman–Kac approach that allows construction of the tracer field along a single one-dimensional section (Legras et al. 2003). The reduced computational expense of a 2D simulations compared with a 3D simulation is therefore particularly important, allowing us to go to greater spatial resolution and carry out multiple realizations of the experiment. The calculations reported below were carried out at high resolution ($1/20^\circ$ or $1/50^\circ$, on the order of 5 or 2 km, respectively), allowing the development of correspondingly small-scale structure in the tracer field (although not necessarily submesoscale features associated with, for example, frontogenesis).

The horizontal velocity field supplied to the calculation was intended to be a representation of the actual velocity field on the neutral density surface corresponding to the DIMES tracer release and subsequent evolution. The approach taken was as follows: The surface velocity field was first estimated from delayed time satellite altimeter data produced by SSALTO/DUACS,

which have been postprocessed and passed through quality control measures.¹ In particular, we used a dataset of weekly sea level anomaly (SLA) merged from two satellites for continuity on a $1/4^\circ$ by $1/4^\circ$ Cartesian grid, from February 2009 to April 2011, in combination with the current mean dynamic topography (MDT) based on 1993–96 SLA. We then postulated, on the basis that previous observational studies (e.g., Phillips and Rintoul 2000) and model studies (e.g., Killworth and Hughes 2002) have shown the ACC flow to be equivalent barotropic, that the flow on the tracer neutral density surface is a constant fraction (the “velocity fraction”) of that at the surface. This is, of course, potentially a gross simplification since even if the ACC is equivalent barotropic, the relationship between the velocity at the tracer level and the surface velocity may not be constant in space and/or time. For the simulations to be analyzed later in the paper, the velocity fraction was chosen as 0.33. Justification for this choice is given in section 2d.

Before use in the advection–diffusion calculation, these velocity fields were interpolated onto finer resolution grids, ($1/20^\circ$ or $1/50^\circ$) as appropriate, and rendered nondivergent at the boundaries before use. The boundaries were provided by the maximum land mask from the altimetry data, and so a fixed maximum sea ice extent is present at all times. The domain was circumpolar in longitude and from roughly 30° to 66°S in latitude, on a spherical polar grid.

In each simulation, the initial tracer concentration field was imposed to have the same distribution at the same location and time of the real tracer release—408 mol of tracer in a cross roughly centered on 58°S , 107°W in early February 2009. The evolution according to the advection–diffusion equation was followed for 2.5 yr, in order to cover the timing of the US2, UK2, and UK2.5 cruises (see Fig. 2).

c. Choice of diffusivity K_d and numerical resolution

We initially tested a simple, second-order central difference advection scheme (Adcroft 1995) but found that a significant percentage of the tracer field became negative within a few weeks of simulation. Instead, we chose a second-order moment advection scheme (Prather 1986) with a limiter ensuring no negative tracer values (although test simulations with the second-order scheme at $1/20^\circ$ showed no significant difference in the roughness measurements). A time step of 6 min was used in all simulations.

¹ Distributed by AVISO, with support from CNES (<http://www.avisioceanobs.com/duacs/>).

A next consideration is whether the chosen diffusivity K_d and the numerical resolution are compatible. This is assessed following the method of Marshall et al. (2006), in which a numerical diffusivity k_{num} is estimated by calculating the decay rate of tracer variance.² This (total) numerical diffusivity is in part because of the diffusivity K_d explicitly included in the advection–diffusion calculation and part because of the effects of finite numerical resolution. Table A1 in the appendix contains the estimated numerical diffusivities for six simulations at a variety of diffusivities and resolutions, each with a fixed velocity fraction of 33%. The numerical diffusivity varies slightly over time, with standard deviations of approximately 14% or 2% of the set value K_d for the $1/50^\circ$ or $1/20^\circ$ simulations, respectively. The numerical diffusivity adds 20%–60% to the prescribed K_d for the 2 and $20 \text{ m}^2 \text{ s}^{-1}$ $1/50^\circ$ simulations and adds 200%–300% to the $0.2 \text{ m}^2 \text{ s}^{-1}$ $1/50^\circ$ simulation. However, all the $1/20^\circ$ simulations remain within a standard deviation of K_d (although there may be locally enhanced values, this is a calculation on global average tracer gradients). This is consistent with the greater variability of the tracer field in the higher-resolution simulations. The reader should note that this spatial and temporal variability of the total numerical diffusivity, a characteristic of numerical simulations, means that the results presented here are only strictly relevant to the numerical simulation setup described here.

Assessing the large-scale eddy diffusivity K_h of our simulated tracer by considering the spatial variance of the tracer distribution as in Tulloch et al. (2014), we found it was largely insensitive to the value of the diffusivity K_d used in our model simulations. This is consistent with the expectation that K_h is dominated by stirring and advection by mesoscale eddies that is well represented by the velocity fields supplied to the calculation. Comparing the predicted K_h for different K_s against that estimated from observations would therefore be a poor approach to choosing a “best” K_s .

d. Choice of velocity fraction

To assess the appropriate scale factor to reduce the surface velocities to tracer-level velocities (which we will call the velocity fraction), we looked at two methods. First, we directly calculated the implied velocities given by RAFOS float locations. Each of the 140 deep floats (designed to remain on the tracer neutral density surface) had its location recorded daily for up to 2 yr from early 2009, and the locations were turned into approximate velocities using a finite-difference approximation. These

were then compared with the weekly surface velocities derived from satellite altimetry mentioned previously and linearly interpolated to the same locations and times. However, because of ballast problems, many of the floats did not stay on the target neutral density surface (LaCasce et al. 2014). Using the temperature data recorded by the floats, we restricted the data points to those where the temperature was within 0.1° of 2.3° , the temperature determined by LaCasce et al. (2014) to be equivalent to being on the tracer neutral density surface. This resulted in a dataset with 16 482 of the original 56 283 points, 29% the size of the original.

This method has the advantage of using a large amount of data points produced directly from observations. However, the satellite altimetry is on larger spatial and temporal scales, and so the features being experienced by the floats may not be well represented in the altimetry and thus the fraction may be inaccurate. The results of this calculation can be seen in Fig. 3, which shows histograms of the ratio between these two derived velocities, where each point is representative of one velocity measurement on 1 day. These have been divided up into four longitudinal sections, with roughly the same number of points in each of the first three most westerly sections but fewer in the fourth as only a small number of floats traveled east of 70°W in the 2 yr of data used. Also shown (gray numbers) are the mode (i.e., most common) fractions from the histogram, chosen because the distributions are skewed. These histograms point to a variable velocity fraction in the range of 33%–43% [the fractions of the velocity components u and v (not shown) show similar values over a slightly larger range of 30%–45%]. This is comparable to the values found within the ACC in the Ocean Circulation and Climate Advanced Model (OCCAM) at depths of ~ 1 – 2 km (see Killworth and Hughes 2002, their Fig. 7).

The second method we used to assess the most suitable velocity fraction was to carry out a variety of simulations at fixed horizontal diffusivity and with variable velocity fraction and to then compare the center of mass of the simulations with the observations. We subsampled the simulated tracer field at the same location that tracer measurements were made on the cruises, which can be seen in Fig. 2. At each station, tracer measurements were made at several depths, and so we compared the vertically integrated measurement, or so-called column integral, at each station. This method is attractive in its simplicity, and by its nature produces the long-time average fraction that allows for the best match between simulation and observation.

Figure 4 shows a direct comparison between the center of mass of the simulations and observations, where the center of mass of all observations on each

² If C is the tracer concentration, $1/2\partial\langle C \rangle/\partial t = -k_{\text{num}}\langle |\nabla C|^2 \rangle$, where $\langle \rangle$ is a global average.

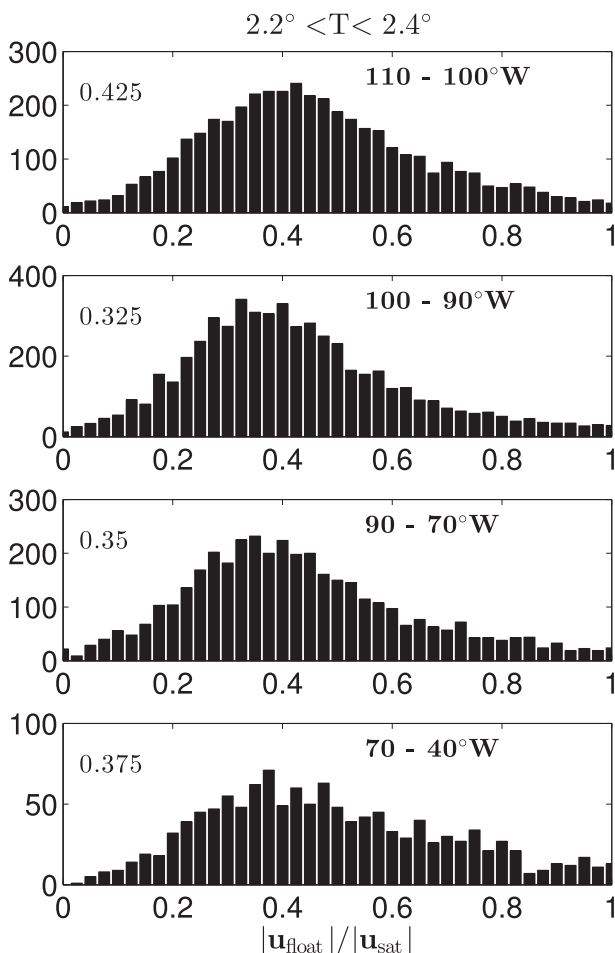


FIG. 3. Histograms of the ratio between the satellite-derived surface velocities and the RAFOS float-derived tracer-level velocities, where the measure temperature fell between 2.2° and 2.4°C (see text), divided into four longitudinal sections. Gray numbers indicate the peak velocity fraction from each histogram.

cruise is marked by a black circle. The center of mass of the simulations, subsampled identically to the observations, is marked with other symbols, with the velocity fraction as labeled. The center of mass of the UK2.5 observations is actually farther upstream than the UK2 observations, despite being measured 3–4 months later. This is because of the higher number of measurements on the upstream S3 line on this cruise. As expected, the spread between the different simulations increases with time (top to bottom), but the observations remain consistently between the 33% and 37% simulations, but closer to 33%, suggesting that a single velocity fraction that would reproduce the observations most closely lies at a value of $\sim 34\%$. Given these results we choose a velocity fraction of 33%, unless explicitly stated otherwise, for the simulations to be considered in the remainder of the paper.

e. Preliminary comparison against observations

Figure 5 shows snapshots of the simulated tracer field roughly 2 weeks, 6 months, 1 yr, and 2 yr after release, with a horizontal diffusivity of $K_d = 20\text{ m}^2\text{ s}^{-1}$, velocity fraction of 33%, and resolution of $1/50^\circ$.³ The three stages of tracer evolution as described in section 1 can be seen in these snapshots. Initially, the tracer expands isotropically as a single patch, but after 2 weeks, the top-left panel shows that the patch has started to feel the strain of the velocity field and is noticeably wider in the zonal direction. Around 6 months later (top-right panel), the majority of the tracer remains in one large patch to the west of the plot, but several streaks have been, and are in the process of being, created, pulled off, and stirred to the east. Around a year after release (bottom-left panel, during the US2 cruise), a single patch is no longer discernible, all of the tracer is now wrapped around velocity features in streaks, some of which are beginning to merge. Then 2 yr after release (during the UK2 cruise), while there is still inhomogeneity, the tracer streaks have merged to create a large patch many hundreds of kilometers wide.

A first assessment of how the choice of K_d affects the qualitative agreement between observations and simulations is provided by Fig. 6, which shows the observed tracer column integrals (black crosses) on the first three return cruises against the along-track distance for each cruise. The first few stations of the US2 cruise are omitted as these are relatively spaced out and have low values of tracer measured. The UK2 and UK2.5 cruises are split into transects as labeled in Fig. 2 and arranged such that the transects are progressively upstream, or farther to the west, from left to right. Also shown are the tracer fields along the cruise path (colored lines) and the subsampled values at the measurement station locations (colored circles) from three simulations with various horizontal diffusivities $K_d = 20, 50,$ and $100\text{ m}^2\text{ s}^{-1}$ as labeled. All three simulations had a velocity fraction of 33% and a resolution of $1/20^\circ$. Across all of the cruises, the effect of increasing horizontal diffusion can be seen in the smoothing of the tracer field, resulting in less extreme spikes.

The measurements taken on the US2 cruise, 1 yr after release, show good similarity to the simulations, for example, with respect to the position of the main body of the tracer, especially in the second half of the cruise track. The simulations and cruise results are less well matched for the UK2 and the UK2.5 cruises, but we expect the difference between the simulation and the

³ See Table A1 for the variability of the numerical diffusivity for this simulation.

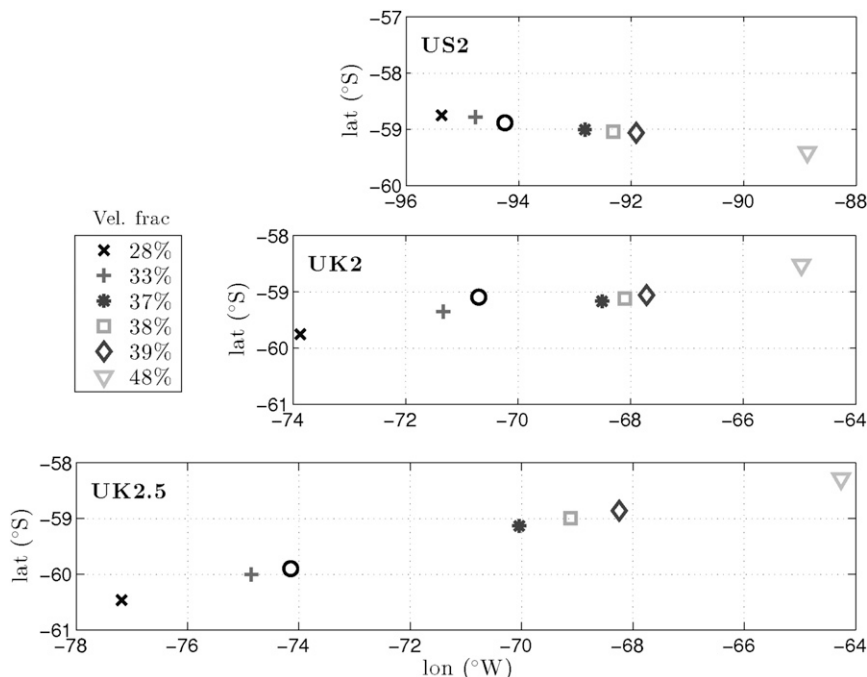


FIG. 4. Center of mass comparison between observations on the various cruises (black circle) and simulations with $K_d = 20 \text{ m}^2 \text{ s}^{-1}$, $1/20^\circ$ resolution, and velocity fractions as indicated.

observations to increase with time because we have seen that the velocity fraction is a spatially varying quantity (Fig. 3); so the velocity fraction that matches the center of mass most accurately is likely to be a domain-

averaged value. Additionally, Fig. 4 shows that the center of mass of the 33% simulation becomes further from the observations with time. The differences between the simulations and measurements is also

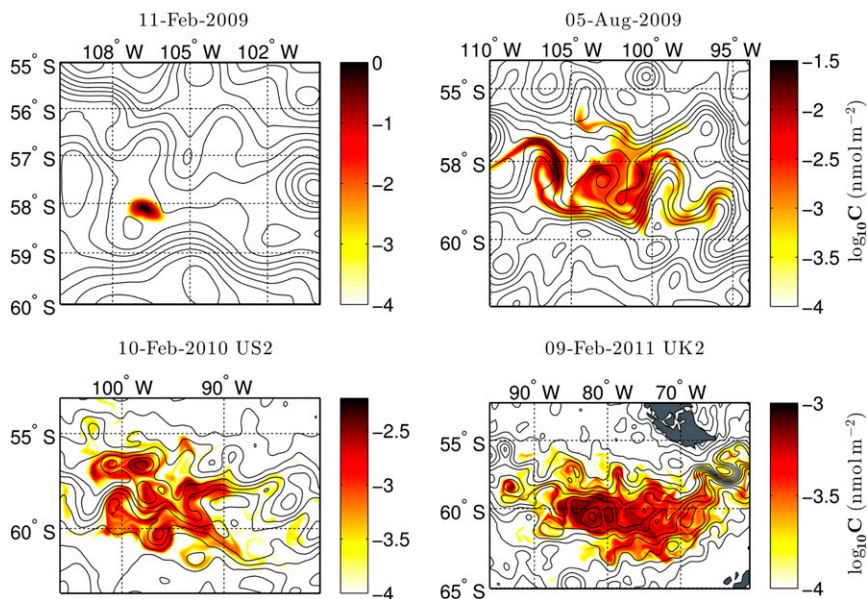


FIG. 5. Snapshots of tracer concentration (shading, log scale) from MITgcm simulation with $K_d = 20 \text{ m}^2 \text{ s}^{-1}$, a velocity fraction of 33%, and $1/50^\circ$ resolution. The contours show instantaneous streamlines. Top-left panel is approximately 2 weeks after release; the tracer has remained in a single patch but has started to be strained by the velocity field. As time progresses, the velocity field pulls the tracer into streaks, which eventually merge on the time scale of years. The bottom two panels coincide approximately with the US2 and UK2 cruises, respectively, as labeled.

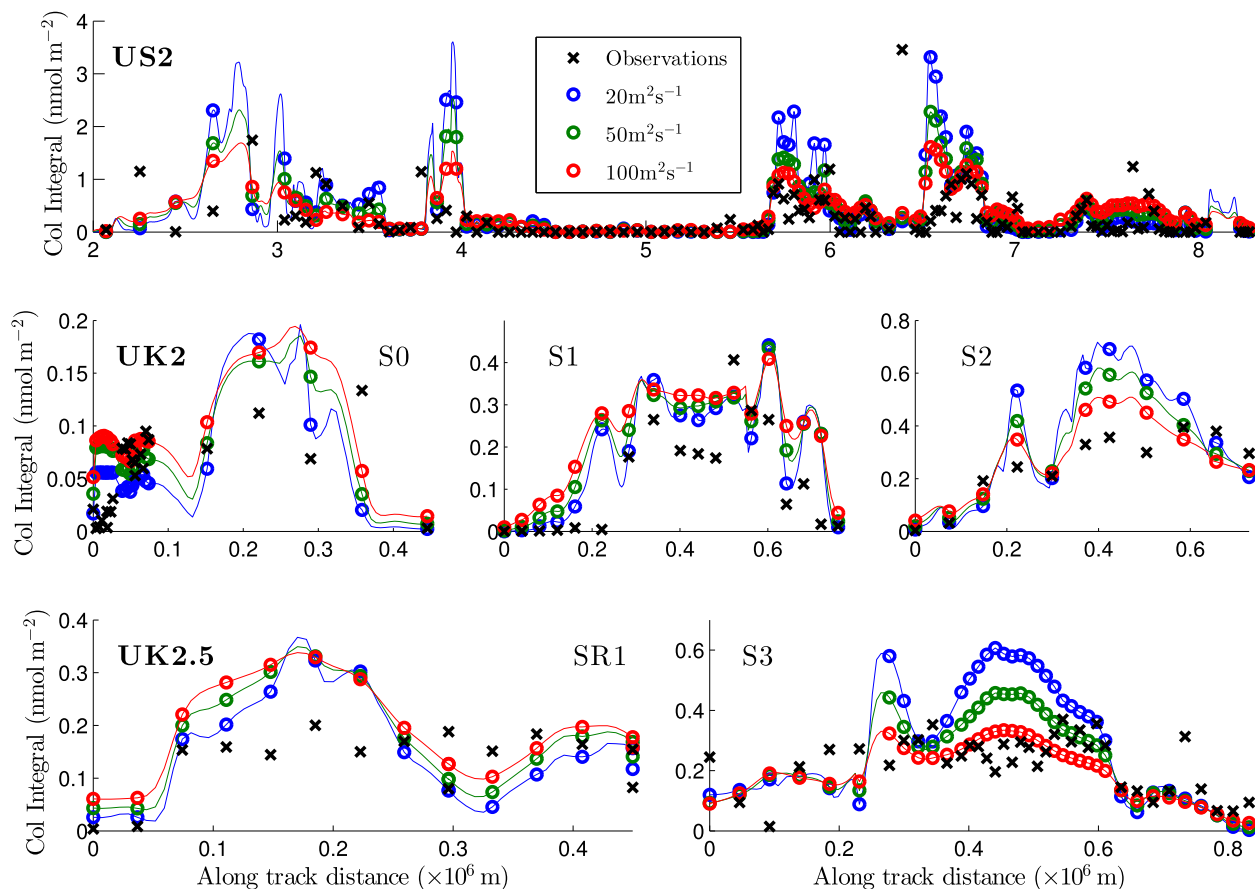


FIG. 6. Tracer measurements (black crosses) from the first three cruises, as shown in Fig. 2, against the cruise along-track distance. Also shown are subsampled $1/20^\circ$ simulations (colored circles) with fixed velocity fraction 33%, $1/20^\circ$ resolution, and different horizontal diffusivities as labeled with the thin colored lines showing the simulated tracer between sample points.

expected to increase with time because of the limitations of the simulations—imperfect knowledge of initial conditions and the velocity field, static boundary conditions, and so on—compounding over time.

Figure 7 shows the same observations as Fig. 6 but with simulations with diffusivities $K_d = 0.2$, 2, and $20 \text{ m}^2 \text{ s}^{-1}$ as labeled and a horizontal resolution of $1/50^\circ$. The higher resolution is used in order to adequately resolve the smaller diffusivities; we can directly compare the $K_d = 20 \text{ m}^2 \text{ s}^{-1}$ simulation with the lower-resolution version (Fig. 6) to ascertain any sensitivity to this choice. Note that the vertical axes have changed scale, but the horizontal axes are as before. While the qualitative form of the simulated measurements has not changed drastically, the lower diffusivities result in peaks far above those seen in the observations.

3. Roughness measure

The key question is how to proceed further determining the diffusivity K_d that gives the best match to

the observations. An elementary point is that we do not expect an exact quantitative match to a simulation at the location of every observation since there are inevitably significant differences between the predicted position of filaments of tracer and the observed position because relatively small errors in the simulation of advection imply large differences in the position of filamentary features relative to their thickness (see, e.g., Methven and Hoskins 1999), and so on. However, we might postulate that the overall “streakiness” or “spikiness” of the simulation could be matched to that of the observations. For a more objective comparison between the streakiness of the simulations and of the observations to assess which simulation most closely matches, we used a roughness measure as previously used in Legras et al. (2003) to match simulations of various diffusivities with observations of ozone profiles in the lower stratosphere. This assesses the roughness of a measured field as a function of the area between two osculating curves fit around the data, over a range of scales, giving a more robust means of comparison rather than a single

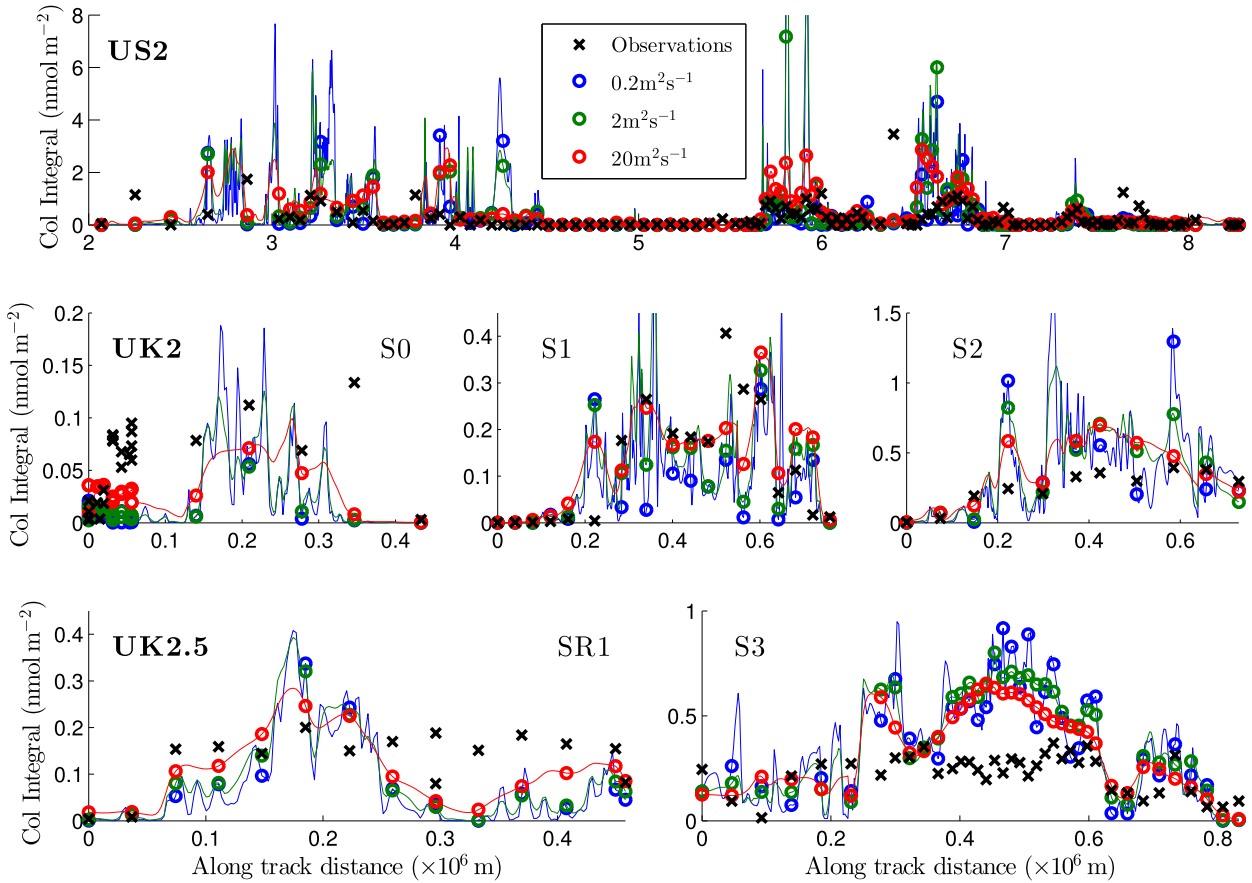


FIG. 7. As in Fig. 6, but with $1/50^\circ$ simulations and different horizontal diffusivities as labeled. Note that the vertical scales are different from Fig. 6.

measure such as variance. Legras et al. (2003) consider other measures of the spatial structure of a field, including the wavenumber power spectrum and the variance of differences measured over different spatial increments and conclude that the roughness measure is the most effective at capturing the spatial structure over a range of scales. As in Legras et al. (2003, 2005), we do not present a physical interpretation of the roughness but use it as a statistical tool, although this would be an interesting area of future study.

The roughness is defined in terms of two curves constructed from a series of parabolas with curvature p of the form

$$2p(y - y_c) = (x - x_c)^2. \quad (1)$$

At each measurement point (x_i, y_i) , where x_i is spatial coordinate and y_i is the value of the tracer, the osculating curve $y = y_p^+(x_i)$ corresponds to the smallest value of y_c such that the parabola with $x_c = x_i$ and curvature p lies above all measurement points. Similarly, $y = y_p^-(x_i)$ corresponds to the largest value of y_c such that the

parabola with curvature $-p$ lies below all points. Examples of two such osculating curves can be seen in Fig. 8, which shows the tracer measurements from the US2 cruise against the along-track distance (solid line) and the two osculating curves y_p^+ and y_p^- (dashed lines) for $p = 0.1$. The roughness $\Phi(p)$ for N measurements is then defined as

$$\Phi(p) = \frac{1}{N} \sum_{i=1}^N [y_p^+(x_i) - y_p^-(x_i)]^2. \quad (2)$$

Comparing the roughness $\Phi(p)$ of the observations with the roughness of the simulated tracer should thus provide an objective way of assessing which diffusivity best matches the streakiness of the observations.

It is important to note that the roughness measure $\Phi(p)$ depends on the absolute magnitude of the field, that is, the field $\chi(x)$ would not have the same roughness as the field $\lambda\chi(x)$ unless $\lambda = 1$. Therefore, since the observations are column-integrated tracer, the simulated field must also be column-integrated tracer, and it is important that the simulations are initialized with the

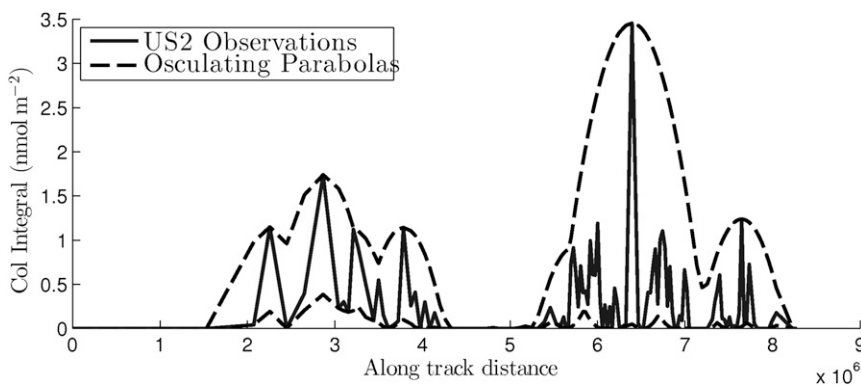


FIG. 8. US2 observations (solid line) and osculating curves y_p^+ and y_p^- (dashed lines) with curvatures $p = 0.1$. The roughness $\Phi(p)$ is a function of the area between these two curves; see text for details of the calculation.

estimated column-integrated tracer as released. The comparisons shown in Fig. 6, for example, provide reassurance that the correct magnitudes of column-integrated tracer are being captured by the simulations.

Examples of the roughness curves from simulations can be seen in Fig. 10 (shown below), which shows results from all three cruises, but with only the S1 transect from UK2 and the S3 from UK2.5. The thick solid lines show the roughness of the cruise-imitating samples (subsamped at the identical times and locations of the cruise observations, as plotted in Figs. 6 and 7, except against time rather than along-track distance) for simulations with $k_h = 0.2$ and $100 \text{ m}^2 \text{ s}^{-1}$, respectively.

Uncertainty in the roughness calculation

Because we are applying the roughness technique to relatively sparsely sampled observations, one is naturally lead to the question of the robustness of the measure—if we are not sampling individual streaks of tracer, does the measure still give a result dependent on the underlying diffusivity K_s ? Here, we look at two techniques to determine the uncertainty in the roughness measure as applied to our results. The first, technique A, assesses how robust the roughness is to small changes in the exact sampling location; that is, if we slightly shift the sampling track in space and/or time, while retaining the same spacing between stations, would we get the same roughness? The second, technique B, assesses how dependent the roughness is to the exact sampling locations along the chosen cruise track.

Technique A involved perturbing the sampling of the simulations in space and time, maintaining the spacing between sample points in space and time equal to the observations. Examples of two such tracks for UK2 S1 can be seen in Fig. 9 (left-hand side). Note that the perturbing is also carried out in time, which is not shown.

After perturbing a maximum of $\pm 3/20^\circ$ in both latitude and longitude and $\pm 17 \text{ h}$ in time (values chosen to provide as large a range as possible without too large a computational burden), we repeated the roughness calculation on each of the 54 new tracks produced and took the maximum and minimum roughness found as the uncertainty limits. Examples of these limits can be seen in Fig. 10, which shows the roughness calculation for the three cruises, the 0.2 and $100 \text{ m}^2 \text{ s}^{-1}$ simulations (solid lines), and uncertainty A (dashed lines).

Technique B was a bootstrapping analysis as follows: we randomly resampled the full-resolution simulated track with the same number of points as observations (allowing for resampling) 1000 times and found the confidence intervals from the distribution of the roughness of these tracks. One example of such a track can be seen in the right-hand side of Fig. 9 for UK2 S1. The confidence interval widths are similar to the uncertainty bands from technique A, although they place the cruise-imitating sampling at the rough end of the uncertainty bands, close to the 75% interval. The 75% and 95% intervals can be seen in Fig. 10 as the error bars.

In general, we found that the uncertainty from both techniques was inversely dependent on the number of points N and the diffusion of the simulation, with lower uncertainties at high N or higher diffusivities. This was confirmed by artificially increasing the number of samples taken from the simulations above that actually sampled, which reduced the uncertainty. Presumably the uncertainty would reach a constant value, reflective of the true variance of the tracer field, once a sufficiently high N was reached. Behavior of this type was found in the optimized sampling investigation in section 5.

The different diffusivity simulations for transects with low N were indistinguishable from one another, and so we have not shown those transects (S0 and S2 from UK2

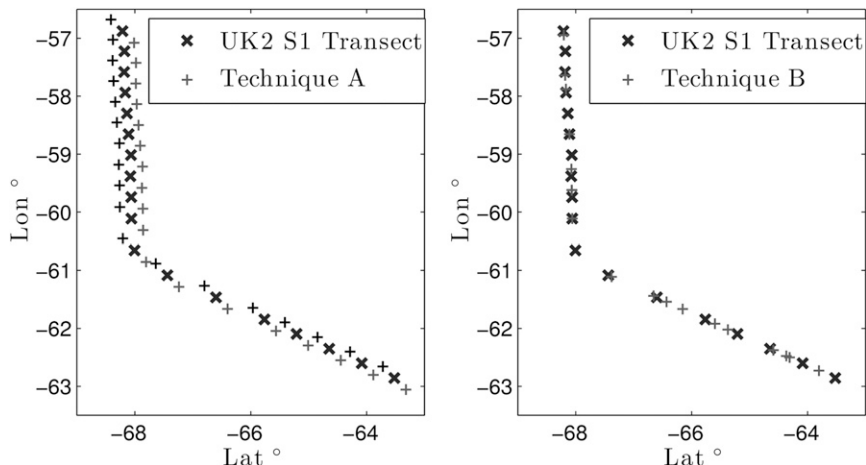


FIG. 9. Examples of tracks used in estimating the uncertainty of the roughness for the UK2 S1 transect. The original cruise transect (large crosses) as well as two from (left) technique A (small plus signs) and one from (right) technique B (small plus signs).

and SR1 from UK2.5). We have included the calculation for S1 from UK2 for reference, but as can be seen, the uncertainty is too high to use the roughness measure to distinguish between the simulations in this case.

Thus, both uncertainty techniques have shown that we can indeed be confident that the roughness measure can be used to distinguish between simulations with varying K_d , provided that the number of samples N is high enough. The dependence of uncertainty on N is explored further in section 5.

4. Roughness calculation results

Figure 11 shows the results of the roughness calculation for all three cruises (black lines) and all six simulations presented so far: $K_d = 0.2, 2, \text{ and } 20 \text{ m}^2 \text{ s}^{-1}$ at $1/50^\circ$ resolution and $K_d = 20, 50, \text{ and } 100 \text{ m}^2 \text{ s}^{-1}$ at $1/20^\circ$ resolution, all with a velocity fraction of 33%. The error bars show uncertainty A, as described in section 3. We choose to use uncertainty A when analyzing the results as it relates directly to the uncertainty in comparing two roughnesses subsampled identically by giving a measure of the uniqueness of the roughness with respect to uncertainties in the exact sampling location. As the observations are taken over time and space, we could choose to set our x axis as either along-track distance (as plotted in Fig. 8) or time before carrying out the roughness calculation. This affects the apparent roughness of the tracer, and so we carried out the roughness calculation for both axes, with the results for along-track distance on the left-hand side of Fig. 11 and the results for time on the right-hand side.

For the US2 cruise, either using along-track distance or time as the x axis for the roughness calculation

resulted in a good match between the $K_d = 20 \text{ m}^2 \text{ s}^{-1}$ simulations and the observations at both resolutions (the good match between the two resolutions shows that our numerical scheme is performing well). The $K_d = 0.2$ and $2 \text{ m}^2 \text{ s}^{-1}$ simulations for this cruise had a large estimated uncertainty as would be expected for different sampling locations missing or hitting peak concentrations associated with thinner tracer streaks. The UK2.5 results again show overlap between the $K_d = 20 \text{ m}^2 \text{ s}^{-1}$ simulations at the two resolutions, and there is more separation between roughness for simulations at different values of K_d . However, the shape of the curves does not match the observations, and as such the simulation with best agreement depends on the curvature p , ranging from between $K_d = 20$ and $50 \text{ m}^2 \text{ s}^{-1}$ and between $K_d = 50$ and $100 \text{ m}^2 \text{ s}^{-1}$. The poor agreement for the UK2.5 results may also be because of the effect of the velocity fraction. As seen in section 2d, the velocity fraction increases downstream, and Fig. 4 shows that the UK2.5 cruise results are better matched by a velocity fraction greater than 33%. A higher diffusivity may match better here because this will transport more tracer downstream, compensating for the low velocity fraction.

We leave investigation into the precise factors that determine the shape of roughness curves to future study, but to assess more completely the effect the velocity fraction has on the roughness of the subsampled tracer, we repeated the roughness calculation as previously for those simulations with variable velocity fraction but with fixed diffusivity $K_d = 20 \text{ m}^2 \text{ s}^{-1}$. Changing the velocity fraction changes the magnitude of the strain felt by the tracer γ , which affects the evolution of the tracer patch, as discussed in section 1. At early times, a larger strain will produce thinner, longer tracer streaks, and so we

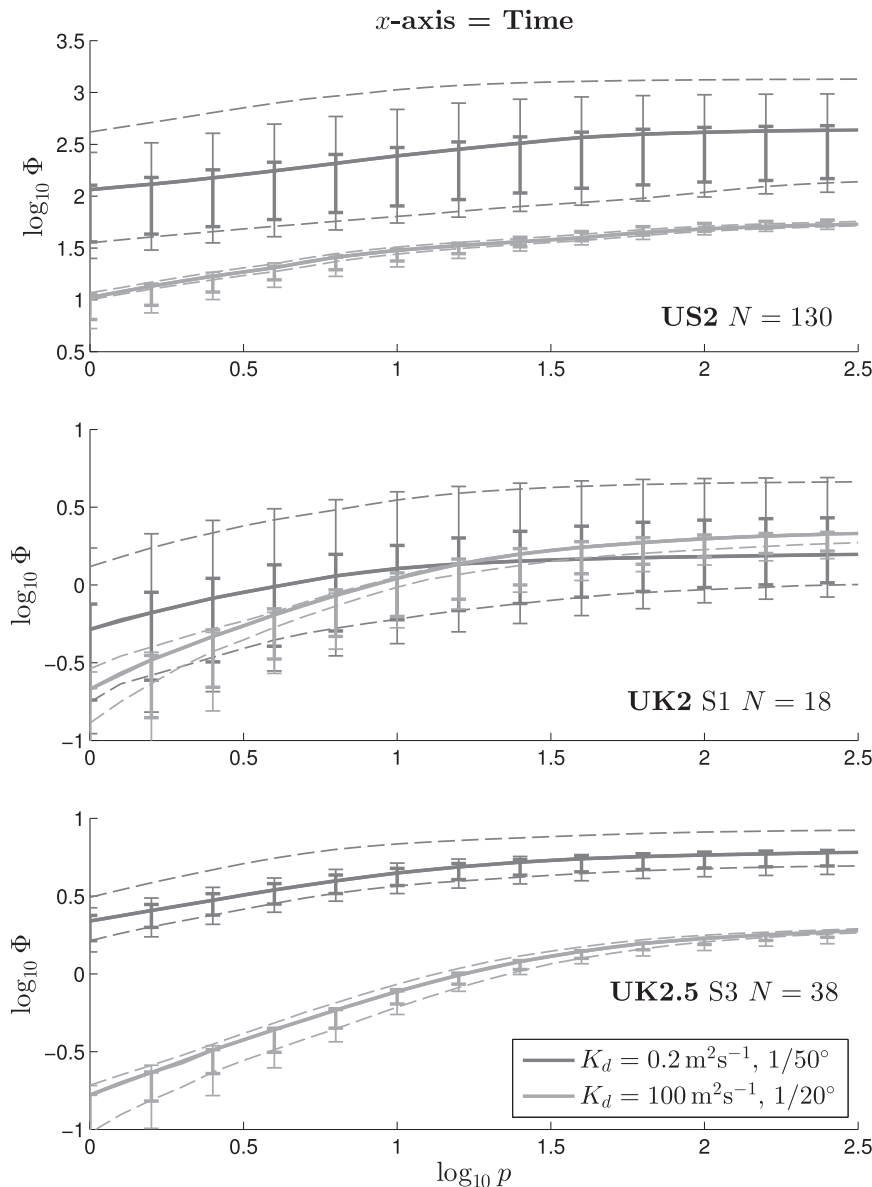


FIG. 10. Roughness as a function of curvature p for the three DIMES cruises, transects as labeled, from two simulations with different diffusivities: $0.2 \text{ m}^2 \text{ s}^{-1}$ (dark gray) and $100 \text{ m}^2 \text{ s}^{-1}$ (light gray). Also shown are uncertainty limits A (dashed lines) and uncertainty limits B (error bars, 75% and 95% confidence intervals); see text for details.

might expect to measure a larger roughness. At later times, a larger strain will merge tracer streaks together quicker, reducing the measured roughness.

Figure 12 shows the results from the calculation in the same form as in Fig. 11. Once again, using either the along-track distance or time as the x axis resulted in qualitatively similar results. The error bars again represent the estimated uncertainty A, calculated as previously, and for each cruise we only show the transect with the largest number of points (US2, UK2 S1, and UK2.5 S3).

For the US2 cruise, increasing the velocity fraction from 28% to 33% increases the roughness, which then decreases at higher velocity fractions. This is consistent with our expectations described above if, at the time of the US2 cruise, which took place roughly 1 yr after the tracer release, the tracer patch is transitioning from a streak-dominated regime to a streak-merging regime. This is also the time scale predicted by the analysis of Garrett (1983); see the discussion in section 1. The observations lie between the 33% and 38% curves, with the uncertainty

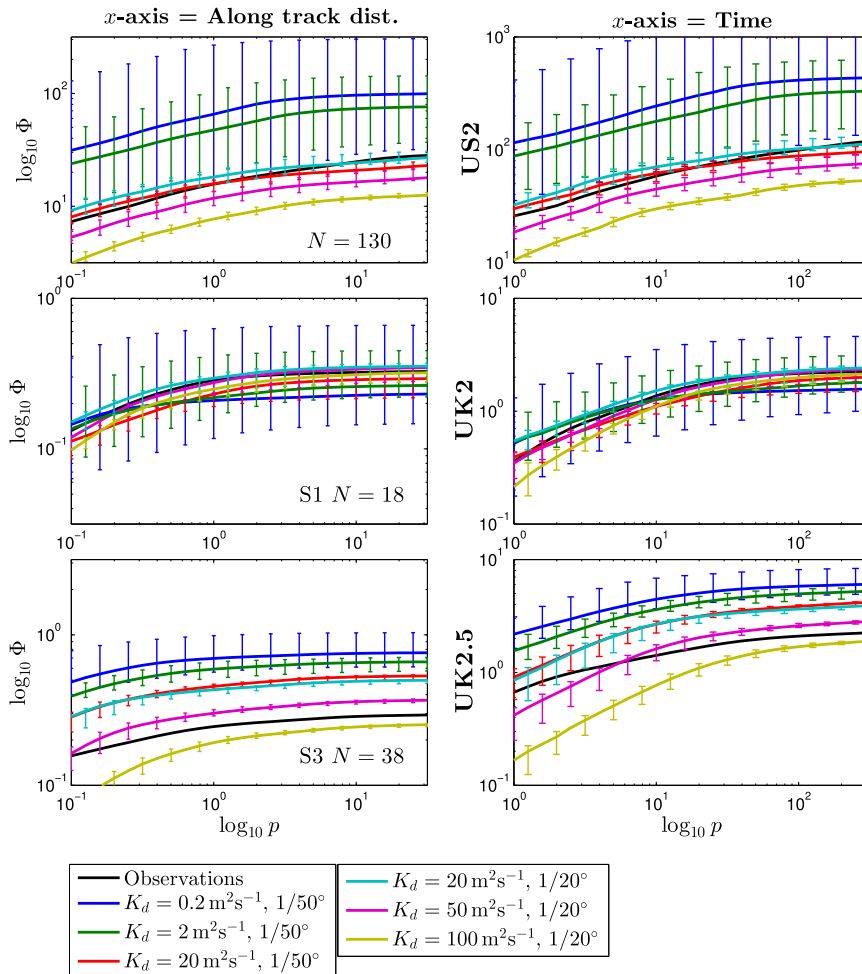


FIG. 11. Roughness Φ vs parabola curvature p for (top to bottom) all three cruises (as labeled) based on observations (black line) and $1/20^\circ$ and $1/50^\circ$ simulations (colored lines, as labeled) for the US2 cruise, the UK2 cruise, S1 transect, and the UK2.5 cruise S3 transect. The x axis for the calculation is either (left) along-track distance or (right) time. See text for details of calculation. Error bars indicate uncertainty A .

interval for the 28% simulation also overlapping, which does not contradict the choice of 33% as the closest fit.

For UK2 S1, we again see a mix of increasing and decreasing roughness with the velocity fraction, and although (as mentioned previously) we do not believe there are enough observation points to make a robust comparison, the 33% simulation again appears to be the closest match to observations.

For UK2.5 S3, all simulations show a decreasing roughness with increasing velocity fraction, showing that the tracer patch is well within the streak-merging stage, and the observations lie closest to the 38% simulation but are also close to the 33% confidence interval for low curvature when the x axis is time. Taken in conjunction with Fig. 11, this implies that the closest match to UK2.5 would be achieved with a velocity fraction 33%–38%

(slightly lowering the roughness of the simulations) and so $K_d = 20\text{--}50 \text{ m}^2 \text{ s}^{-1}$. This can be seen more clearly by comparing the difference between the roughness of simulations and observations $\Delta \log_{10} \Phi$ on a phase diagram of diffusivity and velocity fraction, as shown in Fig. 13. On the phase diagrams, the color of the symbol indicates the magnitude of the difference $\Delta \log_{10} \Phi$, with the scale given by the color bar. The size of the circle indicates the spread in possible results from uncertainty A , as indicated in by the error bars in Figs. 11 and 12; that is, a smaller symbol indicates a more robust difference. Looking at the roughness calculated on either x axis (along-track difference or time), one can see that the minimum difference between simulations and observations is likely found with a velocity fraction 33%–38% and $K_d = 20\text{--}50 \text{ m}^2 \text{ s}^{-1}$.

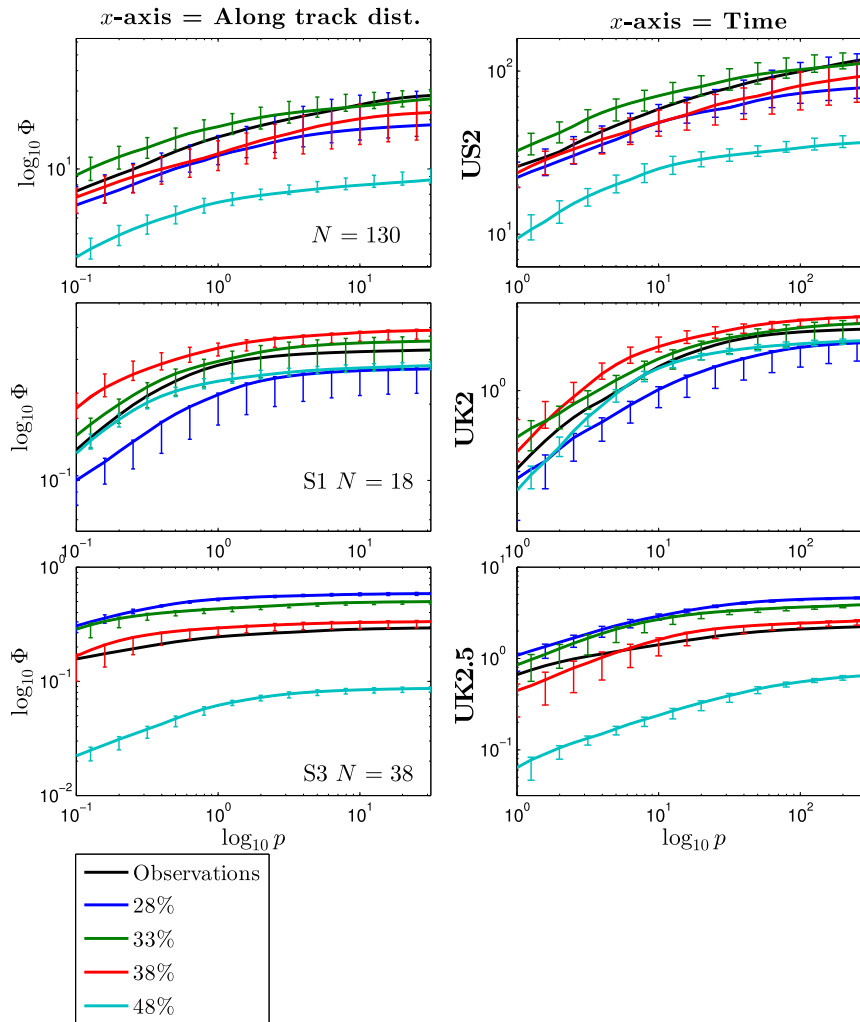


FIG. 12. As in Fig. 11, but with fixed $K_d = 20 \text{ m}^2 \text{ s}^{-1}$, $1/20^\circ$ resolution, and different velocity fractions as labeled.

In summary, these results show that a simulation with 33% and $K_d = 20 \text{ m}^2 \text{ s}^{-1}$ provides the closest match to the “roughness” of the US2 observations. The results also suggest that the tracer measured in the UK2.5 cruise experienced a higher diffusivity ($20\text{--}50 \text{ m}^2 \text{ s}^{-1}$) and a higher velocity fraction (33%–38%), although the roughness calculation does not give good agreement with a single simulation.

5. Using the roughness calculation to optimize sampling choices

Tracer release experiments in the deep ocean, including DIMES, have been designed and executed with the main objective of measuring diapycnal diffusivity. This objective calls for as accurate as possible a horizontal average of the diapycnal distribution of the tracer and thus calls for covering the patch as uniformly

as possible. A second objective has been to measure along-isopycnal dispersion at the mesoscale and requires sampling over a sufficiently large area to delimit the tracer patch. However, arguably the most interesting and least understood mixing processes exposed, albeit imperfectly, by tracer release experiments occur at scales smaller than the mesoscale. Researchers in the field appreciate this aspect and often set aside resources of time, and sometimes instrumentation, to measure the submesoscale features of a tracer patch. The present analysis can help in planning such efforts and, in particular, in deciding on the right balance between station spacing and coverage to estimate submesoscale mixing parameters.

One might expect there to be a balance between sampling at high enough resolution to capture the streakiness of the tracer while sampling across a wide enough region to ensure that the measurement

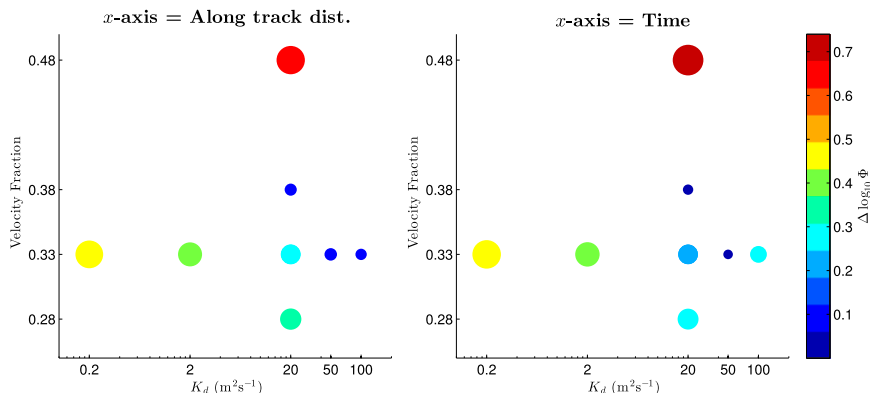


FIG. 13. UK2.5 cruise indicates the magnitude of the difference between the roughness of simulations and observations as plotted in Figs. 11 and 12 ($\Delta \log_{10} \Phi$), as indicated by the color bar, for either (left) along-track difference or (right) time, plotted on a phase diagram of simulation diffusivity (K_d) and velocity fraction. The size of the symbol indicates the spread in possible results from uncertainty A, as indicated by the error bars in Figs. 11 and 12. Both figures indicate that the minimum difference between observations and simulation would be found for a velocity fraction of 33%–38% and $K_d = 20$ – $50 \text{ m}^2 \text{ s}^{-1}$.

is representative of the full field. While testing this systematically for all possible tracks across the full three-dimensional parameter space (longitude, latitude, and time) was beyond the scope of this study, we made a simple test of these ideas as follows: Taking a

30° full-resolution longitudinal transect of the simulation tracer field in February 2010 (the time of the US2 cruise) from the $1/50^\circ$ $K_d = 20 \text{ m}^2 \text{ s}^{-1}$ velocity fraction = 33% simulation (see Fig. 14a), we limited our maximum sampling resolution to $1/50^\circ$ and our maximum number of

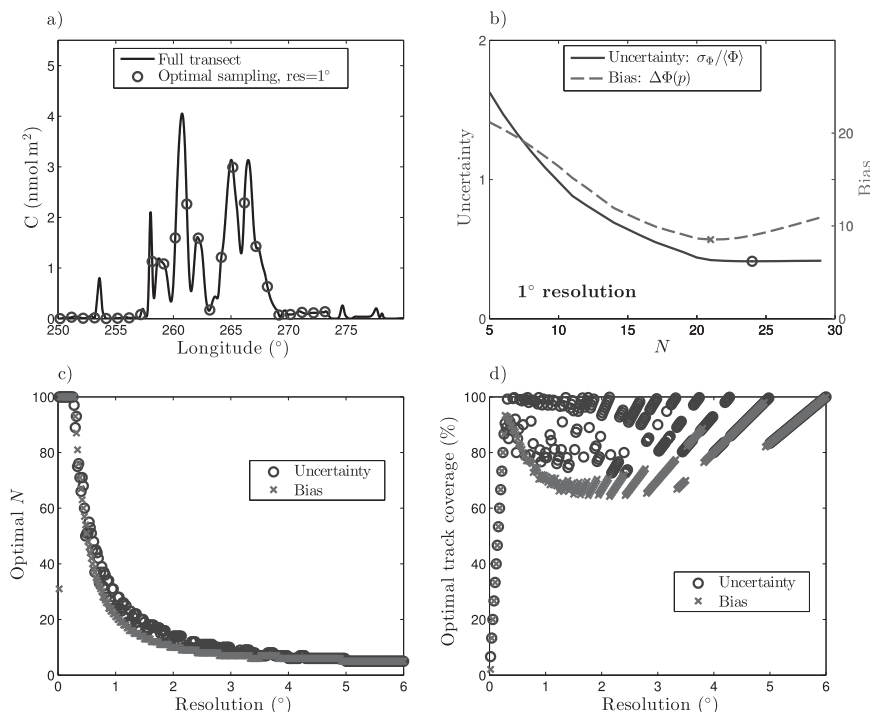


FIG. 14. (a) Transect from $1/50^\circ$ simulation, tracer concentration vs longitude (line), an example of an optimally spaced sampling at 1° (circles). (b) Uncertainty (solid line) and bias (dashed line) against N for tracks sampled at 1° , with minima marked. (c) Optimal N vs sample resolution and (d) optimal sample track coverage vs resolution—most robust w.r.t. the uncertainty (circles) and least biased (crosses).

samples N to 100 and sought to find the optimal sampling technique for a given N .

For each given N and resolution, we took the roughness of all possible tracks covering the transect seen in Fig. 14a, allowing for tracks to be reentrant, scaling the roughness by the length of the transect (the roughness measure is an area and so proportional to the transect length). We then took the standard deviation of the roughness of these tracks, averaged over the curvature p , which gave an estimate of uncertainty, similar to uncertainty A described in section 3. Because the roughness is compared on a log scale, we scaled the standard deviation by the mean of the roughness at each p before taking the mean over p . An example of this estimate of uncertainty for 1° resolution can be seen in Fig. 14b (solid line). We also calculated the mean difference between roughness of the subsampled tracks (again scaled by transect length) and the “true roughness”—the roughness of the full $N = 1500$ $1/50^\circ$ transect, which we called the bias. The rms bias $\langle \Delta\Phi(p)^2 \rangle^{1/2}$, averaged over all possible tracks and then p , for 1° resolution can be seen in Fig. 14b (dashed line). The uncertainty decreases with increasing N , as expected, and then plateaus. The bias has an optimal N , so that too many or too few points can lead to a greater spread in values away from the true roughness. This behavior is found in general at all resolutions, apart from at the lowest resolutions, where the largest N (100) is the most optimal and least biased, as this is the limit we put on the sampling.

For each resolution, the optimal number of points N was defined as that which minimized uncertainty, that is, the circle in Fig. 14b. For reference, we also calculated the N with least bias, that is, the cross in Fig. 14b. One might expect that as the number of samples increases, the true roughness is approached and that the roughness changes little once the sampling resolves some representative scale, perhaps the smallest streak width. We might also expect that the least uncertain roughness might be the least biased; that is, sampling that reveals a measurement closest to the true roughness is also the most robust. Figure 14b shows that this is somewhat the case; the least biased N is also close to the N at which the uncertainty plateaus, the point at which you gain relatively little improvement from increasing N further, although the minima are not exactly collocated.

The optimal N w.r.t. both the uncertainty and the bias can be seen plotted against resolution in Fig. 14c. As can be seen, the optimal N is broadly similar but slightly larger for the uncertainty and decreases with decreasing resolution for both measures. We would expect the curve to continue upward at lower resolutions if we removed the $N = 100$ limit. For 1° resolution, the optimal N w.r.t. uncertainty is 24, and 24 example samples at this

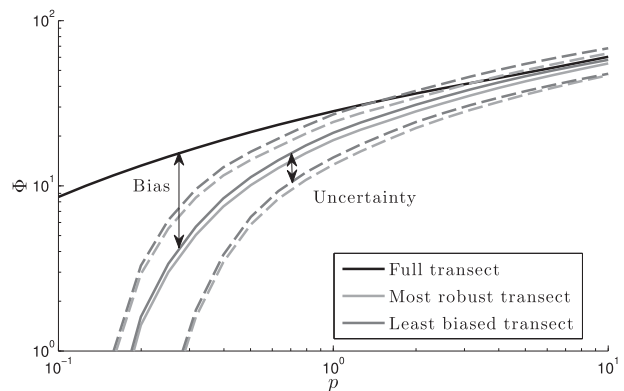


FIG. 15. Roughness of the full transect in Fig. 14a (black line), the most robust resolution track for $N = 20$ (light gray line), and the least biased track (dark gray line). The dashed lines indicate one standard deviation over all possible tracks, that is, the uncertainty.

resolution are seen in Fig. 14a by the circles. The least biased, however, is $N = 21$. It is more desirable that the sampling technique chosen is robust to uncertainty than that it is less biased, given that one will be comparing roughness measures sampled identically, as in Figs. 6 and 7. However, it appears that the least biased N value gives a good indication of where more measurements do not lead to as great gains. The overall most robust and least biased transects were found with the maximum $N = 100$.

The length of the subsampled transect is $N \times$ resolution, and Fig. 14d shows the optimal transect length as a percentage of the full width (30°) against resolution. This rises quickly from $\sim 2\%$ for $1/50^\circ$ to $\sim 100\%$ at $\sim 0.3^\circ$ for both the least biased (crosses) and least uncertain resolutions (crosses). The transect coverage slowly drops to close to 60% at around 2° before rising back to 100% for 6° for the least biased N , with the least uncertain N varying much more widely between full coverage (100%) and similar values to the least biased coverage. Figure 15 shows the roughness curve for the full transect in Fig. 14a (black line), along with both the least biased and most robust roughness for 1° resolution, with uncertainty and the bias labeled. It can be seen here that there is little difference between the roughness or uncertainty of the least biased roughness curve (dark gray line) and the optimal roughness curve (light gray line).

Figures 14c and 14d show that choosing the most robust sampling scheme is not as simple as measuring as many samples as possible; indeed, after a certain point there is little relative gain found from increasing N . This value of N clearly depends on the resolution but also presumably depends on the tracer filament width, determined by the diffusivity and strain experienced by the

TABLE 1. Optimal sampling resolutions for past cruises, assessed by reducing the uncertainty on the $1/50^\circ 20 \text{ m}^2 \text{ s}^{-1}$ simulation.

Cruise	Transect	N	Mean transect resolution	Optimal resolution
US2		130	0.56°	0.56°
UK2	S0	9	0.42°	0.24°
	S1	18	0.38°	0.26°
	S2	11	0.60°	0.34°
UK2.5	SR1	15	0.28°	0.16°
	S3	38	0.20°	0.14°

tracer (see discussion in section 1). The exact relation would be of interest for further study, although the filament width is an unknown in the studies we envisage this technique being applied to.

If one looks at the problem from the perspective of fixed N and looks at how the bias and uncertainty depend on resolution (not shown), there is not such a simple relationship between the least biased and least uncertain resolution, so one cannot use the least biased resolution as a guide to the least uncertain. Additionally, for low N it appears that it is preferable to choose a lower resolution than the highest available in order to cover a larger distance.

A more thorough investigation would be required to discover if these findings are robust and applicable in general, but this method could provide a scheme for designing cruise transects by utilizing simulations validated against previous cruises.

We can also use this concept to assess the suitability of the previous cruise sampling schemes for measuring the roughness of the tracer. For each transect, we repeated the process as described above for the full $1/50^\circ 20 \text{ m}^2 \text{ s}^{-1}$ simulated transect, but keeping N the same as the actual number of observations, and calculated the optimal resolution w.r.t. the uncertainty. This meant that we could not assess resolutions lower than the mean resolution of the observations, as this would have required us to define a wider transect, and we chose to limit the problem to assessing the roughness of the given transect with a fixed N . Thus, we could only assess whether a higher resolution would have been the most robust and not a lower one.

Table 1 shows the mean resolution of the observed transects as well as the optimal resolution, assessed as described previously. We can see that for the US2 cruise, the optimal resolution was equal to or lower than that of the observations, that is, $\geq 0.56^\circ$. However, for the UK2 cruise and UK2.5, the roughness of all transects would have been more robust at higher resolutions.

For our test case (Fig. 14), the optimal resolution increased with increasing N . However, for the actual

cruises, there is a general increase in the optimal resolution with time, resulting in a much higher optimal resolution for the UK2.5 transects than for the US2. This suggests that as the tracer peak values become lower as time passes and the field becomes more diffuse, the roughness is harder to distinguish at low resolutions. Sampling aimed at quantifying roughness in future experiments could be guided by numerical simulations of the sort presented here. In the case of the later cruises in DIMES, it appears that a higher resolution, at the expense of less coverage, would have been optimal, at least for this purpose.

However, as mentioned previously, further work would be required to assess the robustness of this result, especially as we did not take into account the effect of small temporal or cross-transect shifts on the uncertainty for the cruise transects. Additionally, for future cruises, one would need to assess whether other uncertainties expected in the simulation of the tracer itself, introduced by the assumptions of only along-neutral density surface advection, the invariant sea ice field, and so on, which compound with time, would become large enough that such comparisons would not be meaningful.

6. Summary

In this study, we have examined in detail the possible application to ocean tracer measurements of the roughness measure described in Legras et al. (2003, 2005). One important consideration was the robustness of the roughness measure, that is, its sensitivity to the details of the measurements. We assessed the robustness in the measure via two different techniques, either perturbing the sampled track location or altering the sampling resolution. It was found that in general, for either technique, robustness increased with increasing the number of observations.

We used the roughness measure, along with the associated uncertainties, to estimate interior isopycnal diffusivities in the Southern Ocean using altimetry-derived surface velocity fields to advect a conserved 2D tracer field, representing the actual column-integrated tracer, in the offline mode of MITgcm using non-divergent versions of those velocity fields. A Prather advection scheme was used to avoid negative tracer values, although test cases with a second-order central difference scheme showed roughness measurements almost indistinguishable from those found here. The diffusivity estimate was obtained through comparison with a tracer release experiment (DIMES).

When comparing the roughness of simulations with the cruise measurements from the DIMES field campaign, it

was found that some of the cruise tracks did not contain enough measurements to accurately distinguish their roughness because of the large uncertainty in the measure. However, the US2 cruise and the UK2.5 S3 transect both contained sufficient measurements to distinguish the roughness of simulations with different diffusivities.

To obtain the velocity field at the depth of the tracer (neutral density surface $\gamma_n = 27.9 \text{ kg m}^{-3}$), we adjusted the surface velocity fields from altimetry by a constant velocity fraction, under the assumption of an equivalent-barotropic flow. To assess the most suitable value, we utilized 140 RAFOS floats released at the tracer depth in the experiment region during the experimental cruises. The velocities derived from the float paths were compared with the altimetry-derived surface velocities. These results suggested a longitudinally dependent velocity fraction, varying from 33% to 43%, which is comparable to the values found in models (Killworth and Hughes 2002). For simplicity, for the purpose of the simulations, we took the velocity fraction to be constant in latitude and longitude across the computational domain. Additionally, comparisons of simulations with a range of domainwide velocity fractions and an isopycnal diffusivity of $20 \text{ m}^2 \text{ s}^{-1}$ revealed that a velocity fraction close to 33% best matched the center of mass of the subsampled simulation with observations from three separate cruises. This value is at the lower end of that estimated from the float data. We would not necessarily expect these results to agree, as the float data produces many local fractions, whereas the center of mass fraction produces a long-time, domain-scale average. We would expect the local velocities experienced by the floats in this eddy-rich sector of the Southern Ocean to be higher than the large-scale mean flow, but it is reassuring that the values overlap.

Therefore, proceeding with a velocity fraction of 33%, we carried out a range of simulations with a range of horizontal diffusivities and resolutions. Comparison of the roughness measure between simulations and observations implied a best-match isopycnal diffusivity K_s of $20 \text{ m}^2 \text{ s}^{-1}$ for the US2 cruise and a best-match K_s of $20\text{--}50 \text{ m}^2 \text{ s}^{-1}$ for the UK2.5 cruise. The hint that the isopycnal diffusivity might be higher in the Drake Passage region sampled by UK2.5 is interesting, bearing in mind the spatial variation in K_V that has already been found in previous observations and model studies (Watson et al. 2013). The precise relation between K_V and K_s is, of course, not clear. If K_s was given by the Haynes and Anglade (1997) expression $K_V \alpha^2$, then increased K_V would imply increased K_s . But Watson et al. (2013) find a 20-fold variation in K_V , far larger than the twofold variation in K_s hinted at by our results. So, for consistency, there would have to be significant spatial variation in the aspect ratio alpha.

In additional simulations, K_s was kept fixed at $20 \text{ m}^2 \text{ s}^{-1}$, and the velocity fraction was varied. The best-fit velocity fraction was then found to be 33% for the US2 cruise and 33%–38% for the UK2.5 cruise, broadly consistent with the observed longitudinal dependence in the RAFOS-derived measurements.

7. Discussion

The estimates for K_s for the Southern Ocean found here from the comparison of roughness measures between observation and simulations, typically $20 \text{ m}^2 \text{ s}^{-1}$, are significantly larger than the estimates of $1 \text{ m}^2 \text{ s}^{-1}$ from estimated streak width and stretching rates in the North Atlantic. There are several potential reasons for this. One is that the relevant stirring and mixing processes are simply different between the Southern Ocean (at $\sim 1\text{-km}$ depth) and the eastern North Atlantic (at $\sim 300\text{-m}$ depth). Another is that, despite the high resolution of the simulation clearly allowing the development of fine, streaklike structures (see Fig. 5), there are still significant eddy stirring effects unresolved below the $\sim 50\text{-km}$ scale of the altimetric velocity fields. In this case, the K_s required to give the best match with the observations might be representing the effects of the unresolved velocity fields. Any underestimate of the strain due to unresolved small-scale features such as frontogenesis would lead to overestimation of streak widths, and so a smaller value of K_s would be required to match the observed streak widths. Thus, while underestimate of the strain is possible, it is not responsible for the larger than expected K_s . In a turbulent flow with a range of active scales, a tracer field on a given scale to some extent feels the velocity field on smaller scales as a turbulent diffusivity and feels the velocity field on larger scales as a large-scale advection that acts to deform the tracer field. This is nicely illustrated, for example, by Koudella and Neufeld (2004) who consider reaction front propagation in an idealized turbulent flow. On this basis one would certainly expect $K_s < K_h$, which is what we find, bearing in mind the estimate $K_h \approx 700 \pm 260 \text{ m}^2 \text{ s}^{-1}$ for the Pacific sector (e.g., by Tulloch et al. 2014). Thus, although our K_s may not be physically representative of what is experienced by Southern Ocean tracers, it is representative of the unresolved features of the altimetric velocity fields and can be used as a benchmark for comparison with future studies of a similar nature, and this study additionally provides a framework for assessing the next generation of altimetry.

Alongside the uncertainty over what exactly our inferred K_s is representing, there are, of course, many potential shortcomings in our approach. One is the use in the tracer simulations of space and time constant

velocity fractions and diffusivities. The fact that the vertical diffusivity K_V has already been estimated to vary from around $0.2 \times 10^{-4} \text{ m}^2 \text{ s}^{-1}$ in the east Pacific rising to $3.6 \times 10^{-4} \text{ m}^2 \text{ s}^{-1}$ in Drake Passage (Ledwell et al. 2011; Watson et al. 2013) is an indication of potential shortcomings of this assumption. The estimates for K_s in this paper are thus space and time averages and must be interpreted in the light of the numerical simulation setup; that is, the values found here represent how to best match the advection scheme using adjusted surface velocity fields in the given simulation domain to the observations and apply only on the scales investigated here, namely, from 5 to 50 km and on times scales of days, and we would not expect these values to necessarily apply in other situations (as is the nature of all such tracer studies).

A second potential shortcoming is the restriction of the simulation to a single horizontal surface. While it is the case that the diapycnal diffusivity experienced by the DIMES tracer is many orders of magnitude smaller than any expected isopycnal diffusivity and, consistent with this, the observed vertical structure of the tracer resembles a Gaussian profile with a width of only 30 m after 1 yr (Ledwell et al. 2011), vertical structure may not necessarily be neglected and indeed may be crucial in determining horizontal structure (Haynes and Anglade 1997; Smith and Ferrari 2009). A very rough estimate of the effect of vertical shear on the DIMES tracer (using an estimate of the vertical shear itself that assumes an equivalent-barotropic ACC; see discussion in section 2d) implies a horizontal separation of different parts of the tracer patch over the same time period of around 10 km, a ratio of approximately 1:1000. On the same time scale, our entirely 2D simulations show streak widths of approximately 10–100 km (see Figs. 5, 14a). Thus, the implications of the vertical shear for the column integral tracer, which is what we are simulating, are expected to be modest. Taking proper account of the effects of vertical shear would require more accurate information on the vertical structure of the velocity field, and, in the observational context, this is simply not available at present. Using information on the horizontal structure of the flow to deduce some kind of equivalent horizontal diffusivity is a practical compromise that has been used previously in the atmospheric context (e.g., Waugh et al. 1997).

Under these simplifying assumptions, our simulations provided a reasonable approximation of the broad characteristics of the tracer field that allowed for the broad comparison of the tracer field structure with observations via the roughness measure, a tool ideally suited to comparison on a range of scales. The robustness testing of the roughness measure showed it could usefully be used to

distinguish between different diffusivities, given enough samples. Despite the equivalent-barotropic assumption only being strictly valid for circumpolar streamlines, applying a crude domainwide velocity fraction to the surface velocities appeared to be good enough for the region and time period investigated here.

The shortcomings of the use of a velocity fraction, the relatively low-resolution altimetry and the unresolved vertical structure could be addressed by instead using velocity fields taken directly from a high-resolution, three-dimensional numerical model, such as the Southern Ocean State Estimate (SOSE, constrained by observations) and would make an interesting follow-up study.

These results show that while caution needs to be taken with a very small number of observations, the roughness measure is a useful tool for determining small-scale diffusivities when one is sampling widely enough not to be certain of resolving tracer streaks or on time scales on which streak merger is expected to have taken place. Additionally, it could have a wide range of applications in future oceanic tracer experiments. This could include providing estimates of diffusivities at the high spatial resolutions of the next generation of ocean circulation models, allowing for direct comparison or as a basis for model tuning.

Considering the use of the roughness measure as a cruise-planning tool, we found that, for a given resolution, there was an optimal number of measurements for the least uncertainty in the measured roughness. The least biased number of measurements, when compared with the roughness computed at simulation resolution, gave an indication of the point beyond which there was little relative gain in reduced uncertainty. However, for a given number of observations, there was an optimal resolution for the least uncertainty in the measured roughness, which was lower than the least biased resolution. Testing the previous DIMES cruises, we found that the US2 cruise was at or above the most robust resolution but that the UK2 and UK2.5 cruises were sampling at too low resolution. There is the need to carry out a more systematic study to more accurately assess the robustness of these results, but this suggests a strategy that could be utilized for planning future DIMES cruises or similar tracer release experiments. A range of sampling techniques could be tested on a simulation of the tracer, and the technique that produced, on average, the most robust roughness would be adopted for the observational campaign.

Acknowledgments. We thank the U.K. Natural Environment Research Council and the U.S. National Science Foundation for funding the DIMES project.

TABLE A1. Numerical simulations carried out with velocity fraction 33%.

K_d ($\text{m}^2 \text{s}^{-1}$)	k_{num} ($\text{m}^2 \text{s}^{-1}$)	Resolution ($^\circ$)
0.2	0.7 ± 0.1	1/50
2	2.78 ± 0.4	1/50
20	28 ± 4	1/50
20	19.7 ± 0.5	1/20
50	49 ± 1	1/20
100	98 ± 2	1/20

APPENDIX

Further Details of Numerical Simulations

Table A1 gives information on the various numerical simulations carried out with velocity fraction 33%; see section 2c for more information.

REFERENCES

- Abernathy, R., J. Marshall, M. Mazloff, and E. Shuckburgh, 2010: Enhanced isopycnal mixing at steering levels in the Southern Ocean. *J. Phys. Oceanogr.*, **40**, 170–184, doi:10.1175/2009JPO4201.1.
- Adcroft, A., 1995: Numerical algorithms for use in a dynamical model of the ocean. Ph.D. thesis, Imperial College London, 116 pp.
- Balluch, M., and P. Haynes, 1997: Quantification of lower stratospheric mixing processes using aircraft data. *J. Geophys. Res.*, **102**, 23 487–23 504, doi:10.1029/97JD00607.
- Ferrari, R., and K. L. Polzin, 2005: Finescale structure of the T - S relation in the eastern North Atlantic. *J. Phys. Oceanogr.*, **35**, 1437–1454, doi:10.1175/JPO2763.1.
- Garabato, A. C. N., D. P. Stevens, A. J. Watson, and W. Roether, 2007: Short-circuiting of the overturning circulation in the Antarctic Circumpolar Current. *Nature*, **447**, 194–197, doi:10.1038/nature05832.
- Garrett, C., 1983: On the initial streakiness of a dispersing tracer in two- and three-dimensional turbulence. *Dyn. Atmos. Oceans*, **7**, 265–277, doi:10.1016/0377-0265(83)90008-8.
- Haynes, P. H., 2001: Vertical shear plus horizontal stretching as a route to mixing. *From Stirring to Mixing in a Stratified Ocean: Proc. 12th 'Aha Huliko'a Hawaiian Winter Workshop*, Honolulu, HI, University of Hawai'i at Mānoa, 73–79. [Available online at www.soest.hawaii.edu/PubServices/2001pdfs/Haynes.pdf.]
- , and J. Anglade, 1997: The vertical-scale cascade in atmospheric tracers due to large-scale differential advection. *J. Atmos. Sci.*, **54**, 1121–1136, doi:10.1175/1520-0469(1997)054<1121:TVSCIA>2.0.CO;2.
- , and J. Vanneste, 2004: Stratospheric tracer spectra. *J. Atmos. Sci.*, **61**, 161–178, doi:10.1175/1520-0469(2004)061<0161:STS>2.0.CO;2.
- Killworth, P. D., and C. W. Hughes, 2002: The Antarctic Circumpolar Current as a free equivalent-barotropic jet. *J. Mar. Res.*, **60**, 19–45, doi:10.1357/002224002762341230.
- Koszalka, I., J. LaCasce, and K. Orvik, 2009: Relative dispersion in the Nordic Seas. *J. Mar. Res.*, **67**, 411–433, doi:10.1357/002224009790741102.
- Koudella, C., and Z. Neufeld, 2004: Reaction front propagation in a turbulent flow. *Phys. Rev.*, **70E**, 026307, doi:10.1103/PhysRevE.70.026307.
- LaCasce, J., R. Ferrari, J. Marshall, R. Tulloch, D. Balwada, and K. Speer, 2014: Float-derived isopycnal diffusivities in the DIMES experiment. *J. Phys. Oceanogr.*, **44**, 764–780, doi:10.1175/JPO-D-13-0175.1.
- Ledwell, J. R., A. J. Watson, and C. S. Law, 1993: Evidence for slow mixing across the pycnocline from an open-ocean tracer-release experiment. *Nature*, **364**, 701–703, doi:10.1038/364701a0.
- , —, and —, 1998: Mixing of a tracer in the pycnocline. *J. Geophys. Res.*, **103**, 21 499–21 529, doi:10.1029/98JC01738.
- , L. C. St. Laurent, J. B. Girton, and J. M. Toole, 2011: Diapycnal mixing in the Antarctic Circumpolar Current. *J. Phys. Oceanogr.*, **41**, 241–246, doi:10.1175/2010JPO4557.1.
- Legras, B., B. Joseph, and F. Lefevre, 2003: Vertical diffusivity in the lower stratosphere from Lagrangian back-trajectory reconstructions of ozone profiles. *J. Geophys. Res.*, **108**, 4562, doi:10.1029/2002JD003045.
- , I. Pisso, G. Berthet, and F. Lefèvre, 2005: Variability of the Lagrangian turbulent diffusion in the lower stratosphere. *Atmos. Chem. Phys.*, **5**, 1605–1622, doi:10.5194/acp-5-1605-2005.
- Lumpkin, R., and S. Elipot, 2010: Surface drifter pair spreading in the North Atlantic. *J. Geophys. Res.*, **115**, C12017, doi:10.1029/2010JC006338.
- Marshall, J., E. Shuckburgh, H. Jones, and C. Hill, 2006: Estimates and implications of surface eddy diffusivity in the Southern Ocean derived from tracer transport. *J. Phys. Oceanogr.*, **36**, 1806–1821, doi:10.1175/JPO2949.1.
- McKeague, I. W., G. Nicholls, K. Speer, and R. Herbei, 2005: Statistical inversion of South Atlantic circulation in an abyssal neutral density layer. *J. Mar. Res.*, **63**, 683–704, doi:10.1357/0022240054663240.
- Methven, J., and B. Hoskins, 1999: The advection of high-resolution tracers by low-resolution winds. *J. Atmos. Sci.*, **56**, 3262–3284, doi:10.1175/1520-0469(1999)056<3262:TAOHR>2.0.CO;2.
- Okubo, A., 1971: Oceanic diffusion diagrams. *Deep-Sea Res. Oceanogr. Abstr.*, **18**, 789–802, doi:10.1016/0011-7471(71)90046-5.
- Phillips, H., and S. Rintoul, 2000: Eddy variability and energetics from direct current measurements in the Antarctic Circumpolar Current south of Australia. *J. Phys. Oceanogr.*, **30**, 3050–3076, doi:10.1175/1520-0485(2000)030<3050:EVAEFD>2.0.CO;2.
- Polzin, K., and R. Ferrari, 2004: Isopycnal dispersion in NATRE. *J. Phys. Oceanogr.*, **34**, 247–257, doi:10.1175/1520-0485(2004)034<0247:IDIN>2.0.CO;2.
- Prather, M., 1986: Numerical advection by conservation of second-order moments. *J. Geophys. Res.*, **91**, 6671–6681, doi:10.1029/JD091iD06p06671.
- Richardson, L. F., 1926: Atmospheric diffusion shown on a distance-neighbour graph. *Proc. Roy. Soc. London*, **A110**, 709–737, doi:10.1098/rspa.1926.0043.
- , and H. Stommel, 1948: Note on eddy diffusion in the sea. *J. Meteor.*, **5**, 238–240, doi:10.1175/1520-0469(1948)005<0238:NOEDIT>2.0.CO;2.
- Smith, K. S., and R. Ferrari, 2009: The production and dissipation of compensated thermohaline variance by mesoscale stirring. *J. Phys. Oceanogr.*, **39**, 2477–2501, doi:10.1175/2009JPO4103.1.
- , and J. Marshall, 2009: Evidence for enhanced eddy mixing at middepth in the Southern Ocean. *J. Phys. Oceanogr.*, **39**, 50–69, doi:10.1175/2008JPO3880.1.
- Sundermeyer, M. A., and J. F. Price, 1998: Lateral mixing and the North Atlantic Tracer Release Experiment: Observations and

- numerical simulations of Lagrangian particles and a passive tracer. *J. Geophys. Res.*, **103**, 21 481–21 497, doi:[10.1029/98JC01999](https://doi.org/10.1029/98JC01999).
- Tulloch, R., and Coauthors, 2014: Direct estimate of lateral eddy diffusivity upstream of Drake Passage. *J. Phys. Oceanogr.*, **44**, 2593–2616, doi:[10.1175/JPO-D-13-0120.1](https://doi.org/10.1175/JPO-D-13-0120.1).
- Watson, A. J., J. R. Ledwell, M.-J. Messias, B. A. King, N. Mackay, M. P. Meredith, B. Mills, and A. C. N. Garabato, 2013: Rapid cross-density ocean mixing at mid-depths in the Drake Passage measured by tracer release. *Nature*, **501**, 408–411, doi:[10.1038/nature12432](https://doi.org/10.1038/nature12432).
- Waugh, D. W., and Coauthors, 1997: Mixing of polar vortex air into middle latitudes as revealed by tracer-tracer scatterplots. *J. Geophys. Res.*, **102**, 13 119–13 134, doi:[10.1029/96JD03715](https://doi.org/10.1029/96JD03715).
- Young, W., P. Rhines, and C. Garrett, 1982: Shear-flow dispersion, internal waves and horizontal mixing in the ocean. *J. Phys. Oceanogr.*, **12**, 515–527, doi:[10.1175/1520-0485\(1982\)012<0515:SFDIWA>2.0.CO;2](https://doi.org/10.1175/1520-0485(1982)012<0515:SFDIWA>2.0.CO;2).
- Zika, J. D., B. M. Sloyan, and T. J. McDougall, 2009: Diagnosing the Southern Ocean overturning from tracer fields. *J. Phys. Oceanogr.*, **39**, 2926–2940, doi:[10.1175/2009JPO4052.1](https://doi.org/10.1175/2009JPO4052.1).