

Simulating Lagrangian Subgrid-Scale Dispersion on Neutral Surfaces in the Ocean

Daan Reijnders^{1,1}, Eric Deleersnijder^{2,2}, and Erik van Sebille^{1,1}

¹Utrecht University

²Université Catholique de Louvain

November 30, 2022

Abstract

To capture the effects of mesoscale turbulent eddies, coarse-resolution Eulerian ocean models resort to tracer diffusion parameterizations. Likewise, the effect of eddy dispersion needs to be parameterized when computing Lagrangian pathways using coarse flow fields. Dispersion in Lagrangian simulations is traditionally parameterized by random walks, equivalent to diffusion in Eulerian models. Beyond random walks, there is a hierarchy of stochastic parameterizations, where stochastic perturbations are added to Lagrangian particle velocities, accelerations, or hyper-accelerations. These parameterizations are referred to as the 1st, 2nd and 3rd order ‘Markov models’ (Markov-N), respectively. Most previous studies investigate these parameterizations in two-dimensional setups, often restricted to the ocean surface. On the other hand, the few studies that investigated Lagrangian dispersion parameterizations in three dimensions, where dispersion is largely restricted to neutrally buoyant surfaces, have focused only on random walk (Markov-0) dispersion. Here, we present a three-dimensional isoneutral formulation of the Markov-1 model. We also implement an anisotropic, shear-dependent formulation of random walk dispersion, originally formulated as a Eulerian diffusion parameterization. Random walk dispersion and Markov-1 are compared using an idealized setup as well as more realistic coarse and coarsened (50 km) ocean model output. While random walk dispersion and Markov-1 produce similar particle distributions over time when using our ocean model output, Markov-1 yields Lagrangian trajectories that better resemble trajectories from eddy-resolving simulations. Markov-1 also yields a smaller spurious diapycnal flux.

1 Simulating Lagrangian Subgrid-Scale Dispersion on 2 Neutral Surfaces in the Ocean

3 Daan Reijnders¹, Eric Deleersnijder², Erik van Sebille¹

4 ¹Institute for Marine and Atmospheric research Utrecht, Utrecht University, Utrecht, the Netherlands

5 ²Université catholique de Louvain, Institute of Mechanics, Materials and Civil Engineering (IMMC) &
6 Earth and Life Institute (ELI), Louvain-la-Neuve, Belgium

7 Key Points:

- 8 • We create a 3D isoneutral version of the Markov-1 Lagrangian dispersion model,
9 similar to Redi's isopycnal rotation of the diffusion tensor.
- 10 • Dispersion from Markov-1 includes ballistic and diffusive regimes, making trajec-
11 tories more realistic than those from random walk models.
- 12 • Markov-1 produces a much smaller spurious dianeutral diffusivity than Markov-
13 0 (random walk).

Corresponding author: Daan Reijnders, b.j.h.r.reijnders@uu.nl

Abstract

To capture the effects of mesoscale turbulent eddies, coarse-resolution Eulerian ocean models resort to tracer diffusion parameterizations. Likewise, the effect of eddy dispersion needs to be parameterized when computing Lagrangian pathways using coarse flow fields. Dispersion in Lagrangian simulations is traditionally parameterized by random walks, equivalent to diffusion in Eulerian models. Beyond random walks, there is a hierarchy of stochastic parameterizations, where stochastic perturbations are added to Lagrangian particle velocities, accelerations, or hyper-accelerations. These parameterizations are referred to as the 1st, 2nd and 3rd order ‘Markov models’ (Markov-N), respectively. Most previous studies investigate these parameterizations in two-dimensional setups, often restricted to the ocean surface. On the other hand, the few studies that investigated Lagrangian dispersion parameterizations in three dimensions, where dispersion is largely restricted to neutrally buoyant surfaces, have focused only on random walk (Markov-0) dispersion. Here, we present a three-dimensional isoneutral formulation of the Markov-1 model. We also implement an anisotropic, shear-dependent formulation of random walk dispersion, originally formulated as a Eulerian diffusion parameterization. Random walk dispersion and Markov-1 are compared using an idealized setup as well as more realistic coarse and coarsened (50 km) ocean model output. While random walk dispersion and Markov-1 produce similar particle distributions over time when using our ocean model output, Markov-1 yields Lagrangian trajectories that better resemble trajectories from eddy-resolving simulations. Markov-1 also yields a smaller spurious diapycnal flux.

Plain Language Summary

Turbulent eddies stir and disperse material in the ocean. Depending on the resolution of ocean models, these eddies can have length scales that are too small to be resolved explicitly, so they need to be represented by parameterizations. This implies that when particle pathways are computed in Lagrangian simulations, the effect of eddy dispersion also needs to be parameterized. This is traditionally done by adding a random walk on top of successive particle positions. An improvement of this parameterization, referred to as the Markov-1 model, adds random perturbations to particle velocities instead. Dispersion parameterizations have been studied primarily at the surface in two dimensions. In contrast, eddies in the ocean interior predominantly stir and disperse along tilted surfaces of neutral buoyancy. We present a novel three-dimensional formulation of the Markov-1 model and compare it to the random walk model in an idealized setup, as well as using more realistic coarse and coarsened (50 km) ocean model output. Particle distributions produced by both models are similar, but the trajectories produced by Markov-1 better resemble trajectories from simulations that explicitly resolve eddies. Markov-1 also is better able to restrict particle movement to the tilted neutral buoyancy surfaces.

1 Introduction

Turbulent stirring in the ocean disperses tracers and suspended material over time. The eddies, jets, and fronts that characterize this turbulent motion occur at a range of spatial and temporal scales. Since ocean models have a finite resolution, structures with spatial scales of the order of the grid resolution or smaller are not resolved explicitly. Current state-of-the-art global ocean models use nominal $1/48^\circ$ grid resolutions (Su et al., 2018; Fox-Kemper et al., 2019), resolving the mesoscale and part of the submesoscale spectrum. Still, computational constraints limit the simulation length of models at such resolutions to only a few years. Many of the latest generation of Earth system models that are used for CMIP6 use ocean grid resolutions of 1° and $1/4^\circ$ (Hewitt et al., 2020). The models at 1° do not resolve any mesoscale eddies. While the $1/4^\circ$ models are eddy-

64 permitting in parts of the ocean, much higher resolutions are required to resolve the first
 65 baroclinic Rossby radius at higher latitudes, such as in the Southern Ocean, where it is
 66 $\mathcal{O}(10\text{ km})$ (Chelton et al., 1998). Parameterizations of mesoscale eddies therefore remain
 67 vital to ocean modeling.

68 The spreading of tracers due to unresolved eddies is typically parameterized as a
 69 diffusive processes, with the evolution of a tracer concentration C governed by the advection-
 70 diffusion equation:

$$\frac{\partial C}{\partial t} + \bar{\mathbf{u}} \cdot \nabla C = \nabla \cdot (\mathbf{K} \cdot \nabla C), \quad (1)$$

71 where $\bar{\mathbf{u}}$ is the resolved, large-scale velocity, and \mathbf{K} is the diffusivity tensor. This practice
 72 traces back to Boussinesq’s concept of eddy viscosity (Boussinesq, 1877) and G.I.
 73 Taylor’s work on diffusion (Taylor, 1922), and is still ubiquitous in ocean modeling (Fox-
 74 Kemper et al., 2019). Much research has focused on determining and formulating \mathbf{K} in
 75 order to best represent ocean eddies. This includes aspects like the isopycnal or isoneu-
 76 tral orientation of eddies in the ocean interior (Redi, 1982), their advective effect (Gent
 77 & McWilliams, 1990; Griffies, 1998; Haigh et al., 2021), their diffusivity strength (Abernathey
 78 et al., 2013; Griesel et al., 2014; Wolfram et al., 2015; Nummelin et al., 2020), and their
 79 anisotropy (Le Sommer et al., 2011; Bachman et al., 2020).

80 Spreading of tracers and suspended material can also be investigated through the
 81 Lagrangian framework. Through Lagrangian particle simulations, we can study the path-
 82 ways of fluid parcels and suspended material forward and backward in time (van Sebille
 83 et al., 2018). The Lagrangian framework is an especially useful alternative for the Eu-
 84 lerian framework in studying tracer transport when dealing with point sources (Spivakovskaya
 85 et al., 2007; Wagner et al., 2019). Lagrangian simulations use Eulerian ocean model fields
 86 to advect virtual particles. This means that Lagrangian simulations also require param-
 87 eterizations to represent missing dispersion due to the unresolved scales in the Eulerian
 88 input data.

89 The simplest Lagrangian sub-grid scale dispersion model consists of adding a ran-
 90 dom walk onto a particle’s successive locations. It can be shown that this method is con-
 91 sistent with the advection-diffusion equation (1) (Heemink, 1990; Visser, 1997; Spagnol
 92 et al., 2002), hence it is often referred to as ‘diffusion’ in Lagrangian literature. It is the
 93 simplest member of a hierarchy of stochastic parameterizations that is Markovian in na-
 94 ture, and we will refer to it here as *Markov-0* (Berloff & McWilliams, 2002). ‘Markovian’
 95 relates to the Markov property that each successive displacement in the random walk
 96 is independent from the previous.

97 One shortcoming of Markov-0 is that, just like the eddy diffusion approximation
 98 in Eulerian models, it assumes that eddies have infinitely short time scales. Put differ-
 99 ently, it assumes that there is no autocorrelation in the turbulent velocity of the Lagrangian
 100 particles. This assumption does not hold true for mesoscale eddies, which transport La-
 101 grangian particles coherently (Haller & Yuan, 2000; Berloff & McWilliams, 2002). Eddy
 102 coherence leaves an imprint on the Lagrangian velocity autocorrelation, which can be
 103 separated into an exponentially decaying part and an oscillatory part that is the result
 104 of phase differences between the eddies and background flow (Veneziani et al., 2004; Klocker,
 105 Ferrari, & LaCasce, 2012). Due to this imprint, Markov-0 is only accurate at time scales
 106 when the autocorrelation has decayed away, meaning $t \gg T_L$. Here, T_L is the Lagrangian
 107 timescale, equal to the e -folding timescale of the exponential decay of the autocorrela-
 108 tion (LaCasce, 2008). T_L may vary between timescales of a day (Koszalka et al., 2013)
 109 to several weeks (see section 4.2), depending on the characteristics of the ocean domain
 110 at hand. If one is concerned with timescales equal to or smaller than T_L , Markov-0 is
 111 inadequate for parameterizing subgrid-scale dispersion. Regardless, this is often the only

112 scheme for parameterizing subgrid-scale dispersion implemented in community Lagrangian
113 modeling frameworks (van Sebille et al., 2018).

114 Parameterizations higher in the hierarchy of stochastic models add Markovian noise
115 not on particle locations, but on their velocities (Markov-1), accelerations (Markov-2),
116 or even hyper-accelerations (Markov-3) (Sawford, 1991; Rodean, 1996; Griffa, 1996; Berloff
117 & McWilliams, 2002). In doing so, these models are capable of better representing dis-
118 persion at shorter timescales (for which $t \not\gg T_L$), and they can be informed by statis-
119 tical variances in velocity, acceleration, and hyper-acceleration, respectively, as well as
120 the timescales over which the autocorrelations of these quantities decay. Further improve-
121 ments have been formulated that include the looping of particles due to eddy coherence
122 (Reynolds, 2002; Veneziani et al., 2004), as well as the relative dispersion between dif-
123 ferent particles (Piterbarg, 2002).

124 Previous ocean applications of this hierarchy of stochastic models in the Lagrangian
125 framework have been restricted to the horizontal plane (e.g. Haza et al. (2007); Kosza-
126 lka et al. (2013)). However, dispersion through stirring in the interior occurs primarily
127 along sloping surfaces of neutral buoyancy (McDougall, 1987), which are closely related
128 to isopycnals (surfaces of constant potential density). Spivakovskaya et al. (2007) there-
129 fore investigated an isopycnal formulation of the random walk dispersion model. Shah
130 et al. (2011) and Shah et al. (2013) further investigated how the spurious diapycnal flux
131 due to numerical integration can best be minimized.

132 In this study, we discuss, implement, and test an isoneutral formulation of the Markov-
133 1 subgrid-scale dispersion model. We compare the Markov-0 and Markov-1 models when
134 applied to coarse-resolution and coarsened model output data. Specifically, we apply these
135 parameterizations to a channel model of the Southern Ocean, with scales and model set-
136 tings comparable to contemporary global and basin-scale ocean models. This allows us
137 to also assess the spurious dianeutral flux associated with interpolating discrete ocean
138 model output fields.

139 Furthermore, we also consider an anisotropic, shear-dependent formulation of the
140 diffusive/Markov-0 model, formulated by Le Sommer et al. (2011) (*LS* hereafter), which
141 accounts for anisotropy due to shearing and stretching brought about by mesoscale ed-
142 dies. Our aim here is to show how one of the many enhancements proposed to the Eu-
143 lerian diffusion parameterization can be extended to an isoneutral Lagrangian formula-
144 tion.

145 This study focuses on how the isoneutral form of the Markov-1 model, as well as
146 the anisotropic and shear-dependent form of the Markov-0 model, can best be implemented,
147 and to which qualitative differences they lead in the dispersion of Lagrangian particles
148 when compared to a dispersionless case and the isotropic Markov-0 parameterization.
149 We also assess errors of the parameterizations in terms of spurious diffusivities. We aim
150 to use sensible orders of magnitude for the model parameters, but parameter estimation
151 is not our final goal. We are chiefly concerned with formulating an isoneutral form of the
152 Markov-1 model, laying the groundwork for isoneutral subgrid-scale Lagrangian mod-
153 els beyond the isotropic diffusive/Markov-0 parameterization. Higher order stochastic
154 models beyond Markov-1 and extensions thereof will be left out of the scope of this pa-
155 per. These should nonetheless benefit from the ideas discussed here. The advective ef-
156 fect of eddies as captured by the Gent-McWilliams parameterization (Gent & McWilliams,
157 1990) is also not considered.

158 In section 2, we give isoneutral formulations of the Markov-0 and Markov-1 param-
159 eterizations, as well the anisotropic LS formulation of the Markov-0 parameterization.
160 Then, in section 3, we implement and apply these parameterizations to Lagrangian sim-
161 ulations in an idealized situation, and in section 4 to ocean model data output. We as-
162 sess the performance qualitatively and quantitatively. Qualitatively, we compare indi-

163 vidual particle trajectories and the dispersion of particles in a tracer-like patch with the
 164 dispersion in a fine-resolution eddy-resolving model. For the Markov-1 model we also look
 165 at the Lagrangian timescale and associated asymptotic diffusivity, to assess to which extent
 166 we can reproduce these profiles in a fine-resolution setting. Quantitatively, we investigate
 167 the spurious dianeutral diffusivity of the different models. These models should
 168 keep particles restricted to neutral surfaces, but since we use discrete model output, spurious
 169 dianeutral fluxes will occur due to interpolation and other numerical aspects. We
 170 wrap up this study with concluding remarks in section 5.

171 2 Lagrangian isoneutral subgrid-scale models

172 2.1 Markov-0 (diffusion)

173 When we interpret the (Eulerian) advection-diffusion equation (1) as a Fokker-Planck
 174 equation that gives the probability distribution of particle locations over time (Heemink,
 175 1990), this yields a stochastic differential equation (SDE) describing the evolution of La-
 176 grangian particle positions \mathbf{x} as

$$d\mathbf{x} = [\bar{\mathbf{u}}(\mathbf{x}) + \nabla \cdot \mathbf{K}(\mathbf{x})]dt + \mathbf{V}(\mathbf{x}) \cdot d\mathbf{W}(t). \quad (2)$$

177 Here, \mathbf{V} is computed from \mathbf{K} as $\mathbf{K} = \frac{1}{2}\mathbf{V} \cdot \mathbf{V}^T$, meaning that the random noise on the
 178 particle position is proportional to the elements of the diffusivity tensor. This requires
 179 \mathbf{K} to be symmetric and positive-definite. $d\mathbf{W}(t)$ is a vector whose elements correspond
 180 to independent Wiener increments in each respective coordinate direction. These Wiener
 181 increments are normally distributed random variables $\mathcal{N}(0, dt)$ with zero mean and vari-
 182 ance dt (see also Appendix A from Shah et al. (2011)).

183 The $\nabla \cdot \mathbf{K}$ -term in (2) ensures the well-mixed condition (WMC) when the diffu-
 184 sivity tensor is not spatially uniform, and follows the interpretation of (1) as the Fokker-
 185 Planck equation corresponding to the SDE (2) (Heemink, 1990). Simply put, the well-
 186 mixed condition ensures that a particle distribution that is initially mixed, stays mixed.
 187 This condition is also essential for the forward- and backward-in-time formulations of
 188 the model to be consistent. The WMC is extensively discussed by Thomson (1987).

189 The stirring of tracers and dispersion of particles occurs primarily along sloping
 190 neutrally buoyant surfaces (McDougall, 1987). Due to uncertainty about its strength,
 191 spatial variation, and anisotropy of eddy stirring, the eddy diffusivity is often pragmat-
 192 ically chosen to be a homogeneous and isotropic in the neutral plane, with its strength
 193 expressed by the ‘diffusivity’ constant κ (with units $\text{m}^2 \text{s}^{-1}$). Redi (1982) showed that
 194 a diffusivity tensor with these characteristics can be written in geopotential (z -) coordi-
 195 nates in terms of the slopes of the locally neutral plane:

$$\mathbf{K}_{\text{Redi}} = \frac{\kappa}{1 + S_x^2 + S_y^2} \begin{bmatrix} 1 + \epsilon S_x^2 + S_y^2 & -(1 - \epsilon)S_x S_y & (1 - \epsilon)S_x \\ -(1 - \epsilon)S_x S_y & 1 + S_x^2 + \epsilon S_y^2 & (1 - \epsilon)S_y \\ (1 - \epsilon)S_x & (1 - \epsilon)S_y & \epsilon + S_x^2 + S_y^2 \end{bmatrix}, \quad (3)$$

196 where $\epsilon \equiv \kappa_{\text{dia}}/\kappa$ denotes the ratio of dianeutral (adiabatic) to isoneutral diffusivity, and
 197 S_x and S_y are the slopes of the neutral surfaces. When the neutral surfaces are aligned
 198 with the isopycnals, which is the case for an equation of state that is linear in salinity
 199 and potential temperature, these slopes are found as

$$S_x = -\frac{\partial \rho}{\partial x} / \frac{\partial \rho}{\partial z}, \quad S_y = -\frac{\partial \rho}{\partial y} / \frac{\partial \rho}{\partial z}. \quad (4)$$

200 Cox (1987) showed that the diffusivity tensor (3) can be simplified when these slopes
 201 are small (say $|S| = \sqrt{S_x^2 + S_y^2} < 10^{-2}$, which is generally the case in the ocean), and
 202 when ϵ is small compared to unity, so that it reduces to

$$\mathbf{K}_{\text{Redi,approx}} = \kappa \begin{bmatrix} 1 & 0. & S_x \\ 0. & 1 & S_y \\ S_x & S_y & \epsilon + |S|^2 \end{bmatrix}. \quad (5)$$

203 Particle trajectories can then be computed by integrating equation (2). A κ that
 204 is constant in space and time corresponds to the idealized case of homogeneous and sta-
 205 tionary turbulence. The model has the Markovian property that successive spatial per-
 206 turbations $\mathbf{V} \cdot d\mathbf{W}(t)$ are uncorrelated. This in turn causes successive particle velocities
 207 $\mathbf{v} = \frac{\partial \mathbf{x}}{\partial t}$ to be uncorrelated as well, which is unrealistic at short timescales (i.e. $t \not\gg$
 208 T_L) (LaCasce, 2008).

209 2.2 Anisotropic Shear-dependent Markov-0

210 While the tensors (3) and (5) assume that the diffusivity is isotropic and uniform
 211 in the isoneutral plane and time, the transport and stirring by eddies leads to effective
 212 diffusivities that are highly inhomogeneous and anisotropic (McWilliams et al., 1994; Sallée
 213 et al., 2008; Nummelin et al., 2020). In ocean modeling, the effects of eddies on momen-
 214 tum transfer are represented by an eddy viscosity. To account for the inhomogeneous
 215 effect of eddies on the momentum transfer, the eddy viscosity is often parameterized us-
 216 ing the Smagorinsky parameterization (Smagorinsky, 1963), which relates the strength
 217 of the viscosity to the local shear of the flow based on closure of the momentum equa-
 218 tions. This parameterization can also be used for tracer diffusion (Le Sommer et al., 2011),
 219 and has been applied for spatially-dependent (horizontal) random walk dispersion to pa-
 220 rameterize eddies in Lagrangian studies (Nooteboom et al., 2020).

221 Le Sommer et al. (2011) derived an anisotropic and shear-dependent diffusion pa-
 222 rameterization, related to the Smagorinsky parameterization, that also accounts for the
 223 anisotropy in effective diffusivity due to the shearing and stretching effect from the re-
 224 solved scales on the unresolved scales. This parameterization, here abbreviated as *LS*,
 225 was originally proposed for parameterizing the submesoscale using resolved mesoscale
 226 motions, but Nummelin et al. (2020) suggest that the LS parameterization can be ap-
 227 plied to coarser models in which the mesoscale is not resolved.

228 The isoneutral diffusivity tensor from the LS parameterization is given by

$$\mathbf{K}_{\text{LS}} = \frac{h^2}{2}(1 + \delta^2) \begin{bmatrix} p & r & pS_x + rS_y \\ r & q & rS_x + qS_y \\ pS_x + rS_y & rS_x + qS_y & pS_x^2 + qS_y^2 + 2rS_xS_y \end{bmatrix}, \quad (6)$$

229 with $p = \sqrt{r^2 + a^2} + a$ and $q = \sqrt{r^2 + a^2} - a$. Here, $r = \frac{\partial v}{\partial x} + \frac{\partial u}{\partial y}$ is the rate of shear
 230 strain and $a = \frac{\partial u}{\partial x} - \frac{\partial v}{\partial y}$ the rate of normal strain, both in the horizontal plane. The
 231 underlying assumption is that the largest contribution to the isoneutral dispersion falls
 232 within the horizontal plane. The h -term is the horizontal filter size over which the pa-
 233 rameterization acts, and $\delta = [\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}] / \sqrt{r^2 + a^2}$ is a non-dimensional divergence pa-
 234 rameter. The filter size h is related to the size of the grid and it should be tuned through
 235 an $O(1)$, model-dependent constant C that depends on the underlying advection scheme,
 236 so $h^2 = C dx \cdot dy$. A fixed dianeutral diffusivity $\epsilon\kappa$ can be set if we approximate it as
 237 a vertical diffusivity and add it to $\mathbf{K}_{\text{LS},33}$.

238 This parameterization can readily be used in Lagrangian simulations by using \mathbf{K}_{LS}
 239 (6) for the Markov-0 model (2). The parameterization is inherently local, with each of

240 the parameters computed on the location a Lagrangian particle (or grid cell, in the Eu-
 241 lerian case).

242 2.3 Markov-1

243 Next in the hierarchy of stochastic subgrid-scale dispersion models is the Markov-
 244 1 model, also known as the random acceleration or Langevin model (Berloff & McWilliams,
 245 2002). The Markov-1 model adds a random forcing on particle velocities, which should
 246 be proportional to the velocity variance associated to the unresolved eddies. The model's
 247 governing equations are

$$d\mathbf{x} = [\bar{\mathbf{u}}(\mathbf{x}) + \mathbf{u}']dt, \quad (7a)$$

$$d\mathbf{u}' = [-[\boldsymbol{\theta}^{-1}(\mathbf{x})] \cdot \mathbf{u}' + \tilde{\mathbf{a}}(\mathbf{x}, \mathbf{u}')]dt + \mathbf{b} \cdot d\mathbf{W}(t). \quad (7b)$$

248 The particle location \mathbf{x} evolves through integration of the resolved mean flow $\bar{\mathbf{u}}(\mathbf{x})$ and
 249 a turbulent fluctuation \mathbf{u}' . This fluctuation evolves through the stochastic differential
 250 equation (7b). The deterministic part of this equation consists of two terms: a fading-
 251 memory term, which ensures an exponential decay in the autocorrelation of the parti-
 252 cle's velocity, regulated through the fading-memory time tensor $\boldsymbol{\theta}$ (with time as its di-
 253 mension), and a drift correction term $\tilde{\mathbf{a}}$, which ensures the well-mixed condition. The
 254 stochastic forcing term consists of the Wiener increment $d\mathbf{W}$ and the random forcing
 255 is related as $\mathbf{b}\mathbf{b}^T = 2\boldsymbol{\sigma}\boldsymbol{\theta}^{-1}$. Here, $\boldsymbol{\sigma}$ is the velocity variance tensor, which relates to the
 256 strength of the velocity fluctuations \mathbf{u}' that are to be simulated:

$$\sigma_{ij} = \langle u'_i u'_j \rangle, \quad (8)$$

257 where the angled brackets denote ensemble averages over Lagrangian trajectories.

258 The drift correction term is given by

$$\tilde{a}_i = \frac{1}{2} \frac{\partial \sigma_{ik}}{\partial x_k} - \frac{\sigma_{im}}{2} (\bar{u}_k + u'_k) \frac{\partial [\sigma^{-1}]}{\partial x_k} u'_j. \quad (9)$$

259 See Berloff and McWilliams (2002) for further details and derivations.

260 The nonsingular velocity variance tensor $\boldsymbol{\sigma}$ and the fading-memory time tensor $\boldsymbol{\theta}$
 261 are the free parameters in the Markov-1 model. They can be estimated from velocity fields
 262 in which the turbulent velocity is resolved. For the velocity variance, this is clear from
 263 equation (8). The velocity variance tensor may be anisotropic, inhomogeneous in space,
 264 and evolving over time. An obvious and useful simplification is to use a single, average
 265 velocity variance parameter ν^2 that characterizes the entire system (Koszalka et al., 2013).
 266 In this case $\boldsymbol{\sigma}$ is diagonal with its values equal to ν^2 . Alternatively, the velocity variance
 267 may be a probability distribution rather than an average value in order to account for
 268 the variance in ν^2 found within different regions of a fluid domain (Berloff & McWilliams,
 269 2003).

270 The fading-memory time tensor $\boldsymbol{\theta}$ determines the strength of the exponential de-
 271 cay of the turbulent velocity \mathbf{u}' . The elements of $\boldsymbol{\theta}$ are found by integrating the Lagrangian
 272 autocorrelation $R_{ij}(\tau)$ over all time lags τ :

$$\theta_{ij} = \int_0^\infty R_{ij}(\tau) d\tau, \quad (10)$$

273 where

$$R_{ij}(\tau) = \langle u'_i(t) u'_j(t + \tau) \rangle / (\langle u_i'^2 \rangle \langle u_j'^2 \rangle)^{1/2}. \quad (11)$$

274 Like the turbulent velocity, the Lagrangian autocorrelation exhibits spatial variation in
 275 the ocean, and its anisotropy can be strongly affected by the presence of jets (Griesel et
 276 al., 2010). Still, it is also useful to characterize the fading-memory time of the entire sys-
 277 tem by an average value. In a homogeneous, stationary situation without boundary ef-
 278 fects, the fading memory tensor is diagonal with its values equal to the Lagrangian in-
 279 tegral time T_L .

We characterize the dispersion of particles by the single-particle (sometimes called ‘absolute’) dispersion tensor:

$$D_{ij}(t, \mathbf{x}(\mathbf{0})) = \langle (x_i(t) - x_i(0))(x_j(t) - x_j(0)) \rangle. \quad (12)$$

280 Berloff et al. (2002) note that while the dispersion tensor in the ocean may evolve in a
 281 nonlinear manner, it can be described by different power laws at intermediate timescales:

$$D_{ii}(t) \sim t^{\alpha_{ii}}. \quad (13)$$

282 Single-particle dispersion in the ocean is initially ballistic, meaning $D(t) \sim t^2$ for $t \ll$
 283 T_L . At longer time-scales, it becomes approximately linear in time, i.e. $D(t) \sim t$. Since
 284 such behavior is equivalent to that of a diffusive process, this is also referred to as the
 285 diffusive limit. Unsurprisingly, dispersion simulated by the Markov-0 model is purely dif-
 286 fusive. The Markov-1 model, however, is able to also simulate the initially ballistic be-
 287 havior of particles dispersion. For time scales longer than those characterized by the el-
 288 ements of $\boldsymbol{\theta}$, the Markov-1 model essentially behaves diffusively (Rodean, 1996). In this
 289 limit, assuming homogeneity, stationarity, and absence of boundary effects, we can re-
 290 late the absolute diffusivity, velocity variance and Lagrangian integral time as

$$\nu^2 T_L = \kappa. \quad (14)$$

291 At intermediate time-scales, α_{ii} can take on other values than 1 and 2, which is referred
 292 to as anomalous dispersion (LaCasce, 2008). While the dispersion regimes other than
 293 the ballistic and diffusive cannot be simulated by Markov-1, the higher order Markov-
 294 2 and Markov-3 models, or modifications of Markov-1 are able to account for such be-
 295 havior, such as the oscillatory component of the Lagrangian autocorrelation (Berloff &
 296 McWilliams, 2002; Reynolds, 2002; Veneziani et al., 2005). However, we limit ourselves
 297 here to Markov-1 for its simplicity, as each modification or higher model in the hierar-
 298 chy includes more free parameters.

299 We now formulate an ad-hoc three-dimensional, isoneutral version of the Markov-
 300 1 model in the case of homogeneous and stationary turbulence without boundary effects.
 301 First, we assume that the turbulent velocity perturbations should remain primarily re-
 302 stricted to the local neutral plane, in which it is isotropic. In isoneutral coordinates this
 303 yields

$$\boldsymbol{\sigma}_{\text{iso}} = \begin{bmatrix} \nu^2 & 0 & 0 \\ 0 & \nu^2 & 0 \\ 0 & 0 & \eta\nu^2 \end{bmatrix}, \quad \text{and} \quad \boldsymbol{\theta}_{\text{iso}} = \begin{bmatrix} T_L & 0 & 0 \\ 0 & T_L & 0 \\ 0 & & \varepsilon T_L \end{bmatrix}. \quad (15)$$

304 Assuming there is some dianeutral velocity perturbation $\nu_{\text{dia}}^2 (\ll \nu^2)$, we define $\eta \equiv \nu_{\text{dia}}^2 / \nu^2$.
 305 Similarly, assuming a separate dianeutral Lagrangian integral time $T_{L,\text{dia}}$, we define $\varepsilon \equiv$
 306 $T_{L,\text{dia}} / T_L$.

307 Then, we simply transform $\boldsymbol{\sigma}$ and $\boldsymbol{\theta}$ from isoneutral coordinates to geopotential co-
 308 ordinates in analogy to Redi’s formulation of the isoneutral diffusivity tensor (Redi, 1982).
 309 This yields:

$$\boldsymbol{\sigma}_{\text{geo}} = \frac{\nu^2}{1 + S_x^2 + S_y^2} \begin{bmatrix} 1 + \eta S_x^2 + S_y^2 & -(1 - \eta) S_x S_y & (1 - \eta) S_x \\ -(1 - \eta) S_x S_y & 1 + S_x^2 + \eta S_y^2 & (1 - \eta) S_y \\ (1 - \eta) S_x & (1 - \eta) S_y & \eta + S_x^2 + S_y^2 \end{bmatrix}, \quad (16)$$

310 and

$$\boldsymbol{\theta}_{\text{geo}} = \frac{T_L}{1 + S_x^2 + S_y^2} \begin{bmatrix} 1 + \varepsilon S_x^2 + S_y^2 & -(1 - \varepsilon) S_x S_y & (1 - \varepsilon) S_x \\ -(1 - \varepsilon) S_x S_y & 1 + S_x^2 + \varepsilon S_y^2 & (1 - \varepsilon) S_y \\ (1 - \varepsilon) S_x & (1 - \varepsilon) S_y & \varepsilon + S_x^2 + S_y^2 \end{bmatrix}. \quad (17)$$

311 Note that in order for these tensors to be nonsingular, η and ε should be nonzero, mean-
 312 ing that $\boldsymbol{\sigma}_{\text{geo}}$ and $\boldsymbol{\theta}_{\text{geo}}$ have nonzero diapycnal contributions. We thus have to specify
 313 η and ε in a way such that they are small enough to prevent large dianeutral excursions.

314 While the diffusivity tensor (3) can be simplified (5) by the assumption that slopes
 315 are small, this assumption cannot be applied to the tensors $\boldsymbol{\sigma}_{\text{geo}}$ (16) and $\boldsymbol{\theta}_{\text{geo}}$ (17), since
 316 the terms that are scaled out in the small-slope assumption become dominant in the in-
 317 verses of $\boldsymbol{\sigma}_{\text{geo}}$ and $\boldsymbol{\theta}_{\text{geo}}$, which are used in (7), (9) and when computing \mathbf{b} .

318 A key assumption of Redi's diffusivity tensor \mathbf{K}_{redi} is that the neutral surfaces are
 319 stationary and locally flat. 'Locally' here is related to the length scale associated to the
 320 displacement of a particle over one timestep. The assumption is that when a particle is
 321 advected, the neutral slope at the particle's original location \mathbf{x}_0 at time t_0 is approxi-
 322 mately equal to the neutral slope at the particle's new location \mathbf{x}_1 after a timestep dt .
 323 Any difference in the orientation of the neutral surface over successive timesteps will lead
 324 to some dianeutral movement, but as long as neutral surfaces are locally flat, this dia-
 325 neutral movement is limited and the new local slopes are used for computing the next
 326 neutral displacement.

327 For Markov-1, the situation is more complicated. In this case, the stochastic ve-
 328 locity perturbations of a particle at time t_0 and location \mathbf{x}_0 are oriented parallel to the
 329 local neutral plane. However, since particle velocities (7b) are autocorrelated, the cur-
 330 vature of the neutral surface at a particle's initial location \mathbf{x}_0 can influence a particle's
 331 velocity over several timesteps, as the particle is displaced away from \mathbf{x}_0 . This influence
 332 decays exponentially with the e -folding timescale εT_L . Thus if a neutral surface curves
 333 at spatial scales that are similar to or smaller than the length scale L over which a par-
 334 ticle travels within the timescale εT_L , the signal of the turbulent velocity perturbation
 335 at t_0 influences the particle's net turbulent velocity, causing a dianeutral velocity con-
 336 tribution, and therefore a dianeutral displacement. To combat this dianeutral movement,
 337 the Lagrangian autocorrelation in the dianeutral direction should rapidly decay away at
 338 each timestep. Put differently, εT_L should be so small that a neutral surface can be ap-
 339 proximated as flat over the length scale L . While εT_L should be larger than zero to avoid
 340 singularity of $\boldsymbol{\theta}$, one ad-hoc workaround to rapidly extinguish the signal of velocity per-
 341 turbations at previous timesteps is to set

$$\varepsilon T_L = dt. \quad (18)$$

342 This workaround comes at a price: if the neutral surface curves, the Lagrangian decor-
 343 relation of an initially isoneutral signal may occur more quickly than is prescribed by
 344 $\boldsymbol{\theta}$, since the initially isoneutral perturbation becomes dianeutral over time, which causes
 345 it to decay rapidly due to (18). This effect increases when more curvature is covered by
 346 a Lagrangian particle as it moves in space and time. Properly retaining autocorrelations
 347 on curved surfaces is a complicated matter (Gaspari & Cohn, 1999), so here we take a
 348 pragmatic approach by assuming that the change in isoneutral curvature is small enough
 349 for practical use to warrant our ad-hoc formulation of a three-dimensional Markov-1 model.

350 Finally, when ε is fixed by (18), η can be chosen in such a way that the effective
 351 dianeutral diffusivity in the limit $t \gg T_L$ is controlled as:

$$\epsilon\kappa = \eta\nu^2 \varepsilon T_L. \quad (19)$$

352 This means that if we indeed assume homogeneity, stationarity, and a lack of boundary
 353 effects, the parameters necessary for Markov-1 model may be determined by specifying
 354 the Lagrangian integral time T_L and an effective diffusivity κ , which fix ν^2 through (14),
 355 and by specifying the dianeutral diffusivity ratio ϵ , fixing ε and η (through (18) and (19)).

356 **3 Numerical implementation**

357 **3.1 Discretization**

To use the Markov-0 and Markov-1 models numerically, we need to discretize SDEs
 (2) and (7). The simplest SDE discretization is Euler-Maruyama scheme, which can be
 seen as a stochastic version of the Euler-forward scheme. Given a general stochastic dif-
 ferential equation

$$d\mathbf{X} = \alpha(\mathbf{X}, t)dt + \beta(\mathbf{X}, t)d\mathbf{W}(t), \quad (20)$$

358 with $\alpha(\mathbf{X}, t)$ signifying the deterministic forcing strength and $\beta(\mathbf{X}, t)$ the stochastic forc-
 359 ing strength, the Euler-Maruyama scheme approximates the true solution for X by the
 360 Markov chain Y as

$$Y_{n+1}^k = Y_n^k + \alpha^k \Delta t + \sum_{j=1}^m \beta^{k,j} \Delta W^j, \quad (21)$$

361 where superscripts denote the k -th component of the m -dimensional vectors \mathbf{X} and \mathbf{Y}
 362 and subscripts denote discrete time indices. $\Delta\mathbf{W}$ is an m -dimensional vector of discretized
 363 Wiener increments, which are normally distributed, $\mathcal{N}(0, \Delta t)$, with zero mean and vari-
 364 ance Δt . See Kloeden and Platen (1999) or Iacus (2008) for more details on numerical
 365 SDE schemes. The expressions for α and β can be readily identified in (2) and (7b). In
 366 the case of Markov-1, an additional numerical integration is necessary for (7a). For con-
 367 sistency with the Euler-Maruyama scheme, this can simply be the Euler-Forward dis-
 368 cretization.

369 We implemented the Markov-0 and Markov-1 schemes in the *Parcels* Lagrangian
 370 framework (Delandmeter & van Sebille, 2019). All Lagrangian simulations in this pa-
 371 per are carried out with *Parcels* (van Sebille et al., 2020).

372 **3.2 Idealized test case**

373 We assess the validity of the isoneutral subgrid-scale models using an idealized, sta-
 374 tionary density field for which we can compute the isoneutral slopes exactly, assuming
 375 that here the neutral surfaces align with the isopycnals. We do not consider any actual
 376 fluid dynamical setup, meaning there is no background flow ($\bar{\mathbf{u}} = 0$). This three-dimensional
 377 idealized test case is an extension of the two-dimensional test case from Shah et al. (2011),
 378 and is given by

$$\rho(x, y, z) = \rho_0 \left[1 - \frac{N^2 z}{g} + A_x \sin(k_x x) + A_y \sin(k_y y) \right], \quad (22)$$

379 with ρ_0 a reference density, N the Brunt-Vaisala frequency, g the gravitational accel-
 380 eration, A the amplitude of the wave-like neutral surfaces, and k their wavenumber (sub-
 381 scripts denoting direction). The z -coordinate of the neutral surface corresponding to the
 382 density ρ^* is then found as

$$z_{\text{iso}}(\rho^*, x, y) = \frac{g}{N^2} \left[1 - \frac{\rho^*}{\rho_0} + A_x \sin(k_x x) + A_y \sin(k_y y) \right]. \quad (23)$$

We use a similar choice of parameters as (Shah et al., 2011), which is representative of the large-scale ocean:

$$\begin{aligned} \rho_0 &= 1025 \text{ kg m}^{-3}, & N^2 &= 1 \times 10^{-5} \text{ s}^{-2}, & g &= 10 \text{ m s}^{-2}, \\ A_x &= 1 \times 10^{-3}, & A_y &= 1.1 \times 10^{-3}, & k_x = k_y &= \frac{2}{\pi} \times 1 \times 10^{-5} \text{ m}^{-1}. \end{aligned} \quad (24)$$

This choice of parameters leads to a maximum slope of $\max(|S|) \approx 10^{-3}$, which is a typical value for neutral slopes in the ocean, and for which the small-slope approximation (5) is valid (Mathieu & Deleersnijder, 1998). Although we may not use this approximation in the Markov-1 model due to singularity, as explained in section 2.3, it is useful to compare the small-slope approximation of Markov-0 (5) with its full formulation (3).

3.3 Spurious diffusivity

We can compare the spurious dianeutral diffusivities induced by numerical errors in the discretized Markov-0 and Markov-1 models. We limit this analysis for brevity and refer the reader to Shah et al. (2011) for an extensive discussion of numerical errors introduced by Markov-0. The models considered here have an equivalent effective diffusivity (14) in the limit $t \gg T_L$. We initialize 12,800 particles on a neutral surface, using a regular xy -grid, with the z -coordinates computed from (23) and $\rho^* = 1027.5 \text{ kg m}^{-3}$. We found that results are insensitive to adding more particles. We take into account the periodic topology of the neutral surfaces to make sure crests and troughs are sampled evenly. Then, we numerically integrate the particles for 90 days using several choices of integration timestep Δt . The particle displacements are computed by using the exact density field (22) and its spatial derivatives. From the vertical departure of the particles from the neutral surfaces, we can compute an effective spurious vertical diffusivity,

$$\kappa_{z,\text{spurious}} = \frac{\langle (z - z_{\text{iso}})^2 \rangle}{2T_{\text{int}}}, \quad (25)$$

where the angled brackets denote a particle ensemble average and T_{int} is the total integration time. We use this as an approximation of the spurious dianeutral diffusivity introduced by the numerical approximation of (20).

In the Markov-0 model, we set $\kappa = 1000 \text{ m}^2 \text{ s}^{-1}$ and $\epsilon = 0$, such that the only dianeutral movement of particles is due to numerical errors. We test both \mathbf{K}_{Redi} and $\mathbf{K}_{\text{Redi,approx}}$. We cannot test Markov-0 using \mathbf{K}_{LS} , as we do not consider a fluid setup with flow from which its parameters are computed.

For Markov-1, we use a value of $T_L = 20$ days, and we determine $\nu^2 = \kappa/T_L = 5.79 \times 10^{-4} \text{ m}^2 \text{ s}^{-2}$, so that the effective isoneutral diffusivity in the diffusive limit equals the one used for Markov-0 (see (14)). We also need to specify the nonzero dianeutral fading-memory time and velocity variance in the Markov-1 model to guarantee that (16) and (17) are nonsingular. To ensure rapid decorrelation of \mathbf{u}' in the local dianeutral direction, we set $\varepsilon = \frac{\Delta t}{T_L}$ (18). In order to avoid θ being singular, we also need a nonzero η . However, here we are interested in the dianeutral movement induced by numerical errors, rather than what is specified by the algorithm. Here we need to make a trade-off: we found that if η gets very small ($\eta \lesssim 10^{-10}$), this causes instabilities due to the multiplication of very small and very large terms (inverses of η) when computing the drift correction term (9). This may not necessarily lead to a spurious diapycnal diffusivity, but we found that it can lead to particle accumulation in specific areas. We choose $\eta =$

421 10^{-8} ; a value for which we do not observe noticeable instabilities with the drift correc-
 422 tion term. For small choices of dt , this choice of η will cause the ‘spurious’ diapycnal dif-
 423 fusivity to equal the expected diapycnal diffusivity (computed using (19)), while for larger
 424 timesteps the spurious diffusivity is dominated by numerical errors.

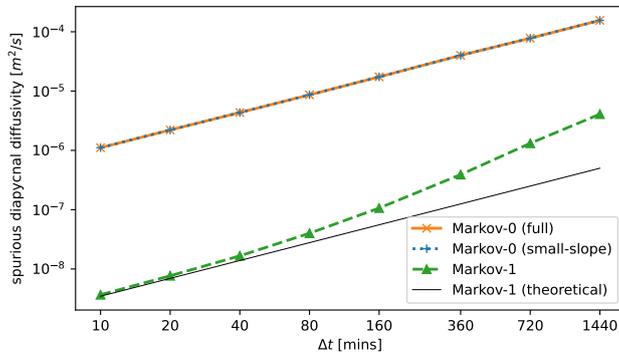


Figure 1. Spurious dianeutral diffusivities after 90 days in the Markov-0 model (with and without the small-slope approximation (5)), and the Markov-1 model, using several timesteps Δt . For Markov-1, we also plot the diapycnal diffusivity that is theoretically imposed through our choice of η . The Markov-1 model has a much smaller spurious dianeutral flux for each timestep. Using the small-slope approximation for Markov-0 leads to negligible differences in the spurious diapycnal diffusivity.

425 Figure 1 shows that the spurious dianeutral diffusivity after 90 days of integration
 426 is much smaller for Markov-1 than for Markov-0. Recall that both use the same Euler-
 427 Maruyama discretization scheme (21). The difference in dianeutral diffusivity is due to
 428 the fact that the expected turbulent displacement for a single timestep in Markov-1 is
 429 $E(\|\mathbf{u}'\| \Delta t) = \nu \Delta t$ (see (7)), while that in Markov-0 is $E(\mathbf{V} \cdot d\mathbf{W}) = \sqrt{2\kappa\Delta t}$, (see (2))
 430 where E denotes the expected value and $\|\cdot\|$ the vector norm. The turbulent excu-
 431 sion of Markov-1 in one timestep is therefore much smaller than that of Markov-0 over
 432 the range of Δt investigated here, and thus Markov-1 introduces less dianeutral move-
 433 ment as the neutral surfaces curve. Also note that over this range of Δt and with our
 434 choice of κ , ε and η , as dt increases, the diapycnal diffusivity diverges from the theoret-
 435 ical diapycnal diffusivity imposed through η . This divergence is caused by numerical er-
 436 rors, meaning these start dominating for the larger values in our range of dt . We con-
 437 clude that Markov-1 generally performs significantly better in keeping particles on ideal-
 438 ized neutral surfaces. Note that the spurious diapycnal diffusivity depends on the slopes
 439 of the idealized neutral surfaces, determined by A_x , A_y , k_x , and k_y (Shah et al., 2011).

440 Several studies propose the use of higher order numerical schemes to reduce the spu-
 441 rious dianeutral flux resulting from numerical integration (Shah et al., 2011; Gräwe, 2011;
 442 Gräwe et al., 2012) or the use of adaptive time-stepping methods (Shah et al., 2013). While
 443 higher order schemes, such as the first order Milstein scheme (see Kloeden & Platen, 1999),
 444 indeed perform better in the idealized configuration, we find that this improvement is
 445 negligible when applied to discrete ocean model data using commonly used spatial and
 446 temporal output resolutions (see section 4.1), and a Lagrangian timestep of 40 minutes,
 447 indicating that the error introduced by interpolating Eulerian data dominates that of
 448 the numerical method.

449

3.4 Well-mixedness

450

451

452

453

454

455

456

457

458

459

460

The equations for the Markov-1 model, including the drift-correction term (9), are rigorously derived in Berloff and McWilliams (2002). However, since we create an ad-hoc adaption of this model for use in three-dimensional isoneutral situations, it is important that we verify whether we did not inadvertently violate the well-mixed condition. Rather than rigorously proving the WMC, we take a pragmatic approach here and visually inspect particle distributions to see if we can find spurious accumulation. We choose pragmatism over rigor of proof, because in applications with discrete Eulerian ocean model output, Lagrangian simulations with Markov-0 and Markov-1 are both affected by numerical errors due to discretization and interpolation. These numerical aspects will violate the WMC in any case, hence a pragmatic visual verification of the WMC satisfies our needs.

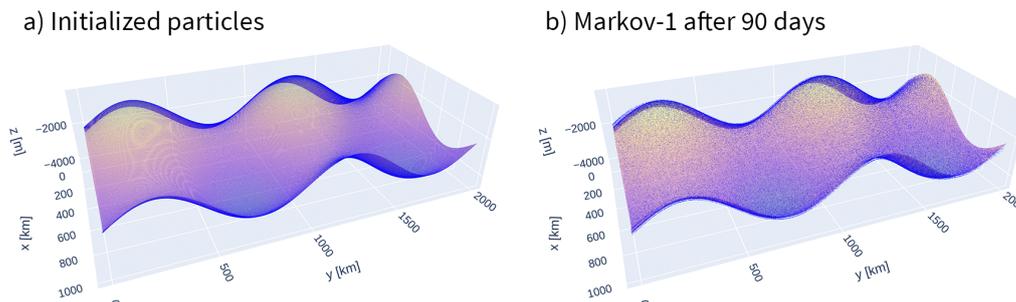


Figure 2. a) 204,800 particles on an idealized neutral surface, initialized in a regular xy -grid. b) the same particles after 90 days of integration with the Markov-1 model, with $T_L = 20$ days and $\nu^2 = 5.79 \times 10^{-4} \text{ m}^2 \text{ s}^{-2}$. Particles remain close to the neutral surface they were released on. We do not observe any distinct zones in which particles accumulate.

461

462

463

464

465

466

467

468

469

470

471

472

473

474

To visually inspect any spurious particle accumulation, which would indicate a WMC-violation, we integrate 204,800 particles with the Markov-0 and Markov-1 models for 90 days and investigate particle distributions. Figure 2 shows the initial and final particle distributions on our idealized neutral surfaces for Markov-1. We again set $T_L = 20$ days and $\nu^2 = \kappa/T_L \approx 5.79 \times 10^{-4} \text{ m}^2 \text{ s}^{-2}$, so that the effective diffusivity after 90 days (in the diffusive limit) is approximately $\kappa \approx 1 \times 10^3 \text{ m}^2 \text{ s}^{-1}$. Figure S3 in the supporting information shows the initial and final particle concentrations in the xy -plane, obtained by binning particles and dividing by the area of curved neutral surface per bin. We do not observe any distinct zones in which particles accumulate. Since the input to the Markov-1 model in this test case solely consists of the σ and θ tensors, whose elements in turn depend on the slopes of the neutral surfaces, any spurious accumulation should manifest itself at specific slope levels. Since we do not observe this, this indicates that in this stationarity situation without background flow the WMC is not violated by our ad-hoc isoneutral formulation of Markov-1.

475

4 Dispersion in an Antarctic Circumpolar Current Channel Model

476

477

478

479

480

481

We also compare the Markov-0 and Markov-1 models through Lagrangian simulations using the output of an ocean model. We use two types of Eulerian model fields at a 50 km horizontal spacing: one is the output of an ocean model run at this *coarse* resolution, and the other is a *coarsened* output of a *fine-resolution* 5 km model. The fine-resolution data serves as an eddy-resolving reference case. While the coarse-resolution data is most representative of the coarse models for which Lagrangian subgrid-scale mod-

482 els are useful, the coarsened data allows for easier comparison to the fine-resolution ref-
 483 erence case.

484 First, we look at how well Markov-1 reproduces the specified Lagrangian integral
 485 timescale and effective diffusivity in the diffusive limit. Then, we qualitatively compare
 486 particle trajectories produced by Markov-0 and Markov-1 with those produced by ad-
 487 vection only. We also compare the spread of a patch of Lagrangian particles, in analogy
 488 to a tracer patch experiment. Finally, we estimate the spurious diapycnal diffusivities
 489 introduced by the different models.

490 In each experiment, we use single values for the isoneutral Lagrangian integral time
 491 and isoneutral velocity variance. This means that we assume a homogeneous and sta-
 492 tionary situation without boundary effects. The stationarity assumption is valid for the
 493 coarsened and coarse fields, but the other assumptions are not. To deal with inhomog-
 494 eneity, we could use space-dependent and anisotropic tensors for σ and θ , but since fu-
 495 ture applications are likely to use constant parameters, we choose the pragmatic route
 496 and do so as well.

497 Since we use Eulerian data with boundaries, we need to consider boundary condi-
 498 tions. In a two-dimensional stationary and homogeneous setting, perfect reflection sat-
 499 isfies the WMC (Wilson & Flesch, 1993). Although neutral surfaces in the Southern Ocean
 500 can outcrop at the surface (Marshall & Speer, 2012), we use the assumption that neu-
 501 tral slopes at the lateral boundaries are near-flat, and adopt perfect reflection as our choice
 502 as well. The isoneutral slopes in certain areas of the model data may be unrealistically
 503 large due to spurious effects, so we use a tapering scheme based on that of Danabasoglu
 504 and McWilliams (1995) to lower or turn off turbulent displacements in such regions. De-
 505 tails of the tapering mechanism are found in supporting information Text S1.

506 4.1 Eulerian model description

507 We use a simplified model of the Antarctic Circumpolar Current run in MITgcm
 508 (Marshall et al., 1997; Campin et al., 2020), similar to the channel model used by Abernathey
 509 et al. (2011) and Balwada et al. (2018). We use an adaptation that is extensively described
 510 in MITgcm’s documentation, also available at: [https://mitgcm.readthedocs.io/en/
 511 latest/examples/reentrant_channel/reentrant_channel.html](https://mitgcm.readthedocs.io/en/latest/examples/reentrant_channel/reentrant_channel.html). It consists of a zon-
 512 ally re-entrant channel that is 1000 km long in the zonal (x) direction, 2000 km wide in
 513 the meridional (y) direction, and 3980 m deep. The model consists of 49 vertical levels
 514 that range from 5.5 m depth at the surface to 149 m at depth. It is forced by a constant
 515 sinusoidal wind stress and a temperature relaxation at the surface and northern bound-
 516 ary. The equation of state is set linearly dependent to potential temperature only, caus-
 517 ing the neutral surfaces to coincide with surfaces of constant potential temperature. This
 518 allows us to compute neutral slopes using (4). To break zonal symmetry, a meridional,
 519 Gaussian-shaped ridge is placed in the center of the domain, going up to 2382.3 m depth.
 520 The ridge has a small opening in the center, causing a strong barotropic jet to develop.

521 The model is spun up for 100 years and run at two horizontal resolutions: once at
 522 5 km resolution (*fine-resolution*), at which the mesoscale eddies are resolved, and once
 523 at 50 km resolution (*coarse-resolution*) where eddies cannot develop. Daily averages of
 524 the output data are used for the Lagrangian simulations. The coarse-resolution flow is
 525 in steady-state, exhibiting no temporal variability. We also create a coarsening of the
 526 fine-resolution model in space and time, by taking a yearly time-average of the flow and
 527 spatially averaging velocities and temperature fields over 50 kilometers. These coarsened
 528 fields thus include the effect of eddies on the mean flow. Snapshots and means of the vor-
 529 ticity and speed fields in the fine, coarsened and coarse runs are found in Figure S1. The
 530 derivatives of the density field, used for computing the neutral slopes, are computed by
 531 means of grid-aware central differences using the *XGCM* package (Abernathey et al., 2021).

532

4.2 Parameter estimation

533

534

535

536

537

538

539

540

To use the two Markov models in our experiments, we need to identify κ for Markov-0 (except when using the LS parameterization) and T_L and ν^2 for Markov-1. We can estimate globally representative values from Lagrangian quantities of the fine-resolution flow field. To do so, we first compute Lagrangian particle trajectories with the fine-resolution model output. We initialize 64,860 Lagrangian particles released regularly spaced apart 20 km in the horizontal and 200 m in the vertical, with $-200 \text{ m} \geq z \geq -1600 \text{ m}$ in order to stay away from the mixed layer and the ridge. We then integrate the trajectories using a 4th order Runge-Kutta scheme, with a timestep $\Delta t = 40$ minutes for 180 days.

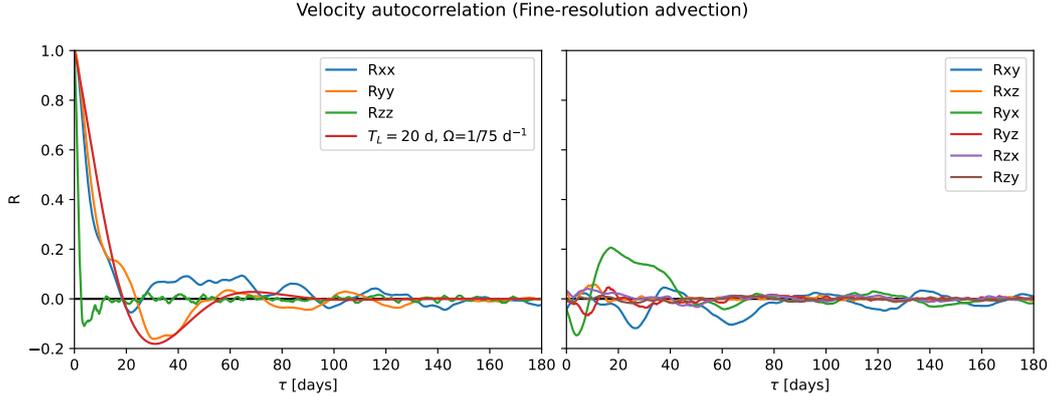


Figure 3. Lagrangian autocorrelations in the fine-resolution model, including an exponentially decaying and oscillatory function (26) with $T_L = 20$ days and $\Omega = 75$ days.

541

542

543

544

545

The Lagrangian integral time is related to the Lagrangian autocorrelation (11). Figure 3 shows the Lagrangian autocorrelation estimated from particle trajectories in the fine-resolution model. We can clearly see the oscillatory and exponentially decaying behavior of the horizontal autocorrelations. Similar to Sallée et al. (2008), we approximate the Lagrangian autocorrelation to be decomposable as

$$R(\tau) = \cos(2\pi\Omega) e^{-\tau/T_L}, \quad (26)$$

546

547

548

549

550

551

where Ω is the frequency of the oscillation. While the parameters T_L and Ω can be estimated using a least-square fit, we are only interested in approximate values for the parameters. A choice of $\Omega = 1/75$ per day and $T_L = 20$ days approximates the autocorrelation functions well enough for our purposes. Bear in mind, though, that we only continue with T_L , as Markov-1 cannot reproduce the oscillatory behavior of particle dispersion in the ocean.

552

553

554

555

Having fixed T_L , we only need to estimate κ , since this will readily give us an average value of ν^2 that reproduces the correct diffusivity in the dispersive regime through (14) (Koszalka et al., 2013). The absolute diffusivity tensor (LaCasce, 2008) is found by integrating the Lagrangian autocovariance:

$$K_{ij}(\mathbf{x}, \tau) = \int_0^\tau \langle u'_i(t_0|\mathbf{x}, t_0) u'_j(t_0 + \tilde{\tau}|\mathbf{x}, t_0) \rangle d\tilde{\tau}. \quad (27)$$

556

557

To find the isoneutral diffusivities, i and j should coincide with the principal directions of the neutral plane at each location. However, since the isoneutral slope in our model

558 is small (generally of order 10^{-3}), we will estimate the isoneutral diffusivity from K_{xx}
 559 and K_{yy} .

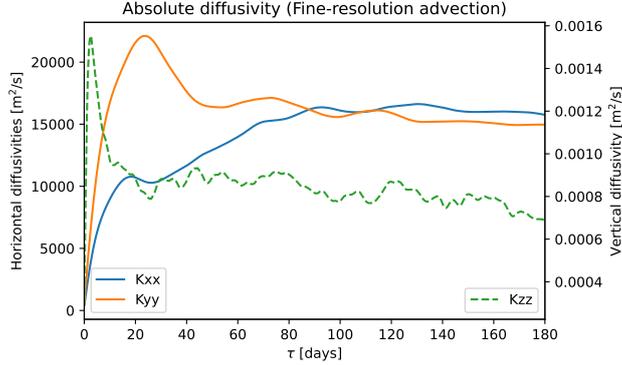


Figure 4. Absolute diffusivities K_{xx} , K_{yy} , and K_{zz} , in the fine-resolution model, computed through (27).

560 Figure 4 shows the horizontal and vertical absolute diffusivities over time. The ab-
 561 solute diffusivity corresponding to the diffusive limit, in which Markov-0 is valid, is found
 562 at $\tau \gg T_L$, for which the diffusivity should take on a near-constant value. Theoretically,
 563 it is found by integrating (27) to infinity, but in practice, it can be found by inte-
 564 grating past the negative and positive lobes associated with the oscillatory component
 565 of the Lagrangian autocorrelation, when the diffusivity becomes near-constant (Klockner,
 566 Ferrari, Lacasce, & Merrifield, 2012; Griesel et al., 2014). From Figure 4, we estimate
 567 the isoneutral diffusivity to be similar to the horizontal absolute diffusivity, with a value
 568 of $\kappa = 1.5 \times 10^4 \text{ m}^2 \text{ s}^{-1}$.

569 4.3 Lagrangian integral time and diffusivity from the Markov-1 model

570 Now we initialize particles in the same lattice as used in section 4.2 and apply the
 571 Markov-1 parameterization. We simulate trajectories by integrating the stochastic dif-
 572 ferential equations (7) using the Euler-Maruyama scheme (21) for 180 days, with $\Delta t =$
 573 40 minutes. We set $T_L = 20$ days, and specify $\nu^2 = 8.68 \times 10^{-3} \text{ m}^2 \text{ s}^{-2}$ in order to ob-
 574 tain an effective diffusivity of $1.5 \times 10^4 \text{ m}^2 \text{ s}^{-1}$ in the diffusive limit. We also set η and
 575 ε in such a way that the effective diapycnal diffusivity in the limit $t \gg T_L$ is $1 \times 10^{-5} \text{ m}^2 \text{ s}^{-1}$.
 576 These settings are used in the remainder of this study. Derivatives of Eulerian quanti-
 577 ties that are necessary for computing the tensor elements of σ and θ (and later \mathbf{K}) are
 578 computed with central differences and successively interpolated linearly in space. Our
 579 aim is to see how well the model reproduces the diffusivity and Lagrangian timescale that
 580 we specified, to verify our ad-hoc diapycnal formulation of Markov-1.

581 Figure 5 shows the Lagrangian autocorrelation and absolute diffusivity of parti-
 582 cles simulated using the Markov-1 subgrid-scale model using the coarsened field, simi-
 583 lar to Figures 3 and 4. Figure S2 provides a similar diagram for the coarse field. The
 584 exponential decay with an e -folding timescale of 20 days can be clearly seen in the au-
 585 tuncorrelation. There is a clear absence of the oscillatory component, which Markov-1 is
 586 unable to simulate.

587 The absolute diffusivity of $1.5 \times 10^4 \text{ m}^2 \text{ s}^{-1}$ is not fully reproduced. In the x -direction,
 588 values reach up to approximately $1.4 \times 10^4 \text{ m}^2 \text{ s}^{-1}$, but in the y -direction, they are much
 589 smaller, with a maximum of $1.0 \times 10^4 \text{ m}^2 \text{ s}^{-1}$ and a decrease at larger time lags. There
 590 are two reasons why values do not reach $1.5 \times 10^4 \text{ m}^2 \text{ s}^{-1}$. First, in regions where the

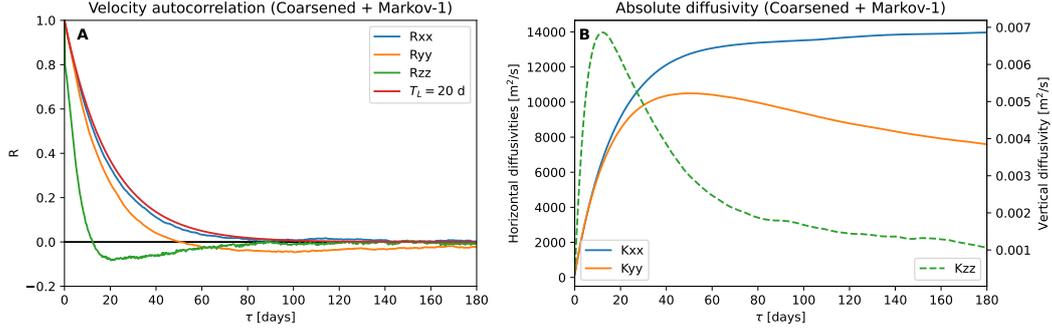


Figure 5. Lagrangian autocorrelation and absolute diffusivity produced by the Markov-1 model when applied on the coarsened field. The Lagrangian autocorrelation in the x -direction best resembles that of an exponentially decaying function with a 20-day e -folding timescale (in red for reference).

591 slope is unrealistically high, for example in the direct vicinity of the meridional ridge,
 592 turbulent velocities are tapered to zero (see supporting information Text S1), which de-
 593 creases the absolute diffusivity computed from the particle ensemble. Second, the lat-
 594 eral domain boundaries limit the dispersion of material and therefore also cause a de-
 595 crease in diffusivity, as D_{yy} cannot grow linearly over long timescales. While the effect
 596 of tapering likely plays a role for both K_{xx} and K_{yy} , only K_{yy} is affected by boundaries,
 597 which causes it to decrease over time. We clearly see that R_{zz} has a much shorter e -folding
 598 time than 20 days. This is likely due to the effect of curvature in the neutral surfaces,
 599 and the rapid decorrelation we impose in the dianeutral directions (18).

600 **4.4 Individual trajectories**

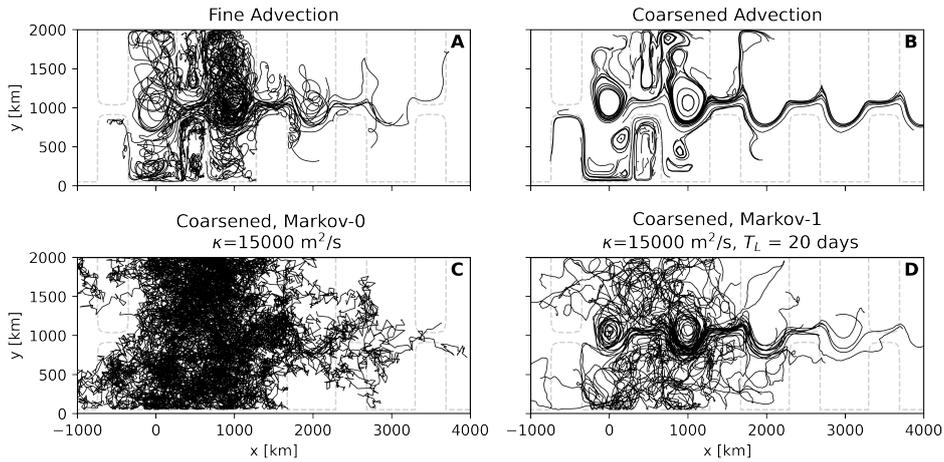


Figure 6. 100 randomly subsampled trajectories from 180 days of simulation on (a) fine-resolution and (b) coarsened fields, and using coarsened fields in combination with (c) the diffusive parameterization and (d) Markov-1. While the domain is periodic, here we tile it in the zonal direction, to separate particles crossing the zonal periodic boundaries. The -3900m isobath is plotted with dashed grey lines, indicating the location of the ridge in the periodic channel.

601 A typical aim of Lagrangian subgrid-scale dispersion models is to construct real-
 602 istic synthetic particle pathways in the absence of turbulent eddies. It is therefore illus-
 603 trative to plot particle trajectories generated by advection using the three model fields
 604 (fine, coarsened and coarse) and compare those with trajectories generated by Markov-
 605 0 and Markov-1. To do so, we randomly subsample 100 trajectories that were initialized
 606 on the same lattice as used in section 4.2. We again use the Runge-Kutta 4 scheme for
 607 advection and Euler-Maruyama for the Markov models, a timestep $\Delta t = 40$ minutes,
 608 and a simulation time of 180 days. Like in the previous section, we tuned Markov-1 to
 609 produce a diapycnal diffusivity of $1 \times 10^{-5} \text{ m s}^{-2}$, and now we do the same for Markov-
 610 0 by setting $\epsilon \kappa$ accordingly. These parameters will also be used for the remainder of this
 611 paper. To more easily identify re-entering trajectories, we record when particles cross
 612 the periodic boundary, so that we can plot particle trajectories as unbroken paths by re-
 613 peating the periodic domain in the zonal direction.

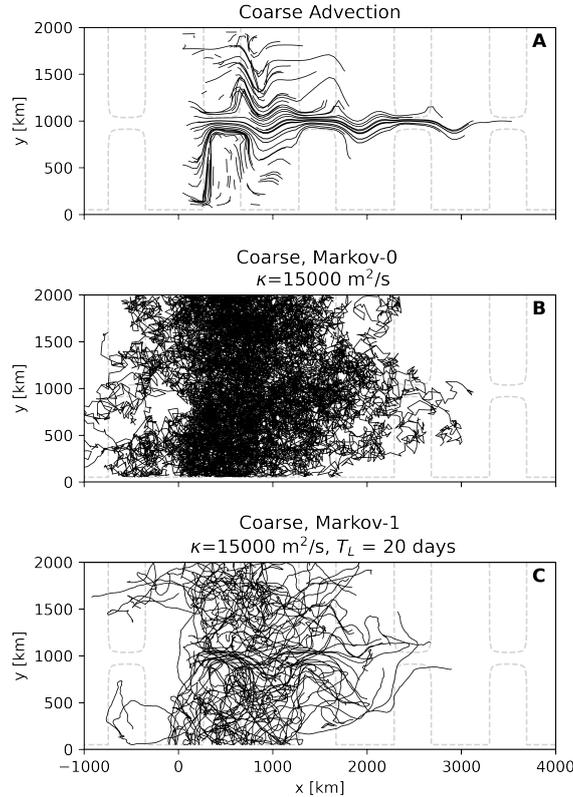


Figure 7. Same as Figure 6, but using coarse-resolution fields.

614 Figure 6 considers 100 trajectories from Markov-0 and Markov-1 in the coarsened
 615 case, compared to advection using fine-resolution and coarsened fields, which serve as
 616 reference. These trajectories are released at different horizontal and vertical locations,
 617 subsampled from the lattice used in the previous two sections. From the trajectories in
 618 Markov-0 we clearly see that there is no autocorrelation in the particle velocities, with
 619 the directions in which a particle moves rapidly changing between recorded timesteps.
 620 Particles simulated with Markov-0 also travel much more, as the turbulent displacement
 621 in this model is much larger than that of Markov-1 (see the discussion in section 3.3).
 622 Markov-1 clearly does a better job at simulating the trajectories from the fine-resolution
 623 reference run. A major difference is that trajectories in the fine run exhibit looping mo-
 624 tions. While the trajectories in Markov-1 veer over time, it is unable to produce the loop-

ing motions that are seen in the fine-resolution run (Veneziani et al., 2005). Bear in mind that in the stochastic perturbations between different particles advected by Markov-0 and Markov-1 are uncorrelated. Instead, each particle ‘feels’ its own turbulent field.

Figure 7 considers the coarse-resolution case. In this case, the underlying flow field has no eddies. When comparing trajectories produced by the Markov models, we thus have no eddying reference case. In the advection-only case, the absence of strong dispersion is clear. One major difference with the results from the coarsened case is the absence of any stationary meanders. Trajectories produced by Markov-1 again seem the most realistic when compared to Figure 6a, albeit less obviously than was the case for 6d.

4.5 Tracer spread

In analogy to studying the spread of a small patch of tracer (Wagner et al., 2019), we qualitatively compare the spread of a patch of Lagrangian particles advected in the fine-resolution, coarsened, and coarse-resolution fields and apply the Markov-0 and Markov-1 subgrid-scale models to the later two flows. For Markov-0, we use the isotropic isoneutral diffusion tensor $\mathbf{K}_{\text{Redi,approx}}$ (5) and the LS parameterization \mathbf{K}_{LS} (6). For the LS parameterization, we set $C = 1$.

We initialize a patch of particles initially located at $z = -736$ m (corresponding to the 25th vertical level) in a radius of 50 km centered around $(x = 250 \text{ km}, y = 1000 \text{ km})$, see Figure 8a.

Figure 8 shows the particle distributions after 180 days of simulation, using advection and the different subgrid-scale models on the coarsened flow data. Again, we repeat the domain in the zonal direction, so that we can distinguish particles that have crossed the periodic boundary. Figure 8c shows the obvious need for modeling subgrid-scale dispersion when turbulent flow features are filtered out.

Figures 8d, e, and f show similar patterns when compared to one another, albeit with the dispersion in the LS case being somewhat weaker, and particles in the Markov-0 case reaching deeper than the others. Note that the diffusivity in the LS parameterization is solely determined by derivatives of the flow fields. The pattern in 8e is qualitatively similar to 8b, which bears testimony to the skill of the LS parameterization. Since the particles in the parameterizations each experience their own independent turbulent fields, coherent structures and filamentation as seen in 8b cannot be reproduced by the Markov models.

In both Markov-0 models and in the Markov-1 model, we see some spurious particle accumulation on the left side of the ridges (at $x = 500 \text{ km} + n * 1000 \text{ km}$, with $n = 0, 1, 2, \dots$). In the LS case, these accumulation patterns (or patterns where particles are fully absent) occur at other places too. In all cases this is likely due to sharp changes in the discrete derivatives used for computing the slopes that are necessary for filling the elements of \mathbf{K} , $\boldsymbol{\sigma}$, and $\boldsymbol{\theta}$. The LS parameterization relies on discrete derivatives of more quantities for computing its tensor elements, since these also depend on the shear of the flow (see (6)). It is therefore more susceptible to violations of the WMC when these discrete derivatives change strongly in space and interpolation is used.

Figure 9 shows the spreading of Lagrangian particles in the coarse model. Again, the isotropic Markov-0 model and Markov-1 show a similar spread of particles, with particles in Markov-0 again reaching slightly larger depths. However, the LS parameterization this time produces very different results, with the dispersion being much more limited, and the particles being more concentrated. This means that in this case the shear-based parameterization leads to much smaller diffusivities in \mathbf{K}_{LS} . This makes sense, as the fine-resolution flow field (and thus the coarsened flow) is full of baroclinic instabil-

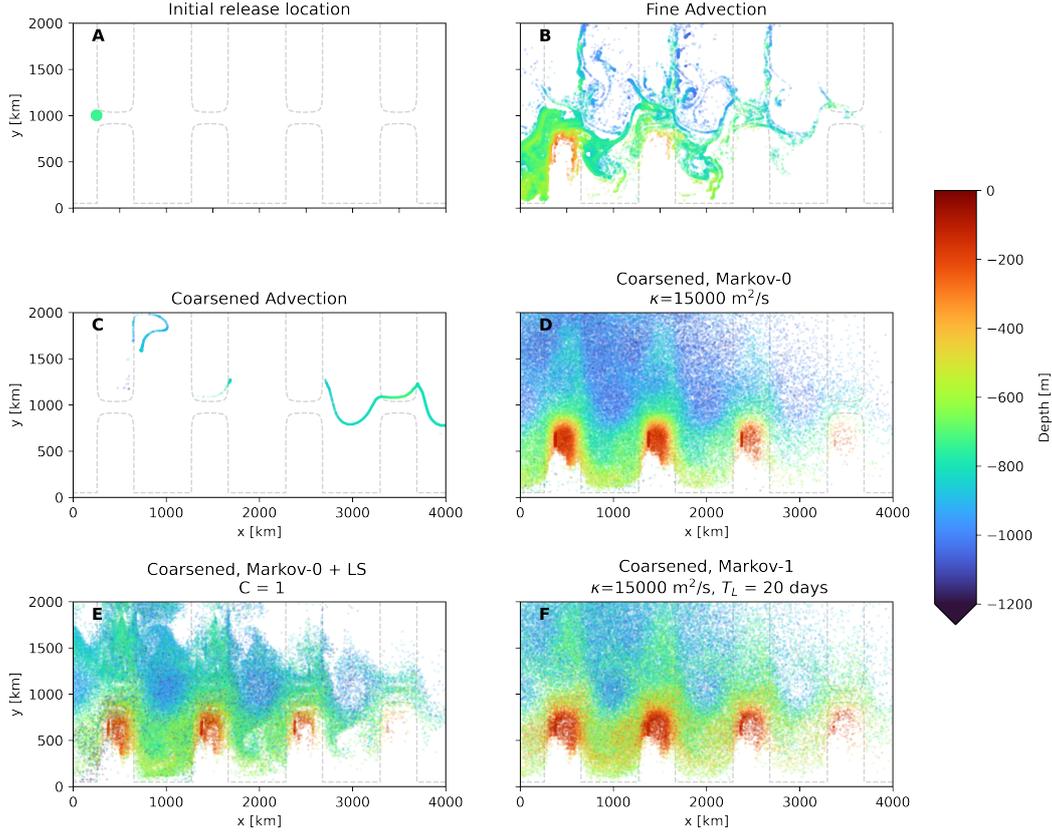


Figure 8. (a) Initial particle positions at $z = -736$ m, (b)-(f) show particle locations and depths after 180 days of simulation with $\Delta t = 40$ minutes. (b) & (c) show particles advected with the fine-resolution and coarsened model fields, while (d)-(f) use the diffusion/Markov-0 and Markov-1 models. Particles that fall within the mixed layer are not shown (see supporting information Text S2)

674 ities that lead to eddies with large shear. The resolution in the coarse model is too low
 675 for these instabilities to develop. Instead, the flow tends to a much smoother steady-state,
 676 with less shear. As this yields smoother derivatives in the temperature field (and in the
 677 velocity fields in the case of LS), this should lead to less spurious accumulation. Indeed
 678 we see no clear regions where particles accumulate.

679 4.6 Spurious dianeutral diffusivity

680 Two possible causes of spurious dianeutral tracer fluxes are numerical integration
 681 and interpolation of discrete, time-evolving Eulerian flow fields. The spurious dianeutral
 682 flux can be expressed as a diffusivity, and this diffusivity should be as small as possible
 683 compared to the vertical diffusivity that is specified to represent dianeutral processes.
 684 For example, in the Southern Ocean, the average diapycnal diffusivity at 1500 m depth
 685 is estimated to be $1.3 \pm 0.2 \times 10^{-5} \text{ m}^2 \text{ s}^{-1}$ (Ledwell et al., 2011). It is important to assess
 686 how large the dianeutral diffusivities in our Lagrangian simulations become, and how
 687 they compare to the dianeutral diffusivity that we specify. In this section, we will assess
 688 these spurious dianeutral diffusivities. In these experiments, we specified an explicit di-
 689 aneutral diffusivity of $1 \times 10^{-5} \text{ m}^2 \text{ s}^{-1}$. Moreover, in the case of the Markov models, we

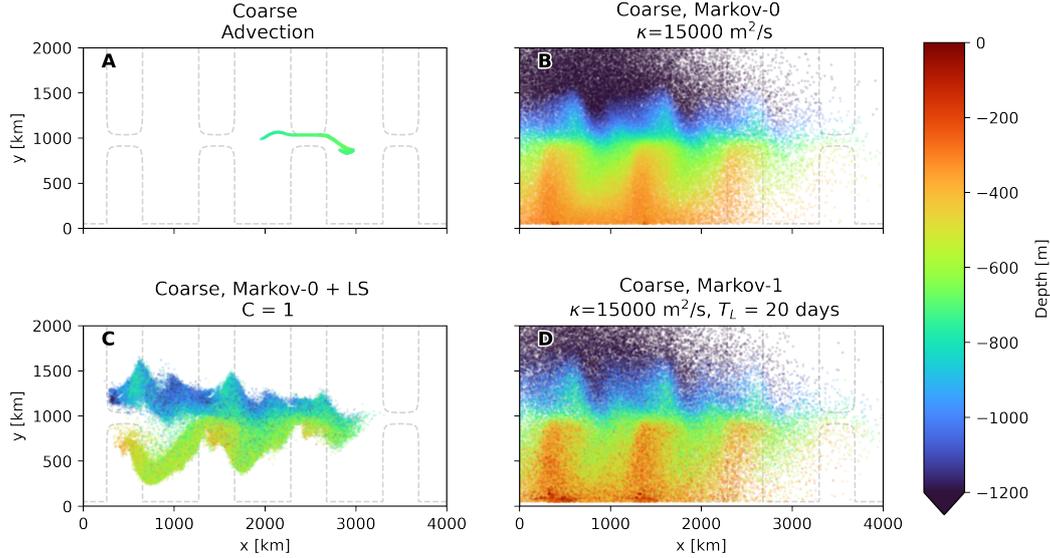


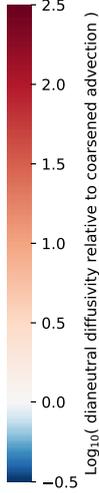
Figure 9. Like 8, with (a) advection in coarse-resolution model, (b)-(d) using the different subgrid-scale models.

690 test several values of (effective) isoneutral diffusivities, keeping $T_L = 20$ days in the case
 691 of Markov-1. For Markov-0 combined with the LS parameterization, we choose differ-
 692 ent tuning parameters C at $\mathcal{O}(1)$, which affect the strength of the diffusivity.

693 We compute the effective dianeutral diffusivity in the case of pure advection using
 694 the fine-resolution, coarsened, and coarse-resolution fields, and using the Markov-
 695 0 and Markov-1 model. This dianeutral diffusivity is approximated as follows: for each
 696 particle, we record its initial local water density. Then, after simulating the particle’s
 697 movement for 180 days, at the particle’s new horizontal location, we compute the depth
 698 z_{iso} of the neutral surface corresponding to the original local water density. Comparing
 699 this depth with the particle’s new depth, we can compute a spurious vertical diffusiv-
 700 ity (similar to (25)). This again assumes that the dianeutral diffusivity is closely aligned
 701 with the vertical direction. We separate the results for three depth classes on which par-
 702 ticles were released. Particle trajectories that at any point reach depths of -50 m or higher
 703 are excluded in these computations, in order to filter out effects related to particles enter-
 704 ing the mixed layer (see supporting information Text S2).

705 The results are found in Table 1 for the coarsened flow and in Table 2 for the coarse-
 706 resolution flow. In all cases, the effective dianeutral diffusivities are larger than the value
 707 of $1 \times 10^{-5} \text{ m}^2 \text{ s}^{-1}$ that we explicitly set, meaning that the spurious dianeutral diffu-
 708 sivities due to errors in interpolation and the numerical schemes dominate. This is al-
 709 ready the case for simulations that only use advection. We found that halving the timestep
 710 does not make a difference here, indicating that the error in the case of advection is likely
 711 not due to the time discretization. In the case of advection using fine-resolution data,
 712 the distance that a particle covers over the course of one flow snapshot, compared to the
 713 length of a grid cell, is relatively larger than in the case of coarse-resolution data, where
 714 it takes longer to traverse the larger cells. The dianeutral error can then be reduced by
 715 using more frequent snapshots of the data (e.g. 6-hourly snapshots instead of daily), such
 716 that temporal interpolation occurs over a smaller time window (Qin et al., 2014). This
 717 could however come at a large expense in storage, memory and I/O. Here, we are solely
 718 interested in comparing the errors between different Lagrangian simulations, so we ac-
 719 cept that the dianeutral diffusivities are larger than specified. In the coarsened and coarse-

Table 1. Effective dianeutral diffusivity (in $\text{m}^2 \text{s}^{-1}$) for different depth classes with parameterizations applied on the *coarsened* flow field, after numerical integration for 180 days, with $\Delta t=40$ minutes. The color scale indicates the logarithm of the relative dianeutral diffusivity, when divided by the dianeutral diffusivity in the coarsened case per depth class as reference. This indicates the orders of magnitude that the dianeutral diffusivity differs from that in the simulations with only advection using coarsened fields. Of the parameterizations, Markov-1 has the smallest dianeutral diffusivity, in some cases even smaller than in the simulation with advection only.

	All depths	$-200 \geq z > -600$	$-600 \geq z > -1200$	$-1200 \geq z \geq -1600$	
Fine advection	4.2E-04	1.1E-03	2.1E-04	1.9E-04	
Coarsened advection	3.8E-04	9.3E-04	1.9E-04	1.7E-04	
Markov-0, $\kappa=1500$	9.2E-04	1.5E-03	6.4E-04	9.1E-04	
Markov-0, $\kappa=5000$	3.1E-03	2.1E-03	2.5E-03	4.2E-03	
Markov-0, $\kappa=15000$	1.4E-02	5.0E-03	1.2E-02	1.8E-02	
Markov-0 + LS, C=0.3	8.6E-04	1.8E-03	5.8E-04	6.0E-04	
Markov-0 + LS, C=1	1.6E-03	2.6E-03	1.2E-03	1.5E-03	
Markov-0 + LS, C=3	5.0E-03	5.3E-03	3.9E-03	6.2E-03	
Markov-1, $\kappa=1500$	6.2E-04	1.3E-03	4.5E-04	4.0E-04	
Markov-1, $\kappa=5000$	5.6E-04	9.2E-04	5.6E-04	3.6E-04	
Markov-1, $\kappa=15000$	2.1E-04	3.7E-04	2.0E-04	1.5E-04	

720 resolution fields, we use steady-state flows, meaning that the errors are due to spatial
 721 interpolation of coarse data, with time-interpolation playing no role.

722 Both Tables 1 and 2 show that for each experiment Markov-0 produces a much larger
 723 spurious dianeutral diffusivity than Markov-1. This corroborates the findings of section
 724 3.3. A likely explanation is that the isoneutral turbulent displacement in each of the mod-
 725 els becomes somewhat dianeutral as discrete neutral surfaces ‘curve’, while the displace-
 726 ments in Markov-1 are much smaller than is the case for Markov-0. In the case of Markov-
 727 0, we see the error increasing as the diffusivity increases. This pattern cannot be seen
 728 for Markov-1, where in some cases, the error decreases with increasing effective diffusiv-
 729 ity. Unfortunately, we do not have an explanation for this pattern.

730 Since the dianeutral diffusivity in the case of Markov-0 can become several
 731 orders of magnitude larger than is the case for only advection, future studies should be care-
 732 ful with applying this subgrid-scale dispersion parameterization. Here we implemented
 733 the Euler-Maruyama scheme. Higher-order schemes, such as the first order Milstein scheme,
 734 are able to greatly reduce the dianeutral error in idealized situations (Shah et al., 2011;
 735 Gräwe, 2011; Shah et al., 2013). However, we found that the Milstein-1 scheme produces
 736 similar dianeutral errors to Euler-Maruyama when applied on our coarsened and coarse-
 737 resolution flows, further indicating that the cause of the error lies in interpolation com-
 738 bined with large turbulent displacements.

739 5 Conclusion

740 We achieved two main goals: formulating an isoneutral description of the Markov-
 741 1 model, and extending an anisotropic tracer diffusion parameterization to the random

Table 2. Same as Table 1, but using *coarse-resolution* flow fields. Again, Markov-1 has the lowest dianeutral diffusivity of the three parameterizations.

	All depths	$-200 \geq z > -600$	$-600 \geq z > -1200$	$-1200 \geq z \geq -1600$	
Coarse advection	2.6E-05	1.5E-05	2.4E-05	3.7E-05	
Markov-0, $\kappa=1500$	6.3E-04	1.9E-03	2.5E-04	1.8E-04	
Markov-0, $\kappa=5000$	1.8E-03	5.8E-03	7.4E-04	3.9E-04	
Markov-0, $\kappa=15000$	3.8E-03	8.6E-03	2.5E-03	2.6E-03	
Markov-0 + LS, C=0.3	3.6E-04	6.1E-04	2.4E-04	3.2E-04	
Markov-0 + LS, C=1	5.2E-04	8.4E-04	4.0E-04	4.2E-04	
Markov-0 + LS, C=3	1.1E-03	2.1E-03	8.8E-04	7.7E-04	
Markov-1, $\kappa=1500$	2.6E-04	5.7E-04	1.3E-04	1.9E-04	
Markov-1, $\kappa=5000$	3.2E-04	9.4E-04	1.1E-04	1.3E-04	
Markov-1, $\kappa=15000$	1.0E-04	3.9E-04	1.5E-05	2.0E-05	

742 walk dispersion/Markov-0 model. With these goals, we aim to improve the parameter-
 743 ization of unresolved isoneutral turbulent motions due to eddies in Lagrangian studies.

744 Because of the inclusion of a velocity autocorrelation, the Markov-1 model is able
 745 to produce both the ballistic and diffusive dispersion regime, and it produces particle tra-
 746 jectories and dispersion patterns that are more realistic than those produced by Markov-
 747 0. Our formulation of Markov-1, inspired by Redi’s diffusion tensor, also has a much smaller
 748 spurious dianeutral flux than Markov-0, due to the smaller turbulent displacement in each
 749 timestep. Large turbulent displacements in the isoneutral direction in the presence of
 750 curvature in the neutral surfaces lead to dianeutral excursions. Therefore, our three-dimensional
 751 isoneutral formulation of Markov-1 will hopefully be useful to the Lagrangian commu-
 752 nity, with the many benefits of higher order stochastic models beyond Markov-1 given
 753 by previous studies (Griffa, 1996; Berloff & McWilliams, 2002; Veneziani et al., 2004).
 754 We also believe that the isoneutral formulation of the parameter tensors (16) and (17)
 755 is extendable to the parameter tensors of the higher order stochastic models beyond Markov-
 756 1, as well as other improvements to this model, like the inclusion of looping motions.

757 Further research into the isoneutral formulation of Markov-1, as well higher order
 758 stochastic models, may focus on better retaining the velocity autocorrelation on curved
 759 surfaces, which remains a complex issue (Gaspari & Cohn, 1999). Next to that, it may
 760 also further investigate boundary conditions further, as well as how Lagrangian parti-
 761 cle models can transition from isoneutral dispersion in the ocean interior to horizontal
 762 and vertical mixing in the mixed layer, which has been left out of this study (see sup-
 763 porting information S2). Moreover, future studies employing isoneutral dispersion mod-
 764 els may benefit from improved computation of neutral surface slopes (Groeskamp et al.,
 765 2019).

766 We hope that future Lagrangian studies using coarse fields, such as the output of
 767 coupled Earth system models, may also benefit from the LS parameterization, as well
 768 as other Eulerian anisotropic parameterizations based on closure. This may help auto-
 769 matically determine the strength of the eddy diffusivity in different regions in the do-
 770 main. When applied to the coarsened flow field, the LS parameterization was able to pro-
 771 duce particle distributions similar to the isotropic Markov models, meaning that LS may
 772 obviate the need for explicit parameter estimation in Markov-0. Our discussion of the
 773 LS parameterization may inspire further investigation into the application of closure schemes

774 in Lagrangian simulations. Similarly, such closures could be further studied for the Markov-
 775 1 model, although so far Berloff and McWilliams (2003) tested a related closure based
 776 on shear with negative results.

777 Data and Software Availability Statement

778 Lagrangian datasets (CC-BY) and data generation and analysis scripts (MIT li-
 779 cense) for this research are available at <https://doi.org/10.24416/UU01-RXA2PB>. This
 780 includes MITgcm model generation scripts and documentation, data post-processing scripts,
 781 Parcels Lagrangian simulation scripts and analysis scripts for generating figures and ta-
 782 bles.

783 Acknowledgments

784 DR was supported through funding from the Netherlands Organization for Scientific Re-
 785 search (NWO), Earth and Life Sciences, through project OCENW.KLEIN.085. ED is
 786 an honorary research associate with the Belgian Fund for Scientific Research (F.R.S.-
 787 FNRS). EvS was supported through funding from the European Research Council (ERC)
 788 under the European Union’s Horizon 2020 research and innovation program (grant agree-
 789 ment 715386). We thank Pavel Berloff and Arnold Heemink for a useful discussion on
 790 the topic of this study and Hyder Shah for tips on numerically implementing the SDE
 791 schemes. We also thank Stephen Griffies and two anonymous reviewers for helpful feed-
 792 back that improved this manuscript.

793 References

- 794 Abernathey, R., Busecke, J., Smith, T., Banihirwe, A., jldleauna, Bot, S., ...
 795 Rath, W. (2021, May). *xgcm/xgcm: v0.5.2*. Zenodo. Retrieved from
 796 <https://doi.org/10.5281/zenodo.4821276> doi: 10.5281/zenodo.4821276
- 797 Abernathey, R., Ferreira, D., & Klocker, A. (2013). Diagnostics of isopycnal mixing
 798 in a circumpolar channel. *Ocean Modelling*, *72*, 1–16. doi: 10.1016/j.ocemod
 799 .2013.07.004
- 800 Abernathey, R., Marshall, J., & Ferreira, D. (2011). The Dependence of Southern
 801 Ocean Meridional Overturning on Wind Stress. *Journal of Physical Oceanogra-
 802 phy*, *41*(12), 2261–2278. doi: 10.1175/JPO-D-11-023.1
- 803 Bachman, S. D., Fox-Kemper, B., & Bryan, F. O. (2020). A Diagnosis of Anisotropic
 804 Eddy Diffusion From a High-Resolution Global Ocean Model. *Journal of Ad-
 805 vances in Modeling Earth Systems*, *12*(2). doi: 10.1029/2019MS001904
- 806 Balwada, D., Smith, K. S., & Abernathey, R. (2018). Submesoscale Vertical
 807 Velocities Enhance Tracer Subduction in an Idealized Antarctic Circum-
 808 polar Current. *Geophysical Research Letters*, *45*(18), 9790–9802. doi:
 809 10.1029/2018GL079244
- 810 Berloff, P. S., & McWilliams, J. C. (2002). Material Transport in Oceanic Gyres.
 811 Part II: Hierarchy of Stochastic Models. *Journal of Physical Oceanography*, *32*,
 812 34. doi: 10.1175/1520-0485(2002)032<0797:MTIOGP>2.0.CO;2
- 813 Berloff, P. S., & McWilliams, J. C. (2003). Material Transport in Oceanic Gyres.
 814 Part III: Randomized Stochastic Models. *Journal of Physical Oceanography*,
 815 *33*, 30.
- 816 Berloff, P. S., McWilliams, J. C., & Bracco, A. (2002). Material Transport
 817 in Oceanic Gyres. Part I: Phenomenology. *JOURNAL OF PHYSICAL
 818 OCEANOGRAPHY*, *32*, 33.
- 819 Boussinesq, J. (1877). Essai sur la theorie des eaux courantes. *Memoires presentes
 820 par divers savants a l’Academie des Sciences de l’Institut National de France*,
 821 *23*(1).
- 822 Campin, J.-M., Heimbach, P., Losch, M., Forget, G., edhill3, Adcroft, A., ...

- 823 raphael dussin (2020, July). *Mitgcm/mitgcm: mid 2020 version*. Zen-
824 odo. Retrieved from <https://doi.org/10.5281/zenodo.3967889> doi:
825 10.5281/zenodo.3967889
- 826 Chelton, D. B., Deszoeke, R. A., Schlax, M. G., Naggar, K. E., & Siwertz, N. (1998).
827 Geographical Variability of the First Baroclinic Rossby Radius of Deformation.
828 *Journal of Physical Oceanography*, *28*, 28.
- 829 Cox, M. D. (1987). Isopycnal diffusion in a z-coordinate model. *Ocean Modelling*.
830 Danabasoglu, G., & McWilliams, J. C. (1995). Sensitivity of the Global Ocean Cir-
831 culation to Parameterizations of Mesoscale Tracer Transports. *Journal of Cli-*
832 *mate*, *8*, 2967–297.
- 833 Delandmeter, P., & van Sebille, E. (2019). The Parcels v2.0 Lagrangian framework:
834 New field interpolation schemes. *Geoscientific Model Development*, *12*(8),
835 3571–3584. doi: 10.5194/gmd-12-3571-2019
- 836 Fox-Kemper, B., Adcroft, A., Böning, C. W., Chassignet, E. P., Curchitser, E.,
837 Danabasoglu, G., . . . Yeager, S. G. (2019). Challenges and Prospects
838 in Ocean Circulation Models. *Frontiers in Marine Science*, *6*, 65. doi:
839 10.3389/fmars.2019.00065
- 840 Gaspari, G., & Cohn, S. E. (1999). Construction of correlation functions in two and
841 three dimensions. *Quarterly Journal of the Royal Meteorological Society*, *723*–
842 *757*.
- 843 Gent, P. R., & McWilliams, J. C. (1990). Isopycnal Mixing in Ocean Circulation
844 Models. *Journal of Physical Oceanography*, *20*(1), 150–155.
- 845 Gräwe, U. (2011). Implementation of high-order particle-tracking schemes in a water
846 column model. *Ocean Modelling*, *36*(1), 80–89. doi: 10.1016/j.ocemod.2010.10
847 .002
- 848 Gräwe, U., Deleersnijder, E., Shah, S. H. A. M., & Heemink, A. W. (2012). Why the
849 Euler scheme in particle tracking is not enough: The shallow-sea pycnocline
850 test case. *Ocean Dynamics*, *62*(4), 501–514. doi: 10.1007/s10236-012-0523-y
- 851 Griesel, A., Gille, S. T., Sprintall, J., McClean, J. L., LaCasce, J. H., & Maltrud,
852 M. E. (2010). Isopycnal diffusivities in the Antarctic Circumpolar Current
853 inferred from Lagrangian floats in an eddying model. *Journal of Geophysical*
854 *Research*, *115*(C6), C06006. doi: 10.1029/2009JC005821
- 855 Griesel, A., McClean, J. L., Gille, S. T., Sprintall, J., & Eden, C. (2014). Eule-
856 rian and Lagrangian Isopycnal Eddy Diffusivities in the Southern Ocean of
857 an Eddying Model. *Journal of Physical Oceanography*, *44*(2), 644–661. doi:
858 10.1175/JPO-D-13-039.1
- 859 Griffa, A. (1996). Applications of stochastic particle models to oceanographic prob-
860 lems. In R. J. Adler, P. Müller, & B. L. Rozovskii (Eds.), *Stochastic Modelling*
861 *in Physical Oceanography* (pp. 113–140). Boston, MA: Birkhäuser Boston. doi:
862 10.1007/978-1-4612-2430-3_5
- 863 Griffies, S. M. (1998). The Gent–McWilliams skew flux. *Journal of Physical*
864 *Oceanography*, *28*(5), 831–841. doi: 10.1175/1520-0485(1998)0280831:TGMSF2
865 .0.CO;2
- 866 Groeskamp, S., Barker, P. M., McDougall, T. J., Abernathey, R. P., & Griffies,
867 S. M. (2019). VENM: An Algorithm to Accurately Calculate Neutral Slopes
868 and Gradients. *Journal of Advances in Modeling Earth Systems*, *11*(7), 1917–
869 1939. doi: 10.1029/2019MS001613
- 870 Haigh, M., Sun, L., McWilliams, J. C., & Berloff, P. S. (2021). On eddy transport in
871 the ocean. Part II: The advection tensor. *Ocean Modelling*, *165*, 101845. doi:
872 10.1016/j.ocemod.2021.101845
- 873 Haller, G., & Yuan, G. (2000). Lagrangian coherent structures and mixing in two-
874 dimensional turbulence. *Physica D: Nonlinear Phenomena*, *147*(3-4), 352–370.
875 doi: 10.1016/S0167-2789(00)00142-1
- 876 Haza, A. C., Piterbarg, L. I., Martin, P., Özgökmen, T. M., & Griffa, A. (2007).
877 A Lagrangian subgridscale model for particle transport improvement and ap-

- 878 plication in the Adriatic Sea using the Navy Coastal Ocean Model. *Ocean*
879 *Modelling*, 17(1), 68–91. doi: 10.1016/j.ocemod.2006.10.004
- 880 Heemink, A. W. (1990). Stochastic modelling of dispersion in shallow water.
881 *Stochastic Hydrology and Hydraulics*, 4, 161–174.
- 882 Hewitt, H. T., Roberts, M., Mathiot, P., Biastoch, A., Blockley, E., Chassignet,
883 E. P., . . . Zhang, Q. (2020). Resolving and Parameterising the Ocean
884 Mesoscale in Earth System Models. *Current Climate Change Reports*, 6(4),
885 137–152. doi: 10.1007/s40641-020-00164-w
- 886 Iacus, S. M. S. M. (2008). *Simulation and inference for stochastic differential equa-*
887 *tions: With R examples*. New York, N.Y.: Springer.
- 888 Klocker, A., Ferrari, R., & LaCasce, J. H. (2012). Estimating Suppression of Eddy
889 Mixing by Mean Flows. *Journal of Physical Oceanography*, 42(9), 1566–1576.
890 doi: 10.1175/JPO-D-11-0205.1
- 891 Klocker, A., Ferrari, R., Lacasce, J. H., & Merrifield, S. T. (2012). Reconciling float-
892 based and tracer-based estimates of lateral diffusivities. *Journal of Marine Re-*
893 *search*, 70(4), 569–602. doi: 10.1357/002224012805262743
- 894 Kloeden, P. E., & Platen, E. (1999). *Numerical solution of stochastic differential*
895 *equations* (No. 23). Berlin: Springer.
- 896 Koszalka, I., LaCasce, J. H., & Mauritzen, C. (2013). In pursuit of anoma-
897 lies—Analyzing the poleward transport of Atlantic Water with surface drifters.
898 *Deep Sea Research Part II: Topical Studies in Oceanography*, 85, 96–108. doi:
899 10.1016/j.dsr2.2012.07.035
- 900 LaCasce, J. H. (2008). Statistics from Lagrangian observations. *Progress in Oceanog-*
901 *raphy*, 77(1), 1–29. doi: 10.1016/j.pocean.2008.02.002
- 902 Ledwell, J. R., St. Laurent, L. C., Girton, J. B., & Toole, J. M. (2011). Diapycnal
903 Mixing in the Antarctic Circumpolar Current. *Journal of Physical Oceanogra-*
904 *phy*, 41(1), 241–246. doi: 10.1175/2010JPO4557.1
- 905 Le Sommer, J., d’Ovidio, F., & Madec, G. (2011). Parameterization of subgrid
906 stirring in eddy resolving ocean models. Part 1: Theory and diagnostics. *Ocean*
907 *Modelling*, 39(1-2), 154–169. doi: 10.1016/j.ocemod.2011.03.007
- 908 Marshall, J., Adcroft, A., Hill, C., Perelman, L., & Heisey, C. (1997). A finite-
909 volume, incompressible Navier Stokes model for studies of the ocean on parallel
910 computers. *Journal of Geophysical Research: Oceans*, 102(C3), 5753–5766.
911 doi: 10.1029/96JC02775
- 912 Marshall, J., & Speer, K. (2012). Closure of the meridional overturning circulation
913 through Southern Ocean upwelling. *Nature Geoscience*, 5(3), 171–180. doi: 10
914 .1038/ngeo1391
- 915 Mathieu, P.-P., & Deleersnijder, E. (1998). What is wrong with isopycnal diffusion
916 in world ocean models? *Applied Mathematical Modelling*, 22(4), 367–378. doi:
917 10.1016/S0307-904X(98)10008-2
- 918 McDougall, T. J. (1987). Neutral Surfaces. *Journal of Physical Oceanography*,
919 17(11), 1950–1964. doi: 10.1175/1520-0485(1987)017<1950:NS>2.0.CO;2
- 920 McWilliams, J. C., Weiss, J. B., & Yavneh, I. (1994). Anisotropy and Coherent Vor-
921 tex Structures in Planetary Turbulence. *Science*, 264(5157), 410–413. doi: 10
922 .1126/science.264.5157.410
- 923 Nooteboom, P. D., Delandmeter, P., van Sebille, E., Bijl, P. K., Dijkstra, H. A., &
924 von der Heydt, A. S. (2020). Resolution dependency of sinking Lagrangian
925 particles in ocean general circulation models. *PLOS ONE*, 15(9), e0238650.
926 doi: 10.1371/journal.pone.0238650
- 927 Nummelin, A., Busecke, J. J. M., Haine, T. W. N., & Abernathey, R. (2020).
928 Diagnosing the Scale and Space Dependent Horizontal Eddy Diffusivity
929 at the Global Surface Ocean. *Journal of Physical Oceanography*. doi:
930 10.1175/JPO-D-19-0256.1
- 931 Piterbarg, L. I. (2002). The Top Lyapunov Exponent for a Stochastic Flow Modeling
932 the Upper Ocean Turbulence. *SIAM Journal on Applied Mathematics*, 62(3),

- 777–800. doi: 10.1137/S0036139999366401
- 933
934 Qin, X., van Sebille, E., & Sen Gupta, A. (2014). Quantification of errors induced by
935 temporal resolution on Lagrangian particles in an eddy-resolving model. *Ocean*
936 *Modelling*, 76, 20–30. doi: 10.1016/j.ocemod.2014.02.002
- 937 Redi, M. H. (1982). Oceanic Isopycnal Mixing by Coordinate Rotation. *Journal*
938 *of Physical Oceanography*, 12(10), 1154–1158. doi: 10.1175/1520-0485(1982)
939 012(1154:OIMBCR)2.0.CO;2
- 940 Reynolds, A. (2002). On Lagrangian stochastic modelling of material transport in
941 oceanic gyres. *Physica D: Nonlinear Phenomena*, 172(1-4), 124–138. doi: 10
942 .1016/S0167-2789(02)00660-7
- 943 Rodean, H. C. (1996). *Stochastic Lagrangian Models of Turbulent Diffusion*. Boston,
944 MA: American Meteorological Society. doi: 10.1007/978-1-935704-11-9
- 945 Sallée, J. B., Speer, K., Morrow, R., & Lumpkin, R. (2008). An estimate of
946 Lagrangian eddy statistics and diffusion in the mixed layer of the South-
947 ern Ocean. *Journal of Marine Research*, 66(4), 441–463. doi: 10.1357/
948 002224008787157458
- 949 Sawford, B. L. (1991). Reynolds number effects in Lagrangian stochastic models
950 of turbulent dispersion. *Physics of Fluids A: Fluid Dynamics*, 3(6), 1577–1586.
951 doi: 10.1063/1.857937
- 952 Shah, S. H. A. M., Heemink, A., & Deleersnijder, E. (2011). Assessing Lagrangian
953 schemes for simulating diffusion on non-flat isopycnal surfaces. *Ocean Mod-*
954 *elling*, 39(3-4), 351–361. doi: 10.1016/j.ocemod.2011.05.008
- 955 Shah, S. H. A. M., Heemink, A. W., Gräwe, U., & Deleersnijder, E. (2013).
956 Adaptive time stepping algorithm for Lagrangian transport models: The-
957 ory and idealised test cases. *Ocean Modelling*, 68, 9–21. doi: 10.1016/
958 j.ocemod.2013.04.001
- 959 Smagorinsky, J. (1963). General Circulation Experiments With the Primitive Equa-
960 tions: I. The Basic Experiment. *Monthly Weather Review*, 91(3), 99–164. doi:
961 10.1175/1520-0493(1963)0910099:GCEWTP2.3.CO;2
- 962 Spagnol, S., Wolanski, E., Deleersnijder, E., Brinkman, R., McAllister, F., Cushman-
963 Roisin, B., & Hanert, E. (2002). An error frequently made in the evaluation
964 of advective transport in two-dimensional Lagrangian models of advection-
965 diffusion in coral reef waters. *Marine Ecology Progress Series*, 235, 299–302.
966 doi: 10.3354/meps235299
- 967 Spivakovskaya, D., Heemink, A. W., & Deleersnijder, E. (2007). Lagrangian mod-
968 elling of multi-dimensional advection-diffusion with space-varying diffusivities:
969 Theory and idealized test cases. *Ocean Dynamics*, 57(3), 189–203. doi:
970 10.1007/s10236-007-0102-9
- 971 Su, Z., Wang, J., Klein, P., Thompson, A. F., & Menemenlis, D. (2018). Ocean sub-
972 mesoscales as a key component of the global heat budget. *Nature Communica-*
973 *tions*, 9(1), 775. doi: 10.1038/s41467-018-02983-w
- 974 Taylor, G. I. (1922). Diffusion by continuous movements. *Proceedings of the London*
975 *Mathematical Society*, 2(1), 196–212.
- 976 Thomson, D. J. (1987). Criteria for the selection of stochastic models of particle tra-
977 jectories in turbulent flows. *Journal of Fluid Mechanics*, 180(-1), 529. doi: 10
978 .1017/S0022112087001940
- 979 van Sebille, E., Griffies, S. M., Abernathey, R., Adams, T. P., Berloff, P., Biastoch,
980 A., . . . Zika, J. D. (2018). Lagrangian ocean analysis: Fundamentals and
981 practices. *Ocean Modelling*, 121, 49–75. doi: 10.1016/j.ocemod.2017.11.008
- 982 van Sebille, E., Delandmeter, P., Lange, M., Kehl, C., Reijnders, D., Fischer, R.,
983 . . . Vallès, I. (2020, October). *OceanParcels/parcels: Parcels v2.2.1: a La-*
984 *grangian Ocean Analysis tool for the petascale age*. Zenodo. Retrieved from
985 <https://doi.org/10.5281/zenodo.4108148> doi: 10.5281/zenodo.4108148
- 986 Veneziani, M., Griffa, A., Reynolds, A. M., Garraffo, Z. D., & Chassignet, E. P.
987 (2005). Parameterizations of Lagrangian spin statistics and particle disper-

- 988 sion in the presence of coherent vortices. *Journal of Marine Research*, 63(6),
989 1057–1083. doi: 10.1357/002224005775247571
- 990 Veneziani, M., Griffa, A., Reynolds, A. M., & Mariano, A. J. (2004). Oceanic
991 Turbulence and Stochastic Models from Subsurface Lagrangian Data for the
992 Northwest Atlantic Ocean. *Journal of Physical Oceanography*, 34, 23.
- 993 Visser, A. (1997). Using random walk models to simulate the vertical distribution
994 of particles in a turbulent water column. *Marine Ecology Progress Series*, 158,
995 275–281. doi: 10.3354/meps158275
- 996 Wagner, P., Rühls, S., Schwarzkopf, F. U., Koszalka, I. M., & Biastoch, A. (2019).
997 Can Lagrangian Tracking Simulate Tracer Spreading in a High-Resolution
998 Ocean General Circulation Model? *Journal of Physical Oceanography*, 49(5),
999 1141–1157. doi: 10.1175/JPO-D-18-0152.1
- 1000 Wilson, J. D., & Flesch, T. K. (1993). Flow boundaries in random-flight dispersion
1001 models: Enforcing the well-mixed condition. *Journal of Applied Meteorology*
1002 *and Climatology*, 32(11), 1695–1707. doi: 10.1175/1520-0450(1993)0321695:
1003 FBIRFD2.0.CO;2
- 1004 Wolfram, P. J., Ringler, T. D., Maltrud, M. E., Jacobsen, D. W., & Petersen, M. R.
1005 (2015). Diagnosing Isopycnal Diffusivity in an Eddying, Idealized Midlati-
1006 tude Ocean Basin via Lagrangian, in Situ, Global, High-Performance Particle
1007 Tracking (LIGHT). *Journal of Physical Oceanography*, 45(8), 2114–2133. doi:
1008 10.1175/JPO-D-14-0260.1

Supporting Information for “Simulating Lagrangian Subgrid-Scale Dispersion on Neutral Surfaces in the Ocean”

Daan Reijnders¹, Eric Deleersnijder², Erik van Sebille¹

¹Institute for Marine and Atmospheric research Utrecht, Utrecht University, the Netherlands

²Université catholique de Louvain, Institute of Mechanics, Materials and Civil Engineering (IMMC) & Earth and Life Institute

(ELI), Louvain-la-Neuve, Belgium

Contents of this file

1. Text S1 to S2
2. Figures S1 to S3

Text S1. Tapering scheme

We use central differences to compute the neutral slopes S_x and S_y (see equation (4)) in the discrete Eulerian model data in the experiments in Section 4. Incidentally, the computed neutral slopes can be unrealistically high, for example in the vicinity of the meridional ridge or in the mixed layer (see Text S2). The buoyancy in the mixed layer is mostly uniform, but if we compute neutral slopes in this region, small deviations in the local buoyancy field can lead to huge slopes. It is common practice in Eulerian ocean modeling to limit or turn off isopycnal/isoneutral diffusion in regions with high slopes, in order to prevent numerical instability. This practice is called ‘tapering’.

Here we use a tapering scheme similar to that of Danabasoglu and McWilliams (1995) to smoothly decrease the values of the Markov-0 diffusivity tensors $\mathbf{K}_{\text{redi,approx}}$ and \mathbf{K}_{LS} to zero in regions with high slopes. Similarly, for Markov-1, we use it to smoothly decrease the perturbative velocity \mathbf{u}' to zero in such regions. At each timestep, we respectively multiply $\mathbf{K}_{\text{redi,approx}}$, \mathbf{K}_{LS} , or \mathbf{u}' by a taper function f_{taper} which assumes values between 1 in regions where the isoneutral slopes are well-behaved and 0 in regions where it is unrealistically high. Danabasoglu and McWilliams (1995) choose a taper function

$$f_{\text{taper,DMW}}(S) = \frac{1}{2} \left(1 + \tanh \left[\frac{S_c - |S|}{S_d} \right] \right), \quad (1)$$

where S_c is the slope at which $f_{\text{taper}} = 0.5$ and S_d an acting distance over which f_{taper} changes steeply. If we were to multiply the perturbative velocity \mathbf{u}' in the Markov-1 model (7) with such a function, this causes an exponential decay of the \mathbf{u}' with an e -folding timescale of $\Delta t / \log(f(s))$. This can significantly shorten the effective decorrelation of \mathbf{u}' as set by T_L . For example, in a simulation with $T_L = 20$ days and $dt = 40$ minutes, if $f(S)$ persistently equals 0.999, this causes \mathbf{u}' to exponentially decay with a timescale of 28 days. In conjunction with the exponential decorrelation specified using T_L , this leads to an effective decorrelation with an e -folding timescale of 12 days. This is why we limit the slope values over which tapering happens smoothly to values that differ from S_c by at most $3S_d$. We thus use the following taper function

$$f_{\text{taper}}(S) = \begin{cases} 1 & |S| < S_c - 3S_d \\ \frac{1}{2} \left(1 + \tanh \left[\frac{S_c - |S|}{S_d} \right] \right) & S_c - 3S_d \leq |S| \leq S_c + 3S_d \\ 0 & S_c + 3S_d < |S| \end{cases} \quad (2)$$

Note that $f_{\text{taper}}(S_c - 3S_d) \approx 0.998$ and $f_{\text{taper}}(S_c + 3S_d) \approx 0.002$. In our simulations in section 4, we choose $S_c = 8 \times 10^{-3}$ and $S_d = 5 \times 10^{-4}$. With these values, tapering occurs only in a fraction of the domain, namely near the meridional ridge and in the mixed layer.

Text S2. Treatment of the mixed layer

By definition, potential temperature is approximately homogeneous in the mixed layer. As neutral surfaces appeal to the notion of a strong stratification which inhibits motion in the dianeutral direction, the concept of neutral surfaces does not apply in the mixed layer. That is why the experiments in this study focus on the ocean interior. In the experiments in section 4, particles are released well below the mixed layer. Still, since neutral surfaces in the Southern Ocean can outcrop to the surface (Marshall & Speer, 2012), particles in our model may be transported to the surface. In Figures 8 and 9, we exclude particles that fall within the mixed layer. Similarly, in the computation of the spurious dianeutral diffusivities in section 4.6, we exclude particle trajectories that at any point reach depths of -50 m. The actual mixed-layer, marked by a sharp gradient in potential temperature, lies less deep, but since it varies in space, we use -50 m as a global approximation for computational efficiency.

References

- Danabasoglu, G., & McWilliams, J. C. (1995). Sensitivity of the Global Ocean Circulation to Parameterizations of Mesoscale Tracer Transports. *Journal of Climate*, 8, 2967–297.
- Marshall, J., & Speer, K. (2012). Closure of the meridional overturning circulation through Southern Ocean upwelling. *Nature Geoscience*, 5(3), 171–180. doi: 10.1038/ngeo1391

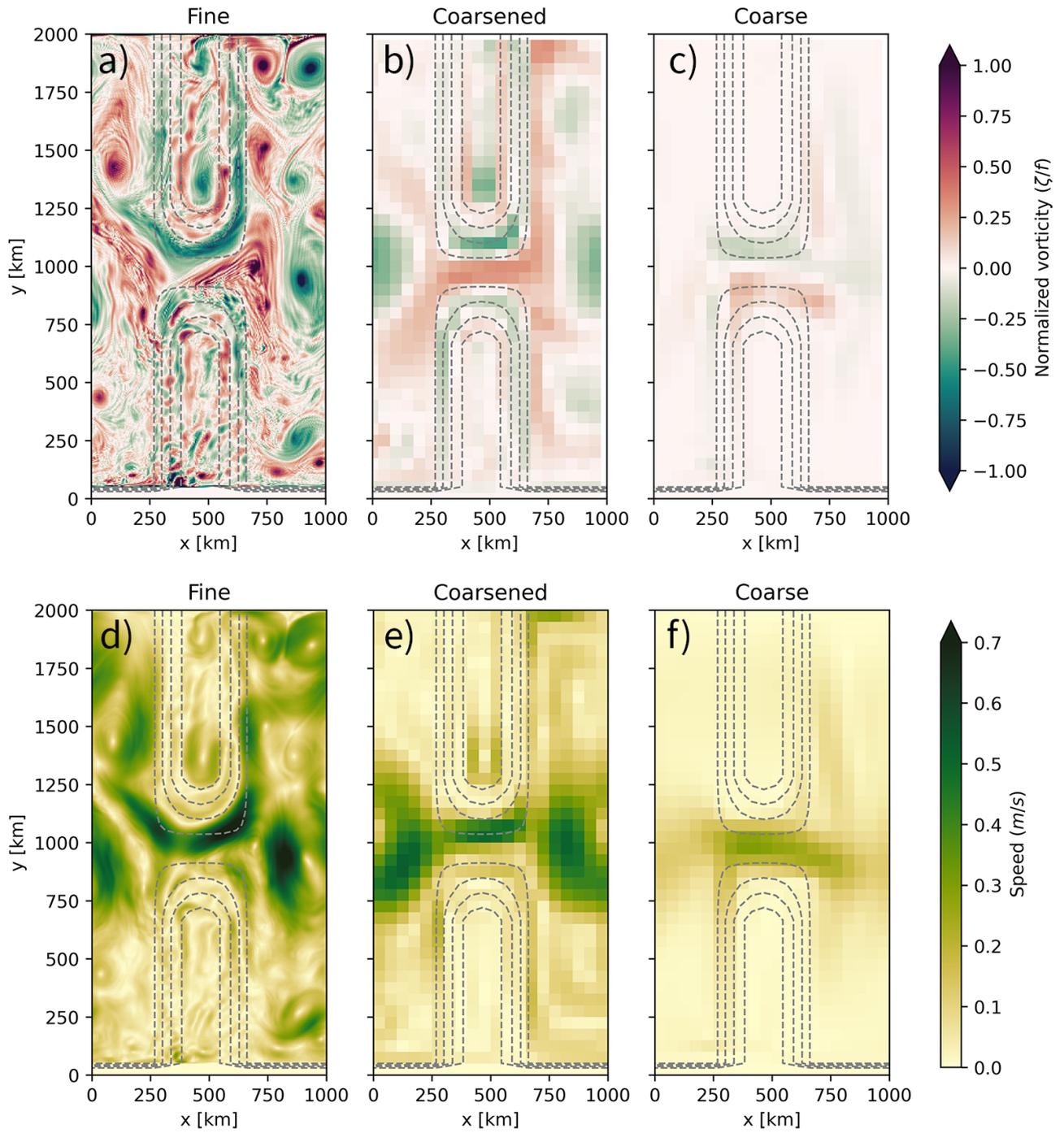


Figure S1. Snapshot of the vorticity (a-c) and speed (d-f) of the fine (a & d), coarsened (b & e), and coarse (c & f) model fields used in this study. The fine fields are daily averages, the coarsened fields are 1-year time averages and 50 kilometer spatial averages, and the coarse model is in steady state. Dashed lines indicate the position of the meridional ridge.

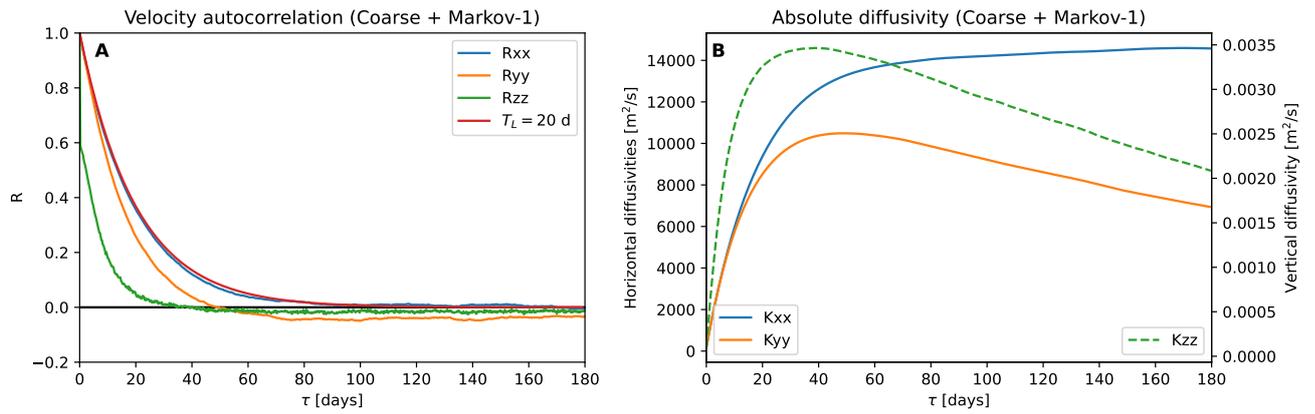


Figure S2. Lagrangian autocorrelation and absolute diffusivity produced by the Markov-1 model when applied on the coarse field (cf. Figure 5). The Lagrangian autocorrelation in the x -direction best resembles that of an exponentially decaying function with a 20-day e -folding timescale (in red for reference).

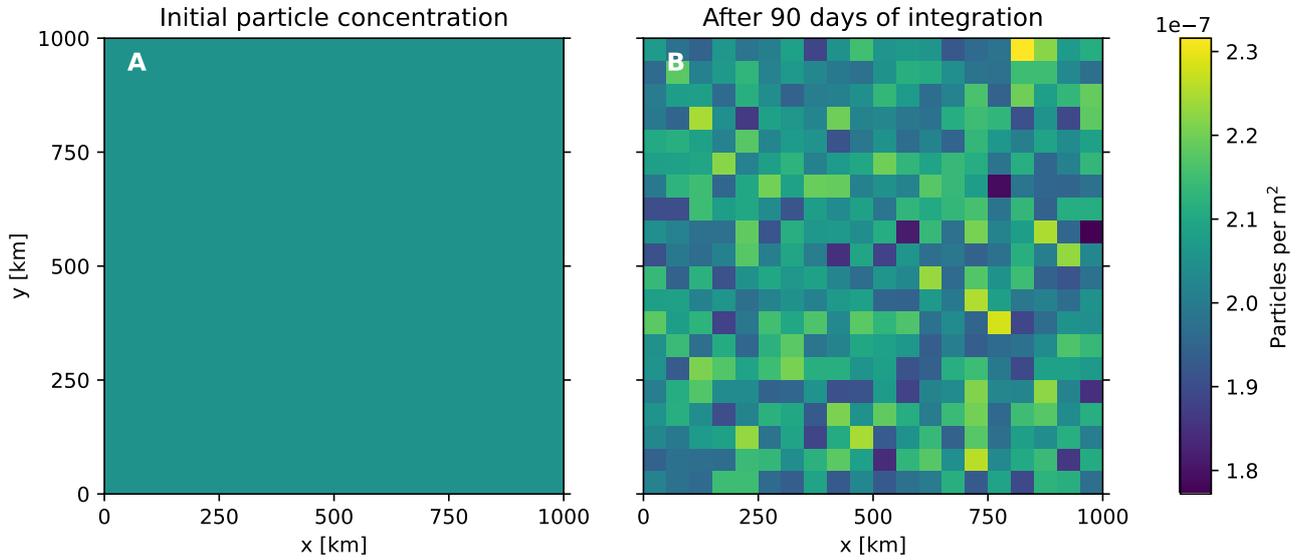


Figure S3. Concentrations of 204,800 particles before (a) and (b) after 90 days of integration using the Markov-1 model, using $T_L = 20$ days and $\nu^2 = 5.79 \times 10^{-4} \text{ m}^2 \text{ s}^{-2}$. We take advantage of the periodicity of the domain and analyze all particles over one wavelength $1/k_x = 1/k_y = 1000$ km by displacing them as $x = x \bmod 1/k_x$, $y = y \bmod 1/k_y$. The concentrations are computed by binning particles and dividing by the total area of curved surface per bin. Particles start out evenly spaced. From (a) it can be seen that the curvature causes only negligible differences in initial concentrations. After 90 days of integration (b), concentrations are much less homogeneous than they were initially, but there are no clear accumulation patterns coinciding with specific features of the idealized neutral surface. If that were the case, it would indicate that the well-mixed condition is violated.