

Stochastic superparameterization and multiscale filtering of turbulent tracers

Yoonsang Lee^{a,*}, Andrew J. Majda^a, Di Qi^a

^a*Center for Atmosphere Ocean Science, Courant Institute of Mathematical Sciences, New York University, New York, NY 10012, USA*

Abstract

Data assimilation or filtering combines a numerical forecast model and observations to provide accurate statistical estimation of the state of interest. In this paper we are concerned with accurate data assimilation of a sparsely observed passive tracer advected in turbulent flows using a reduced-order forecast model. The turbulent flows which contain anisotropic and inhomogeneous structures such as jets are typical in geophysical turbulent flows in atmosphere and ocean science and passive tracers with a mean gradient can exhibit anisotropic transport with intermittent extreme events as shown below. Stochastic superparameterization, which is a seamless multiscale method developed for large-scale models of atmosphere and ocean models without scale-gap between the resolved and unresolved scales, generates large-scale turbulent velocity fields using a significantly smaller degree of freedoms compared to a direct fine resolution numerical simulation. In a large-scale model of the tracer transport, the tracer is advected by the large-scale velocity field generated by superparameterization with a parameterization of eddies, an additional eddy diffusion given by an anisotropic biharmonic diffusion. The turbulent tracer is sparsely observed in space in only the upper surface layer. These observations naturally mix the resolved and unresolved scales so we develop an ensemble multiscale data assimilation algorithm which provides estimates of the resolved scales using the mixed observations. The reduced-order model is 250 times cheaper than the fine resolution solution and thus enables us to increase the number of ensembles for accurate predictions of state distributions. Numerical experiments show positive results in the estimation of the resolved scales of the tracer as well as in capturing anisotropic intermittent extreme events for the unresolved portions of the tracer field.

Keywords: data assimilation, superparameterization, turbulent tracer, multiscale

*Corresponding author

Email address: ylee@cims.nyu.edu (Yoonsang Lee)

1. Introduction

Turbulent diffusion of passive tracers is a major problem in geophysical science and engineering. The passive scalar field $T(\mathbf{x}, t)$ represents the concentration of the tracer immersed in the fluid which is advected by a velocity field and dissipated by, for example, molecular diffusion but does not by itself influence the dynamics of the fluid. When $\mathbf{v}(\mathbf{x}, t)$ is a velocity field, the dynamics of the passive tracer T is given by

$$\frac{\partial T}{\partial t} + \mathbf{v} \cdot \nabla T = \mathcal{D}T \quad (1)$$

where \mathcal{D} is a scale-selective linear dissipation operator. The description of the passive tracer has applications in, for example, the spread of pollutants in environmental science and the anthropogenic and natural tracers in climate change science. This model, which is a linear equation while statistically nonlinear due to the advection flow \mathbf{v} , also serves as a simpler prototype model for a large-scale closure for Navier-Stokes equations [1] and transport of baroclinic potential vorticity in the ocean [2].

When the velocity field \mathbf{v} becomes turbulent, it is important to describe the statistical properties of the tracer such as energy spectra and probability density functions (PDFs) rather than a single realization of the tracer trajectory. For roughly Gaussian background velocity \mathbf{v} as typically observed in turbulent flows, the interaction of many complex factors in \mathbf{v} can create extreme rare fluctuations in the tracer field. Observational data [3] and experimental results [4] show the existence of intermittency in the tracer statistics, which results in fat-tailed PDFs and random spikes in the time series of the tracer. The fat-tailed PDFs can also be generated by numerical simulations in idealized models of (1) [5, 6] whose mechanism for the intermittent behavior is rigorously investigated recently in [7].

In the simulation of (1), due to a wide range of scales of \mathbf{v} , it is computationally prohibitive to resolve all scales of \mathbf{v} . On the other hand, the central quantity of interest is the tracer statistics rather than the background flow field \mathbf{v} , thus the parameterization of \mathbf{v} by simpler approximations is practical and advantageous in real applications. For coarse resolution numerical models, the turbulent diffusion of passive tracers is modeled by additional dissipation which is called eddy diffusion in the turbulent literature [1]. Typically, the eddy diffusivity is estimated in some manner, such as mixing-length theory [8] and isotropic down-gradient diffusion for isotropic flows with Gaussian randomness. For a data-driven prediction model of the turbulent diffusion of passive tracers, see [9] which uses empirical information theory to select optimal model parameters in a training phase.

In this paper we are interested in turbulent diffusion of passive tracers in anisotropic and inhomogeneous flows, which are typical in atmospheric and oceanic flows. In particular, we consider flows dominated by jets in a turbulent background, for example, of β -plane turbulence. Observation data of satellite-tracked drifting buoys show that the tracer transport along the jet direction is stronger than the transport across the jet [10]. The anisotropic tracer parameterization has been investigated in [11, 12] using statistical turbulence closure models. In [13] and [2], the eddy diffusivities in the across-jet

and along-jet directions are analyzed using mixing-length theory and shear dispersion respectively, which supports the observations of larger diffusivity in the along-jet direction than that in the across-jet direction.

The goal of this paper is to develop reduced-order data assimilation methods of the turbulent tracer which provides accurate estimation of the tracer statistics through the combination of sparse noisy spatial observations and a reduced-order forecast model. In ensemble based filters or particle filters, many ensemble members are required to accurately approximate the prior distributions of the state variables. But due to limited computing resources, the practical ensemble numbers are limited and insufficient as the dimensionality of the system increases, which is usually called “the curse of ensemble size” [14]. The incorporation of a coarse reduced-order forecast model in data assimilation enables us to use a large number of ensemble, which alleviates the curse of ensemble size. By contrast, the reduced-order forecast model provides only the large-scale estimation of the tracer field and thus we encounter a problem due to the mixed contributions from both the resolved coarse-scale and unresolved small-scale components in observations. To overcome these problems, we employ the multiscale data assimilation methods developed in [15, 16], which update the resolved large-scale variables using the mixed observations of the resolved and unresolved scale components. In particular, an ensemble multiscale data assimilation method with covariance localization and inflation is applied to stabilize the filtering procedure.

To generate turbulent velocity fields which are anisotropic and inhomogeneous with jets, a two-layer quasigeostrophic equation [14, 17] becomes one desirable choice containing many interesting and important features that can be found in realistic flows. The model is an idealized system for large-scale geophysical turbulence and includes regimes with an inverse cascade of kinetic energy. In order to complete the reduced-order forecast method for the turbulent diffusion of passive tracers, we need a coarse resolution method to generate turbulent velocity fields which reduces the computational complexity. For this model, which is nonlinear, there is no straightforward approach to parameterize the small-scales such as the eddy diffusivity for the passive tracer. In this study we use the stochastic superparameterization method [18, 19] to generate coarse-scale approximate turbulent flows. Stochastic superparameterization is a seamless multiscale method for the parameterization of the unresolved sub-grid scales by solving cheap local problems embedded in a coarse grid and has been successfully applied to the large-scale approximation of the quasigeostrophic equations [18, 20, 19] (see [18] for a comprehensive review of superparameterization including the conventional deterministic superparameterization). Through the use of reduced-order models both for the generation of turbulent velocity fields and the passive tracer, the computational complexity reduces significantly from 256^2 to 48^2 while robust prediction skills of the statistical properties of the tracer field are well maintained.

This paper is organized in the following way. In Section 2 the two-layer quasigeostrophic equation is described which generates anisotropic and inhomogeneous turbulent flows. Also we describe a turbulent diffusion model of the passive tracer forced by a large-scale tracer mean gradient. Preliminary direct numerical simulations are

provided showing non-trivial non-Gaussian statistics of the tracer field. In Section 3 we briefly review stochastic superparameterization for the two-layer quasigeostrophic equation and discuss an eddy diffusion for the turbulent diffusion of the passive tracer modeled by an anisotropic biharmonic diffusion. An ensemble multiscale data assimilation method is discussed in Section 4 which updates the resolved large-scale variables using the mixed observations of the resolved and unresolved sub-grid scales. In Section 5 we test the multiscale data assimilation methods for the passive tracer using sparse observations of the upper layer tracer field followed by conclusions in Section 6.

2. Model equations and a reference simulation

We consider the turbulent diffusion of the passive tracer (1) in two-dimensional turbulent flows which generate anisotropic and inhomogeneous structures with jets. The model to generate the flows is the two-layer quasigeostrophic equation, which is a classical idealized model of synoptic-scale turbulence in atmosphere and ocean science. Beginning with a brief review of a reference simulation of the flow equation, which displays anisotropic and inhomogeneous mean flows with jets, we examine the tracer turbulent diffusion advected by turbulent flows and forced by a large-scale mean tracer gradient. In particular, we consider two tracer fields which are in the along-jet and across-jet directions. In the along-jet tracer field which is forced by the zonal velocity component, the tracer shows a similar pattern aligned with the jets and non-trivial Gaussian PDFs which implies intermittency in the tracer field; while in the across-jet field, the tracer displays no intermittency.

2.1. Two-layer quasigeostrophic equation

The model equation to generate turbulent flows is the following nondimensional two-layer quasigeostrophic equation in a doubly periodic domain

$$\begin{aligned}
\partial_t q_1 &= -\tilde{\mathbf{v}}_1 \cdot \nabla q_1 - \partial_x q_1 - (k_\beta^2 + k_d^2) \tilde{v}_1 - \nu \Delta^4 q_1, \\
\partial_t q_2 &= -\tilde{\mathbf{v}}_2 \cdot \nabla q_2 + \partial_x q_2 - (k_\beta^2 - k_d^2) \tilde{v}_2 - r \Delta \psi_2 - \nu \Delta^4 q_2, \\
q_1 &= \Delta \psi_1 + \frac{k_d^2}{2} (\psi_2 - \psi_1), \\
q_2 &= \Delta \psi_2 - \frac{k_d^2}{2} (\psi_2 - \psi_1).
\end{aligned} \tag{2}$$

Here q_j is the potential vorticity in the upper ($j = 1$) and lower ($j = 2$) layers, the velocity-streamfunction relation $\tilde{\mathbf{v}}_j = (\tilde{u}_j, \tilde{v}_j) = (-\partial_y \psi_j, \partial_x \psi_j)$, k_d is the deformation wavenumber, r is a linear Ekman drag coefficient at the bottom layer of the flows and ν is the hyperviscosity to dissipate small-scales. In order to absorb a downscale cascade of enstrophy while leaving all but the smallest scales nearly inviscid, we use a hyperdissipation, $\Delta^4 q_j$ which is a common choice in turbulence simulations to suppress some of the instabilities at small-scales while capturing interesting dynamics of the large-scales. The nondimensional k_β results from the variation of the vertical projection of

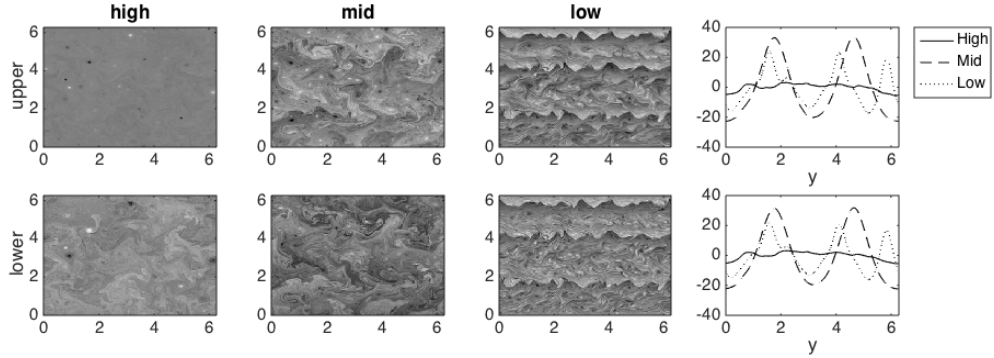


Figure 1: Snapshots of potential vorticities $q_j, j = 1, 2$. Upper layer (top row) and lower layer (bottom row). High (first column), mid (second column) and low (third column) latitudes. The fourth column is the temporally and zonally averaged zonal velocity.

Coriolis frequency with latitude. The equation (2) is forced by an imposed baroclinic instability; a large-scale zonal vertical shear is applied with equal and opposite unit velocities which is related to the terms $(-1)^j(\partial_x q_j + k_d^2 \tilde{v}_j)$. Thus the total velocity, \mathbf{v}_j , to advect the passive tracer for the upper and lower layers are given by

$$\mathbf{v}_j = (-\partial_y \psi_j + (-1)^{j-1}, \partial_x \psi_j), \quad j = 1, 2. \quad (3)$$

For reference simulations, we consider three different regimes studied in [20, 25]. These regimes correspond to high, mid and low latitude ocean dynamics by changing the parameters k_β and r . The parameters corresponding to high, mid, and low latitude regimes are given by

$$\begin{aligned} \text{High : } & k_\beta = 0 \text{ and } r = 8, \\ \text{Mid : } & k_\beta^2 = k_d^2/4 \text{ and } r = 2, \\ \text{Low : } & k_\beta^2 = k_d^2/2 \text{ and } r = 0.5, \end{aligned}$$

while the deformation wavenumber k_d and the hyper viscosity ν are fixed at 25 and 1.28×10^{-15} respectively for all three regimes. The reference simulations use pseudo-spectral numerics with a 256×256 grid points for each layer. For time integration, we use a semi-implicit third order Runge-Kutta time integration which uses an exponential integrator for the stiff linear dissipation term. The time step is fixed at 2×10^{-5} .

Figure 1 shows snapshots of the upper and lower layer potential vorticities $q_j, j = 1, 2$ for three different regimes along with the temporally and zonally averaged zonal velocities. In the high latitude case, which corresponds to a f -plane case, the flow is dominated by small-scale vorticities with spatially homogeneous and isotropic statistics. In the mid and low latitude cases (β -plane), the flows organize into two and three zonal jets respectively, which are anisotropic and inhomogeneous.

2.2. Passive tracer fields with a mean gradient in turbulent flows with jets

Now we consider the passive tracers $T_j, j = 1, 2$, which are advected by the upper and lower turbulent flows of (2). In particular, for the scale-selective dissipation \mathcal{DT}_j

of (1), we choose the same hyper dissipation $-\kappa\Delta^4T_j$ with a hyper diffusion coefficient κ so that the down-scale cascade of the tracer is absorbed in the smallest scales while the large-scales are marginally dissipated

$$\frac{\partial T_j}{\partial t} + \mathbf{v}_j \cdot \nabla T_j = -\kappa\Delta^4T_j. \quad (4)$$

For a tracer field with a large-scale mean tracer gradient, $T_j = T'_j + a_jx + b_jy$, the deviation tracer T'_j is advected by

$$\partial_t T'_j + \mathbf{v}_j \cdot \nabla T'_j = -a_j(\tilde{u}_j + (-1)^{j-1}) - b_j\tilde{v}_j - \kappa\Delta^4T'_j. \quad (5)$$

whose tracer variance is maintained by the zonal and meridional velocity components $\tilde{u}_j + (-1)^{j-1}$ and \tilde{v}_j respectively. The zonal velocity component $\tilde{u}_j + (-1)^{j-1}$ has non-zero mean, which is $(-1)^{j-1}$. To solve the equation (5) in a doubly-periodic domain, we also include external forcing $F_{ext}^j := a_j(-1)^{j-1}$ on the right hand side of (5), which removes the non-zero-mean term

$$\begin{aligned} \partial_t T'_j + \mathbf{v}_j \cdot \nabla T'_j &= -a_j(\tilde{u}_j + (-1)^{j-1}) - b_j\tilde{v}_j - \kappa\Delta^4T'_j + F_{ext}^j \\ &= -a_j\tilde{u}_j - b_j\tilde{v}_j - \kappa\Delta^4T'_j \end{aligned} \quad (6)$$

Note that from the linearity of (6), the deviation tracer T'_j can be reconstructed from

$$T'_j = a_j\chi_j + b_j\phi_j \quad (7)$$

where χ_j and ϕ_j are the solutions to the along-jet and across-jet tracer equations

$$\text{along-jet : } \partial_t\chi_j + \mathbf{v}_j \cdot \nabla\chi_j = -\tilde{u}_j - \kappa\Delta^4\chi_j, \quad (8)$$

$$\text{across-jet : } \partial_t\phi_j + \mathbf{v}_j \cdot \nabla\phi_j = -\tilde{v}_j - \kappa\Delta^4\phi_j. \quad (9)$$

Note that \tilde{u}_j and \tilde{v}_j are the zero-mean fluctuation parts of the total velocity field $\mathbf{v}_j = (u_j, v_j) = (\tilde{u}_j + (-1)^{j-1}, \tilde{v}_j)$.

In the same doubly periodic domain as in the flow advection equation (2), we compute a reference simulation of (8) and (9) using the same method as in the flow equation, that is, the pseudo-spectral method with the semi-implicit 3rd Runge-Kutta method. The spectral resolution uses 256×256 computational grid for both layers with a time step 2×10^{-5} . The hyper-diffusivity κ is set to 10^{-10} . Figure 2 shows the temporally and zonally averaged along-jet χ_j and across-jet ϕ_j tracer fields for both layers along with the spectra of tracer variance about the mean. In the β -plane cases ($k_\beta \neq 0$), as the along-jet tracer is forced by the zonal velocity component \tilde{u}_j with jets, tracer fields show patterns aligned with the jets. On the other hand, the across-jet tracer fields, which is forced by the meridional velocity component \tilde{v}_j with no jets, displays no significant spatial patterns aligned with the jets although it is advected by velocity fields with jets. In the high latitude case ($k_\beta = 0$), as we expect from the flows dominated

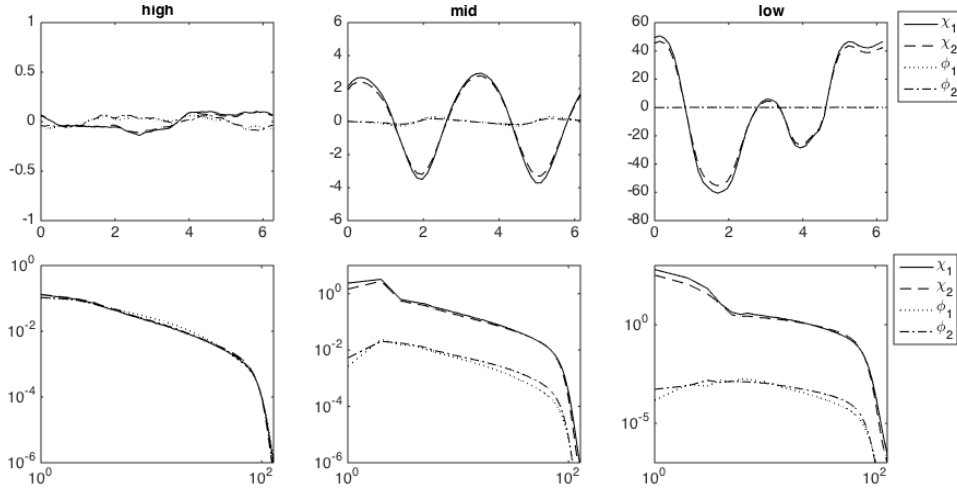


Figure 2: (Top) Temporally and zonally averaged along-jet $\chi_j, j = 1, 2$, and across-jet, $\phi_j, j = 1, 2$ tracer fields. (Bottom) Tracer variance spectra about the mean.

by isotropic small-scale fluctuations, both the along-jet and across-jet tracer fields are comparable with no spatial structure.

To check the intermittency in the tracer fields, Figure 3 shows the log-scale probability distributions of the tracer fields subtracting the mean. In the β -plane cases, the interactions of complex structures in the turbulent flows and tracer fields generate intermittency in the tracer fields. In the low latitude case, which has the strongest β -plane effect, the tracer fields in the along-jet direction, χ_j , have fat-tailed PDFs both for the upper and lower layers [11, 21, 4, 6, 1, 3, 9]. The other tracer fields in the across-jet direction, on the other hand, have almost Gaussian tails as they are forced by an isotropic meridional velocity component. If the β -plane effect diminishes, the intermittency in the along-jet tracer fields also disappear. In the high latitude case, no along-jet tracer field shows significantly fat-tailed PDFs.

3. Reduced-order forecast model

Turbulent flows contain a wide range of scales which require large computational resources and long computational time. In ensemble data assimilation, which provides statistical prediction of the state variables by combining numerical forecast models and observations, a large number of ensemble members is required to obtain accurate prior statistical description of the state. Thus it is essential to develop and apply reduced-order models for turbulent systems to reduce the computational costs and increase the ensemble numbers for accuracy. In Section 3.1, we briefly review stochastic superparameterization [20, 19, 18] which is a seamless multiscale method to generate large-scale velocity fields with significantly reduced computational complexity. In Section 3.2, the passive tracer transports in the along-jet (8) and across-jet (9) directions are approximated by additional anisotropic biharmonic eddy diffusion.

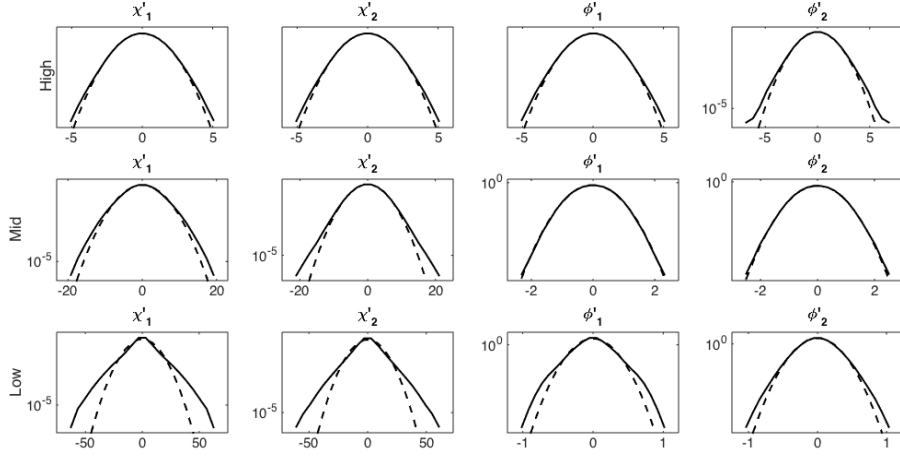


Figure 3: Probability distributions of the along-jet (χ'_j) and across-jet (ϕ'_j) tracer fields subtracting the mean. The dashed lines are Gaussian fits using the same tracer field variances. Note the log-scale in the vertical axis.

3.1. Stochastic Superparameterization

Quasigeostrophic turbulence includes regimes with a net transfer of kinetic energy from small to large scales [22], which is also confirmed in our reference simulations (see Figure 4 for energy spectra of the reference simulations along with the spectra by superparameterization). Thus it is important to accurately model the effects of the unresolved eddies, which is crucial in obtaining correct properties of the resolved large-scales such as energy spectra. To derive a reduced-order model for the quasigeostrophic turbulent flows, we apply a low-pass spatial filter denoted by $(\bar{\cdot})$ to (2) while the hyper-viscous diffusion is replaced by a biharmonic vorticity diffusion to mimic the biharmonic viscosity commonly used in eddy permitting ocean models [23]

$$\begin{aligned}\partial_t \bar{q}_1 &= -\bar{\mathbf{v}}_1 \cdot \nabla \bar{q}_1 - \overline{\bar{\mathbf{v}}_1 \cdot \nabla q_1} - \partial_x \bar{q}_1 - (k_\beta^2 + k_d^2) \bar{v}_1 - \nu \Delta^2 \bar{\omega}_1, \\ \partial_t \bar{q}_2 &= -\bar{\mathbf{v}}_2 \cdot \nabla \bar{q}_2 - \overline{\bar{\mathbf{v}}_2 \cdot \nabla q_2} + \partial_x \bar{q}_2 - (k_\beta^2 - k_d^2) \bar{v}_2 - r \Delta \bar{\psi}_2 - \nu \Delta^2 \bar{\omega}_2.\end{aligned}\tag{10}$$

Here $\omega_j = \Delta \psi_j$ is the relative vorticity. This equation is not closed under the large-scale variables due to the effects of the eddy terms $\bar{\mathbf{v}}_1 \cdot \nabla q_1$ and $\bar{\mathbf{v}}_2 \cdot \nabla q_2$. As there is an inverse energy cascade, these filtered eddy terms must be parameterized to recover the correct dynamics of the resolved large-scale variables.

The idea of superparameterization [20, 19, 18] is to find an efficient way to compute these effects. Among various versions of superparameterization, we choose the stochastic superparameterization method using Gaussian closure for the large-scale equation [20, 19]. Stochastic superparameterization models the eddy variables as randomly oriented plane waves in a direction θ and replaces the nonlinear terms in the eddy equations by additive stochastic forcing and linear deterministic damping. Thus the eddy

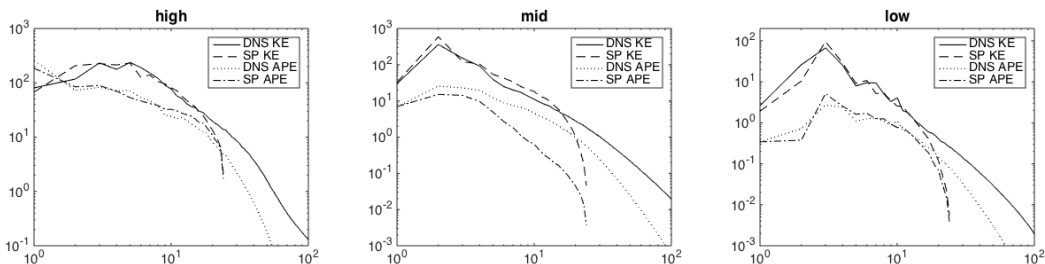


Figure 4: Total kinetic energy (KE) and available potential energy (APE) by direct numerical reference and Superparameterization simulations

equations become quasilinear conditional to the large-scale variables. By solving the quasilinear equations in formally infinite domains, this approach has no scale-gap between the resolved large-scale and unresolved small-scale while conventional superparameterization, which solves deterministic eddy equations in local periodic domains, has scale-gap in the estimation of the eddy terms. Also, the missing instability of the eddy terms in conventional superparameterization can be recovered through stochastic modeling of the eddy terms. Note that similar approximations have often been made in quasigeostrophic turbulence (see [24] and references therein) but stochastic superparameterization uses stochastic approximations only for the eddy variables.

Under this modeling of the eddy terms (randomly oriented plane waves and stochastic modeling of nonlinear dynamics of eddy variables), the effects of the eddy terms on the large-scale dynamics take the following form

$$A\mathcal{I} \left\{ \frac{1}{2}(\partial_y^2 - \partial_x^2)[F_{u,j} \sin(2\theta)] - \partial_{xy}[F_{u,j} \cos(2\theta)] + \partial_x[F_{b,j} \sin(\theta)] + \partial_y[F_{b,j} \cos(\theta)] \right\}, \quad (11)$$

where A is a tunable eddy amplitude and \mathcal{I} is a spatial smoothing operator to smooth out the smallest scales of the forcing [19]. Here the two functions $F_{u,j}$ and $F_{b,j}$ are computed using quasilinear stochastic models of eddies which are conditioned to the large-scale variables. The temporal variation of the eddy terms is modeled a random equation θ in (11),

$$d\theta = \sigma dW \quad (12)$$

where W is a Wiener process on the circle and independent of other coarse grid points. In simulations, the parameter σ is approximated using the decorrelation time of the eddy terms from the reference simulation, which yields $\sigma^2 = 250, 167.6$ and 67.5 for the high, mid and low latitude cases respectively. Note that weaker β -plane case has a larger σ as there is more energy and the subgrid-scale terms decorrelate faster. More details of the version of stochastic superparameterization used in this study can be found in [19] and [25].

The parameter, A , is tuned to match the total kinetic energy (KE) $\frac{1}{2} \int |\nabla\psi_1|^2 + |\nabla\psi_2|^2$ and the available potential energy (APE) $\frac{k_d^2}{2} \int (\psi_1 - \psi_2)^2$ which yields $A =$

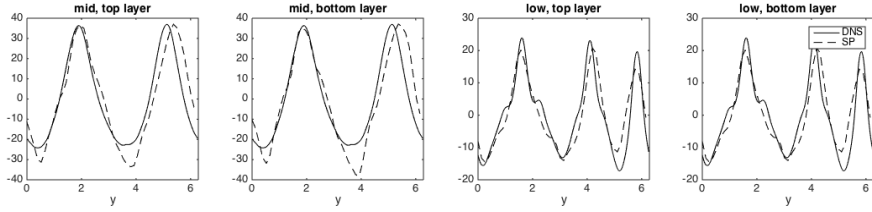


Figure 5: Temporally and zonally averaged zonal velocity components for the mid (left two) and low (right two) latitude cases

6750, 1700, and 350 for the high, mid, and low latitude cases respectively. The kinetic and potential energy spectra by stochastic superparameterization using the tuned eddy amplitude and a resolution of 48^2 for each layer are shown in Figure 4 for all three regimes. Stochastic superparameterization, which uses 250 less computational resources than the reference simulation, captures the correct energy spectra of the large-scales. The zonal jets in the β -plane cases are also captured by stochastic superparameterization (see Figure 5 for the temporally and zonally averaged zonal velocity components). By matching the energy spectra, stochastic superparameterization also provides the correct amplitude and position of the zonal jets.

3.2. Turbulent diffusion using anisotropic biharmonic eddy diffusion

The large-scale approximation of the turbulent diffusion of passive tracers is commonly obtained by adding additional diffusion, which is called ‘eddy diffusion’, while tracers are advected by large-scale velocity fields. For turbulent diffusion of passive tracers in anisotropic and inhomogeneous turbulent flows with jets, the jets act as barriers to meridional tracer transport. Thus, the transport in the along-jet direction becomes stronger than the one in the across-jet direction (see [2] for analysis of along-jet and across-jet diffusivity).

For the reduced-order models of the tracer transport in the along-jet (8) and the across-jet (9) directions, the following tracer transport model is proposed which is advected by the large-scale velocity fields from SP, $\mathbf{v}_{SP,j}$, $j = 1, 2$,

$$\text{along-jet} : \partial_t \chi_j + \mathbf{v}_{SP,j} \cdot \nabla \chi_j = -\tilde{u}_j - \kappa \Delta^4 \chi_j - \kappa_{aniso} \mathcal{D}_{aniso}^2 \chi_j, \quad (13)$$

$$\text{along-jet} : \partial_t \phi_j + \mathbf{v}_{SP,j} \cdot \nabla \phi_j = -\tilde{v}_j - \kappa \Delta^4 \phi_j - \kappa_{aniso} \mathcal{D}_{aniso}^2 \phi_j. \quad (14)$$

The eddy diffusion is modeled by an anisotropic biharmonic diffusion

$$-\kappa_{aniso} \mathcal{D}_{aniso}^2$$

where κ_{aniso} is a diffusion coefficient and the linear dissipation \mathcal{D}_{aniso} is given by

$$\mathcal{D}_{aniso} = \alpha \partial_{xx} + \partial_{yy} \quad (15)$$

with a tunable parameter α controlling the anisotropy of the diffusion. The biharmonic diffusion $-\mathcal{D}_{aniso}^2$ is commonly used in ocean models [23] to account for the effect of eddies on the large-scale dynamics.

	High	Mid	Low
κ_{aniso}	1.1E-3	3.0E-4	1.0E-4
α	1.0	1.1	1.4

Table 1: Tuned parameters, diffusion coefficient κ_{aniso} and α , of the anisotropic biharmonic eddy diffusion (15)

The reduced-order model for the tracer turbulent diffusion, (13) and (14), is solved using the same numerical method as the full-resolution simulation in Section 2 except that only 48×48 computational grids are used in each layer, which is the same resolution of the stochastic superparameterization simulation. The tunable parameters κ_{aniso} and α which determine the strength of the anisotropic biharmonic diffusion, are obtained by tuning the tracer variance spectrum. In the β -plane cases, the across-jet tracer has much smaller variance than the along-jet tracer and thus we mainly focus on matching the along-jet tracer variance. The tracer variance is shown in Figure 6 (see Table 1 for the tuned parameters). In the high latitude case, as there are no jets, the tracer variance spectra of χ_j and ϕ_j are well reconstructed from the reduced-order velocity fields by the superparameterization algorithm. In the β -plane cases (mid and low latitude cases), the along-jet tracer variance spectra show comparable results between the reference and low order simulations. For the across-jet tracer, on the other hand, it is difficult to reconstruct accurate spectra as the variance in the across-jet direction is much smaller than in the along-jet direction. The reduced-order forecast model has biases in the temporally and zonally averaged mean states shown in Figure 7. The mean states by the reduced-order model capture the meridional variation of the true signal as it is related to the velocity fields which maintain the tracer variance, but the amplitude of the mean states are not captured correctly, which is a difficult task using a reduced-order forecast model as there are model errors in the velocity fields and the tracer fields. In Section 5, we will see that the multiscale data assimilation described in the next section significantly improves the mean state estimation by incorporating noisy sparse observations.

Another important feature to generate using a reduced-order method is the tracer intermittency illustrated by the fat-tailed probability distribution functions of the tracer field. The PDFs by the direct numerical reference simulation in Section 2 and the reduced-order method are compared in Figure 8 for the representative low latitude case (the other two cases have almost Gaussian distributions). In the along-jet tracer case, which has fat-tailed PDFs, the reduced-order tracer field advected by superparameterization velocity field has PDFs comparable to the reference simulation. Also note that the low order methods have fake fat-tails in the across-jet tracer PDFs but the variance of the tracer in the across-jet direction is significantly smaller than in the along-jet direction.

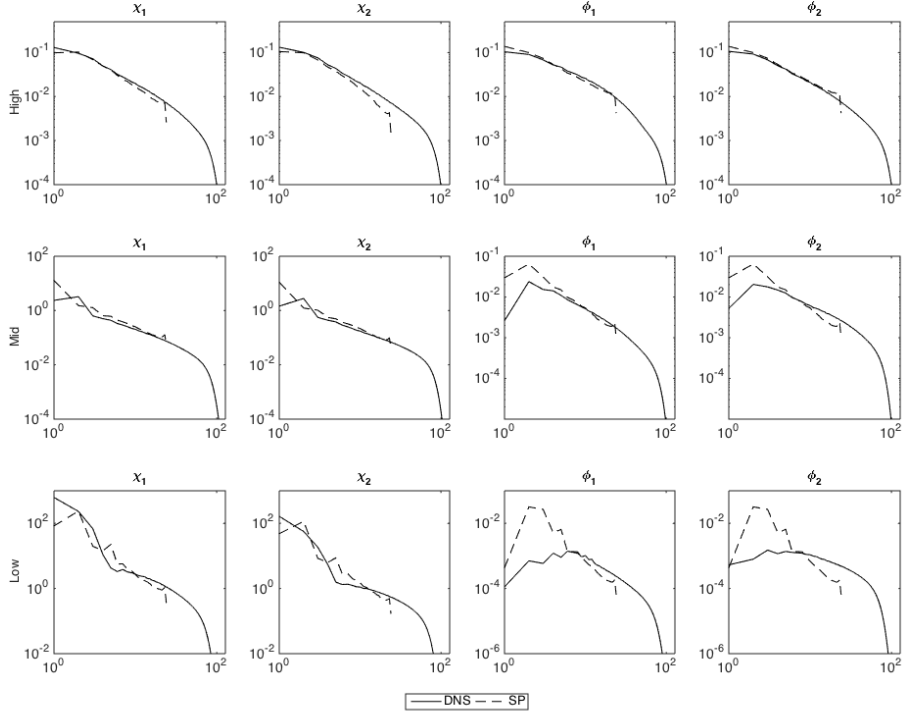


Figure 6: Tracer variance spectrum in the along-jet and across-jet directions by the direct numerical reference simulation (DNS) and reduced-order model advected by superparameterization velocity fields (SP)

4. Multiscale Data Assimilation

Though the reduced-order forecast model displays high skill in capturing many crucial features in the tracer fields, large errors may still appear due to the insufficient characterization of the unresolved small-scale dynamics. In this study we are concerned with using data assimilation of the passive tracer to further improve the imperfect model prediction skill with the observations of the tracer and the reduced-order method as the forecast model. During data assimilation process using a reduced-order model as the forecast method, which provides prior distributions only for the resolved scales, we encounter an important difficulty because the observation contains contributions both from the resolved scales and the unresolved small-scales while only large-scale predictions are available from the forecast model. To alleviate this difficulty, a general multiscale data assimilation framework was proposed in [15] providing the improved estimations for the resolved scales from the mixed observations of both the large- and small-scales. In particular, the ensemble multiscale data assimilation method [16, 15] is used to achieve accurate estimation about the large-scale statistics. In comparison with the particle method in [15], the ensemble method we use here is easy to implement and

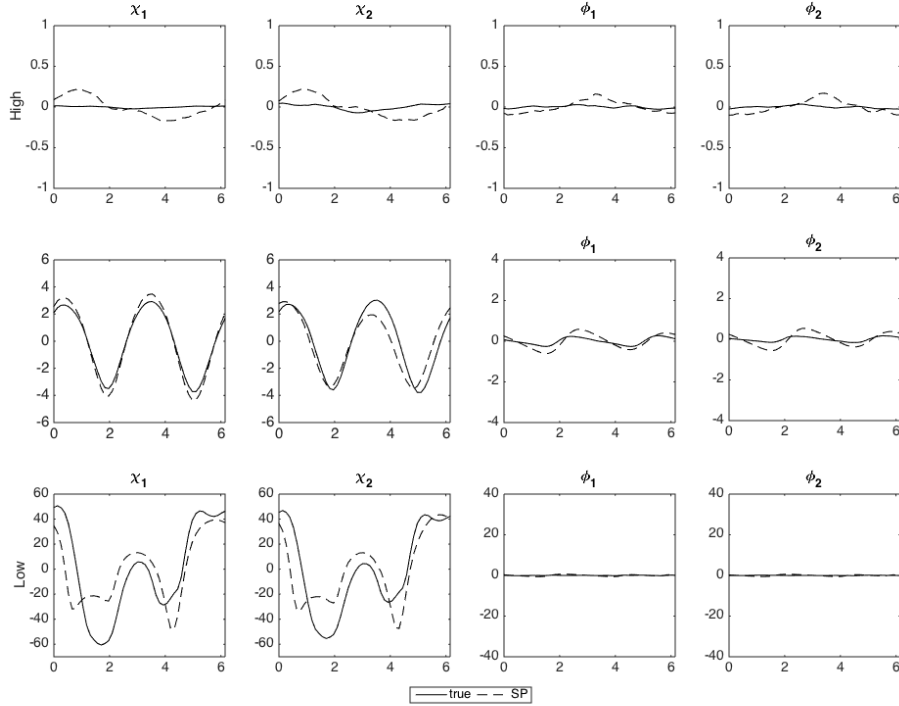


Figure 7: Temporally and zonally averaged tracer mean in the along-jet (χ_j) and across-jet (ϕ_j) directions by the direct numerical reference simulation (DNS) and reduced-order model advected by superparameterization velocity fields (SP)

can be applied to general cases when the dimension of the resolved variables is as large as the tracer transport considered here.

4.1. Treatment for multiscale observation data

For a tracer field c , let us assume that the full state has a decomposition into the following form

$$c = (\bar{c}, c') \quad (16)$$

where \bar{c} and c' represent the resolved and unresolved scale components respectively. Although the general multiscale framework of [15] can be applied to nonlinear observations, we consider the case when the observation w is linearly related to the full state

$$w = \bar{G}\bar{c} + G'c' + \xi_o. \quad (17)$$

\bar{G} and G' are observation matrix corresponding to the resolved and unresolved scales and ξ_o is observation noise which is Gaussian with zero-mean and a covariance $r_o I$ (I is the identity matrix in the observation space) uncorrelated in space and time. To derive an

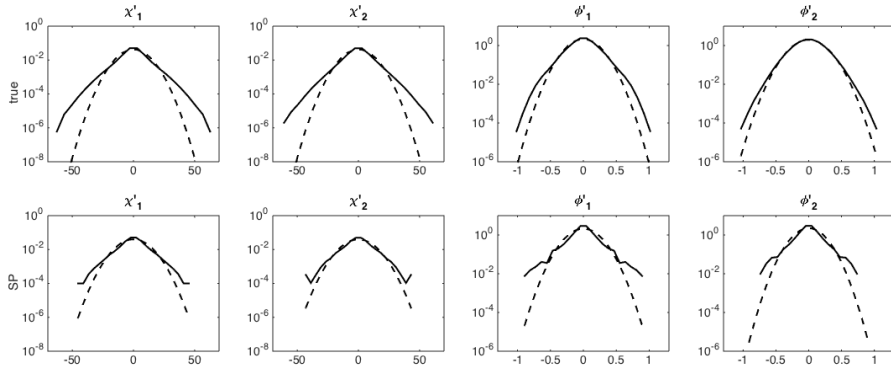


Figure 8: Log-scale probability distributions of tracer fields without (temporally and zonally averaged) mean state by the direct numerical reference simulation (DNS) and the reduced-order model advected by superparameterization velocity fields (SP). Low latitude case. χ'_j : along-jet tracer around the mean. ϕ'_j : across-jet tracer around the mean

update formula for the resolved large-scale \bar{c} , the large-scale is assumed to be Gaussian while the small scale variance is Gaussian conditional to the large-scale with covariances \bar{R} and $R'(\bar{c})$ respectively. Although the unresolved scale is Gaussian conditional to the resolved large-scale, the interaction between the resolved and unresolved scales through the dependence of the unresolved scale on the resolved scale makes interesting statistics other than Gaussian statistics. Note that the unresolved scale is driven by the resolved scale and the interaction between them makes non-trivial statistics. But it is demonstrated in [16] that for zero-mean c' conditionally Gaussian to \bar{c} , there is no correlation between \bar{c} and c' . Thus, the full state covariance R can be put in the following form

$$R = \begin{pmatrix} \bar{R} & \\ & R'(\bar{c}) \end{pmatrix}. \quad (18)$$

Under the previous assumptions, a direct application of the Bayes rule to the prior forecast state c^f with a prior covariance R^f in the form of (18) yields the following formula for the update of the resolved scale \bar{c} in the analysis step of the filtering procedure

$$\begin{aligned} \bar{c}^a &= \bar{c}^f + \bar{K}(w - \bar{G}\bar{c}) \\ \bar{K} &= \bar{R}^f \bar{G}^T (\bar{G} \bar{R}^f \bar{G}^T + G' R'^f G'^T + r_o I)^{-1} \\ \bar{R}^a &= (I - \bar{K} \bar{G}) \bar{R}^f \end{aligned} \quad (19)$$

where f and a denote the prior and posterior variables. Although the Bayes rule provides the update for the unresolved state, no update for the unresolved scale is necessary as the reduced-order model for the tracer transport does not involve the unresolved small-scale.

The formula (19) is the standard Kalman update formula for \bar{c} with the same observation w and an increased observation variance $G' R'^f G'^T + r_o I$ is added to the Kalman

gain matrix due to the effect from the unresolved scales. The difference being the additional observation error $G'R'G'^T$ from the unresolved scale is known as the “representative error” [26]. In other words, the resolved scales trust the forecast more than without the representative error. To update \bar{c} we need the covariance of the unresolved scales which is not provided by the forecast model. Instead, we employ the climatological variance of the unresolved scales and use it to reconstruct a diagonal covariance matrix of the unresolved scales, that is, the variances of the unresolved scales at each coarse grid points and different observation time set to be identical to the climatological value. The covariance matrix of the unresolved scale is diagonal thus the additional observation error covariance also becomes diagonal. With this constant additional observation variance at each observation points, we can use standard ensemble data assimilation methods to update \bar{c} using (19) without any difficulties.

4.2. Covariance inflation using additive noise

The reduced-order model is an imperfect forecast model and thus model errors are unavoidable in the data assimilation. Imperfect model errors yield underestimation of the uncertainty in the forecast model and thus filter trusts the forecast more confident than observations. The covariance inflation in ensemble methods is an effective remedy of particle collapse and mitigates the problem related to model errors which increases the uncertainty in the forecast model by inflating the prior covariance [27, 28]. Here we employ an additional covariance inflation which adds noise to the prior ensemble members

$$\bar{c}^k \leftarrow \bar{c}^k + \xi_a \quad (20)$$

where ξ_a is Gaussian with zero mean and a variance r_a^2 and i.i.d in space and time. This is equivalent to adding $r_a^2 I$ to the prior covariance matrix \bar{P}^f . The additional noise can be regarded as the zero-correlation time limit of the random forcing given by a OU process which is commonly used to improve the controllability and thus the stability of the filter [14]. In the next section, we see that the covariance inflation using additive noise plays an important role in improving the filtering skills.

Now we summarize the general strategy for the multiscale filtering with SP **Algorithm** (Multiscale SP filter)

1. Forecast step
 - (a) Run SP flow model following (10) and (11)
 - (b) Get tracer field prediction from (13) with biharmonic diffusion in (14)
2. Analysis step
 - (a) Inflate the covariance with additive noise (20)
 - (b) Update the prediction with the observation data using proper DA scheme (19)

High	Decor time	Total var	Large-scale var	Small-scale var
χ_1	0.013	1.40	1.09	0.31
χ_2	0.014	1.46	1.13	0.33
ϕ_1	0.015	1.68	1.35	0.33
ϕ_2	0.014	1.53	1.21	0.32
Mid	Decor time	Total var	Large-scale var	Small-scale var
χ_1	0.017	20.39	17.58	2.81
χ_2	0.018	18.22	15.44	2.78
ϕ_1	0.010	0.37	0.30	0.07
ϕ_2	0.009	0.46	0.35	0.11
Low	Decor time	Total var	Large-scale var	Small-scale var
χ_1	0.022	1448	1415	33
χ_2	0.020	1412	1378	34
ϕ_1	0.008	0.040	0.031	0.009
ϕ_2	0.007	0.042	0.034	0.008

Table 2: Stationary state statistical properties of the reference simulation

5. Filter performance

The ensemble multiscale data assimilation described in the previous section is tested for the passive tracer in the along-jet (8) and the across-jet (9) directions. The true signal is generated by the direct numerical method as in Section 2, a fine resolution of 256×256 grid points in each layer and a time step of 2×10^{-5} . The reduced-order prediction models for both the two-layer quasigeostrophic equation (2) and the tracer model use a coarse resolution of 48×48 for each layer and a time step of 1×10^{-4} . Thus the computational cost for the forecast model is 250 times less than that of the direct full-resolution method. From the reference true signal, we define the large-scale variable as the Fourier truncated variable with a cutoff wavenumber $k = 24$ which is the Nyquist wavenumber of the coarse grid points.

In Table 2, the stationary state statistical properties of the true signal are shown for the three different test regimes. As the β -plane effect (k_β) increases, the along-jet tracer fields have longer decorrelation times and larger variances than the across-jet tracer fields. We observe only the upper layer tracer fields on a 48×48 network coinciding with the coarse grid points of the forecast model which is sparse compared to the full true signal. The tracer dynamics in the lower layer is completely unobserved. Based on the stationary state properties of the true signal, the observation intervals are set to 0.01 for all three test regimes, which are comparable to or shorter than the decorrelation time of the along-jet tracer fields. Raw observation errors are 0.07, 1.00 and 72 for the high, mid and low latitude cases corresponding to 5% of the upper layer tracer field variances. In the estimation of the large-scale variables, the small-scale variables add additional observation errors and thus the effective observation errors for the large-scale

variables are given by a combination of these two parts

$$\text{effective obs error} = \sqrt{\text{raw obs error variance} + \text{small-scale variance}}. \quad (21)$$

We expect that the qualitative results are not strongly dependent on a particular choice of ensemble Kalman filters. In this study the ensemble adjustment Kalman filter [29] is applied to update the large-scales in the analysis step (19) with a sequential update of observations and an additional observation error variance given by the small-scale variance. 50 ensemble members are used and as observations are available at each coarse grid point, an extreme localization which ignores influences from other grid points is applied. Each test runs 500 assimilation cycles and the first 200 cycles are ignored in measuring filter performance. Filter performance is measured using two statistics on the posterior mean estimate; time-averaged RMS errors and time-averaged pattern correlations (PC) between the posterior mean and the true signal. For the large-scale posterior mean \bar{c}_m^a and true signal \bar{c}_m^t at the m -th cycle, the time-averaged RMSE and PC are defined as

$$\text{time-averaged RMS error} = \frac{1}{M - m_0} \sum_{m=m_0+1}^M \|\bar{c}_m^a - \bar{c}_m^t\| \quad (22)$$

and

$$\text{time-averaged PC} = \frac{1}{M - m_0} \sum_{m=m_0+1}^M \frac{\sum_{x_i} \bar{c}_m^a \bar{c}_m^t}{\|\bar{c}_m^a\| \|\bar{c}_m^t\|} \quad (23)$$

where $\|\cdot\|$ represents the l_2 norm on the coarse grid points $\{x_i\}$, $m_0 = 200$ and $M = 500$.

5.1. Filter performance without additional covariance inflation

First we present filtering results without additional covariance inflation. Time-averaged RMS errors and pattern correlations are shown in the first column of Table 3 for all three test regimes. The upper layer effective observation error (21) and standard deviation of the stationary state large-scale variables are in the third and fourth columns respectively. Note that the standard deviation of the stationary state large-scale variables is an expected error when the prediction trusts the spatially homogeneous mean 0.

Without covariance inflation, the data assimilation method is unable to generate accurate posterior estimates for the large-scale variables. In the high latitude case RMS errors are worse than the standard deviation of the stationary state with no meaningful correlations between the posterior mean and the true signal. As the β -plane effect increases which has stronger and more stable zonal jet structures, the pattern correlations increases. This result is expected as the prediction model better captures the anisotropic and inhomogeneous structures as the β -plane effect increases. But still the RMS errors are worse than or comparable to the standard deviation of the stationary state.

High	without inflation	with inflation	Effect obs error	stationary std
χ_1	1.56 (0.01)	0.51 (0.87)	0.60	1.04
χ_2	1.61 (-0.01)	0.92 (0.65)	N/A	1.06
ϕ_1	1.65 (-0.00)	0.55 (0.88)	0.62	1.16
ϕ_2	1.70 (0.00)	0.99 (0.57)	N/A	1.10
Mid	without inflation	with inflation	Effect obs error	stationary std
χ_1	4.55 (0.15)	1.75 (0.90)	1.95	4.19
χ_2	4.21 (0.13)	2.83 (0.77)	N/A	3.93
ϕ_1	0.91 (0.11)	0.28 (0.88)	0.29	0.55
ϕ_2	0.85 (0.10)	0.52 (0.60)	N/A	0.60
Low	without inflation	with inflation	Effect obs error	stationary std
χ_1	25 (0.65)	6.77 (0.98)	10.24	37.62
χ_2	23 (0.61)	17.1 (0.96)	N/A	21.50
ϕ_1	0.18 (0.08)	0.10 (0.86)	0.10	0.18
ϕ_2	0.29 (-0.01)	0.68 (0.57)	N/A	0.24

Table 3: RMS errors and pattern correlations in parenthesis of the posterior mean in the estimation of the large-scale variables

5.2. Filter performance with additional covariance inflation

Now we consider the results using the additive covariance inflation (20) in order to recover accurate filtering skill. As the along-jet and across-jet tracer fields are independent of each other and the tracer fields have different statistical properties in different directions, the additional variance levels r_a^2 are picked depending on the different directions. In the along-jet direction tracer field, r_a^2 are set to 1.5, 10 and 25 for the high, mid and low latitude cases. In the across-jet direction tracer field, r_a^2 are set to 1.5, 0.3 and 0.01 for the high, mid and low latitude cases (note that tracer fields are isotropic in the high latitude case). These values are comparable to the variances of the fluctuation tracer field components after subtracting the temporally and zonally averaged mean state. The additional noise is independent at different coarse grid points but to recover skillful estimation in the lower layer tracer field, the upper layer and lower layer use the same additional noise which yields perfectly correlated noise between different layers.

As can be found in the second column of Table 3, additive covariance inflation using additional noise significantly improves the filtering skills in the estimation of the large-scale variables. In the high latitude case, the upper layer tracer field estimation have RMS errors more than 10% smaller than the effective observation errors while the posterior mean and the true signal are correlated more than 87%. The lower layer tracer field estimation is not as skillful as the upper layer as there is no direct observation of the lower layer. But the posterior RMS errors are smaller than the standard deviation of the stationary state. Also the estimation is correlated with the true signal more than 55%.

If the β -plane effect increases, skill in the along-jet tracer field estimation increases.

In the mid latitude case, the posterior estimate and the true signal are correlated more than 90% and the RMS error is about 10% smaller than the effective observation error. In the low latitude case, the filtering performance is even better; RMS error is more than 30% smaller than the effective observation and pattern correlation is larger than 98%. Using the identical additional noise for both layers, the lower layer estimation of the along-jet tracer field has smaller errors than the standard deviation of the stationary state and high pattern correlations. Note that if the additional noise are uncorrelated between different layers, the filter shows no skill in the lower layer estimation (results are not shown here).

The large-scale estimation in the across-jet tracer fields (of the mid and low latitude cases), on the other hand, has no improved filtering skill even with the additive covariance inflation. This result is expected from the poor performance of the reduced-order forecast model in recovering the tracer variance spectra in Section 3.2. The across-jet tracer fields have very small variances in comparison with the along-jet tracer fields and thus it is a difficult task to recover accurate filtering skill for the across-jet tracer fields. The RMS errors are comparable to or worse than the effective observation error (when it is available) or the standard deviation of the stationary state although still the pattern correlations of the observed upper layer field is higher than 85%.

We have seen in Section 3.2 that the forecast model without filtering has model errors, the biased temporally and zonally averaged mean states with incorrect amplitudes (Figure 7). Due to the model errors by an imperfect forecast model, it was essential to use additive covariance inflation to achieve meaningful filtering skill. Figure 9 shows the temporally and zonally averaged mean state of the posterior estimate with the additive covariance inflation. The filtering improves the mean state estimation significantly for the upper layer tracer fields which are on top of the true signal without any significant bias. The unobserved lower layer tracer fields have biased mean states compared to the observed upper layer but the filtered lower layer mean state estimation is significantly improved in comparison with the estimation without filtering.

In the low latitude case, the along-jet tracer fields show non-Gaussian fat tails. In Section 3 we see that the reduced-order prediction model is able to generate non-Gaussian fat tails but the tails are weaker and narrower than the true signal (Figure 8), which implies the prediction model has narrower range of variability than the true signal and thus covariance inflation is necessary in recovering meaningful filtering skill. Figure 10 shows the log-scale PDFs of the posterior estimates in the low latitude case. The filtered signal recovers the full range of the true signal with non-Gaussian fat tails in the along-jet tracer fields. In the filtered posterior across-jet tracer fields, where the prediction model generates fake fat-tails in Figure 8, the fake fat-tails are corrected in the posterior estimation.

6. Conclusions

This paper has been concerned with data assimilation of passive tracer advected by turbulent flows with jets. The forecast is given by a computationally cheap but robust

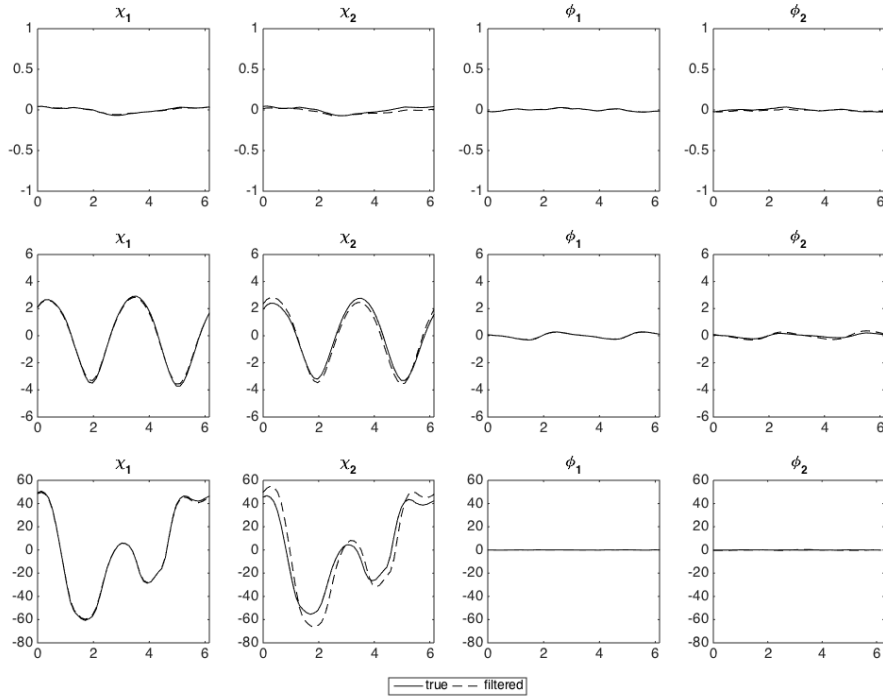


Figure 9: Temporally and zonally averaged tracer in the along-jet (χ_j) and across-jet (ϕ_j) directions of true signal and filtered posterior mean

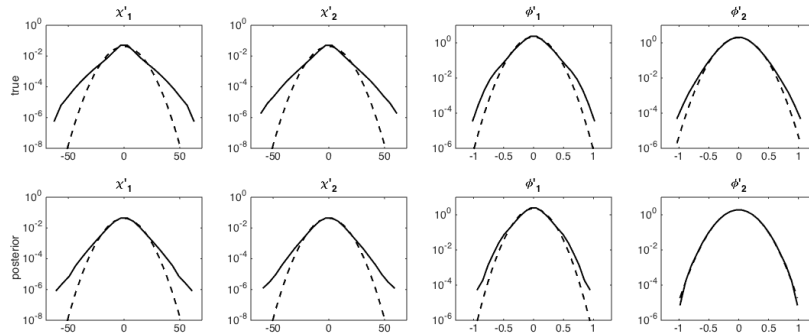


Figure 10: Log-scale probability distributions of tracer fields without (temporally and zonally averaged) mean state. True signal (top row) and posterior estimation (bottom row). Low latitude case. χ'_j : along-jet tracer around the mean. ϕ'_j : across-jet tracer around the mean

reduced-order method which provides predictions for the large-scale component resolved on coarse grid points. To generate large-scale velocity fields, stochastic superparameterization is applied to the two-layer quasigeostrophic equation which is an idealized

model for geophysical turbulent flows in atmosphere and ocean. The large-scale model of the passive tracer transport is then advected by the superparameterization large-scale velocity fields with an additional eddy diffusion term. In the f -plane case when the velocity fields are isotropic with no jets, the eddy diffusion is accurately modeled by an isotropic biharmonic diffusion. In the β -plane case with jets, an anisotropic diffusion is necessary as the jets act as barriers to meridional tracer transport; the zonal diffusion is much stronger than the meridional diffusion.

The reduced-order forecast method reduces the computational complexity significantly compared to a direct numerical forecast model but it requires a treatment for mixed observations of the resolved and unresolved scale components in applying data assimilation. In this study we have employed an ensemble multiscale data assimilation method which provides the large-scale update using the mixed observations. The multiscale data assimilation method treats the unresolved scale component of the tracer as representative error which enables us to use conventional single-scale ensemble filters. We have tested the multiscale data assimilation for three different test regimes with or without jets and numerical results show skillful estimation of the large-scale variable; the intermittency in the low latitude case is also captured by the reduced-order data assimilation approach. Due to large dissipation and model error in the reduced-order forecast method, covariance inflation by additional noise has played an crucial role in obtaining skillful filtering results. Using additional noise correlated between different layers, the data assimilation also achieves large-scale estimation skill for the unobserved lower layer tracer field using the observation of the upper layer tracer field.

An improved reduced-order prediction skill of tracer fields by empirical information theory has been developed [9] which fits the autocorrelation function of tracer fields and finds optimal model parameters in a training phase. Thus it is natural to speculate whether this improved prediction skill can achieve better filtering skills than the reduced-order model using eddy diffusion studied in this paper. Also there is an adaptive inflation technique which improves stability preventing catastrophic filter divergence and also increases filtering skills [30]. This technique could be another approach to improve the filtering performance by adaptively controlling the additional noise level which is essential to obtain meaningful filtering skills in our study.

The incorporation of the reduced-order forecast method for the ensemble multiscale method has been limited to the large-scale estimation as there is no prediction method for the unresolved small-scales. We have modeled the unresolved small-scales as stationary spatially homogeneous variables and thus used the climatological variance as the representative error in the multiscale data assimilation. We expect that an improved filtering performance could be obtained by using efficient models also for the unresolved scales such as linear stochastic models [14] or dynamic stochastic super-resolution methods [31].

Acknowledgements

The research of A.J. Majda is partially supported by Office of Naval Research grant ONR MURI N00014-12-1-0912 and DARPA 25-74200-F4414. Y. Lee is supported as a postdoctoral fellow by these grants. D. Qi is supported as a graduate research assistant by the ONR grant.

- [1] P. R. K. A. J. Majda, Simplified models for turbulent diffusion: theory, numerical modelling, and physical phenomena, *Physics reports* 314 (1999) 237–574.
- [2] K. S. Smith, Tracer transport along and across coherent jets in two-dimensional turbulent flow, *J. Fluid Mech.* 544 (2005) 133–142.
- [3] J. D. Neelin, B. R. Lintner, B. Tian, Q. Li, L. Zhang, P. K. Patra, M. T. Chahine, S. N. Stechmann, Long tails in deep columns of natural and anthropogenic tropospheric tracers, *Geophysical Research Letters* 37, 105804. doi:10.1029/2009GL041726.
- [4] J. P. Gollub, J. Clarke, M. Gharib, B. Lane, O. N. Mesquita, Fluctuations and transport in a stirred fluid with a mean gradient, *Phys. Rev. Lett.* 67 (1991) 3507–3510.
- [5] A. Bourlioux, A. Majda, Elementary models with probability distribution function intermittency for passive scalars with a mean gradient, *Physics of Fluids* 14 (2002) 881–897.
- [6] A. J. Majda, B. Gershgorin, Elementary models for turbulent diffusion with complex physical features : eddy diffusivity, spectrum and intermittency, *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* 371 (1982). doi:10.1098/rsta.2012.0184.
- [7] A. J. Majda, X. T. Tong, Intermittency in turbulent diffusion models with a mean gradient, *Nonlinearity* 28 (2015) 4171–4208.
- [8] J. L. H. Tennekes, *A First Course in Turbulence*, MIT Press, Cambridge, MA, 1972.
- [9] D. Qi, A. J. Majda, Predicting fat-tailed intermittent probability distributions in passive scalar turbulence with imperfect models through empirical information theory, *Comm. Math. Sciences* Accepted.
- [10] S. Bauer, M. S. Swenson, A. Griffa, Eddy mean flow decomposition and eddy diffusivity estimates in the tropical pacific ocean: 2. results, *J. Geophys. Res.* 107C (2002) 3154–3172.
- [11] G. Holloway, S. S. Kristmannsson, Stirring and transport of tracer fields by geostrophic turbulence, *J. Fluid Mech.* 141 (1984) 27–50.

- [12] P. Bartello, G. Holloway, Passive scalar transport in β -plane turbulence, *J. Fluid Mech.* 223 (1991) 521–536.
- [13] K. S. Smith, G. Boccaletti, C. C. Henning, I. N. Marinov, C. Y. Tam, I. M. Held, G. K. Vallis, Turbulent diffusion in the geostrophic inverse cascade, *J. Fluid Mech.* 469 (2002) 13–48.
- [14] A. J. Majda, J. Harlim, *Filtering complex turbulent systems*, Cambridge University Press, 2012.
- [15] Y. Lee, A. J. Majda, Multiscale methods for data assimilation in turbulent systems, *Multiscale Modeling and Simulation* 13 (2015) 691–173.
- [16] I. Grooms, Y. Lee, A. J. Majda, Ensemble kalman filters for dynamical systems with unresolved turbulence, *Journal of Computational Physics* 273 (2014) 435–452.
- [17] I. Grooms, A. J. Majda, Efficient stochastic superparameterization for geophysical turbulence, *Prod. Natl. Acad. Sci. USA* 110 (2013) 4464–4469.
- [18] A. J. Majda, I. Grooms, New perspectives on superparameterization for geophysical turbulence, *Journal of Computational Physics*, 271 (2014) 60–77.
- [19] I. Grooms, Y. Lee, A. J. Majda, Numerical schemes for stochastic backscatter in the inverse cascade of quasigeostrophic turbulence, *Multiscale Modeling and Simulation* 13 (2015) 1001–1021.
- [20] I. Grooms, A. J. Majda, Stochastic superparameterization in quasigeostrophic turbulence, *J. Comput. Phys.* 271 (2014) 78–98.
- [21] B. Gershgorin, A. J. Majda, Filtering a statistically exactly solvable test model for turbulent tracers from partial observations, *Journal of Computational Physics* 230 (2011) 1602–1638.
- [22] J. G. Charney, Geostrophic turbulence, *J. Atmospheric Sci.* 28 (1971) 1087–1095.
- [23] S. M. Griffies, R. W. Hallberg, Biharmonic friction with a smagorinsky-like viscosity for use in large-scale eddy-permitting ocean models, *Monthly Weather Review* 128 (2000) 2935–2946.
- [24] T. Delsole, Stochastic models of quasigeostrophic turbulence, *Surv. Geophys.* 25 (2004) 107–149.
- [25] I. Grooms, Y. Lee, A. J. Majda, Ensemble filtering and low-resolution model error: Additive inflation, stochastic parameterization, and model numerics, *Mon. Wea. Rev.* 143 (2015) 3912–3924.
- [26] S. Cohn, An introduction to estimation theory, *J. Meteorol. Soc. Jpn.* 75 (1975) 257–288.

- [27] J. Anderson, S. Anderson, A monte carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts, *Mon. Wea. Rev.* 127 (1999) 2741–2758.
- [28] H. Mitchell, P. Houtekamer, An adaptive ensemble kalman filter, *Mon. Weather Rev.* 128 (2000) 416–433.
- [29] J. Anderson, An ensemble adjustment kalman filter for data assimilation, *Mon. Wea. Rev.* 129 (2001) 2884–2903.
- [30] X. T. Tong, A. J. Majda, D. Kelly, Nonlinear stability of the ensemble kalman filter with adaptive covariance inflation, *Comm. Math. Sci* in press, arXiv:1507.08319.
- [31] M. Branicki, A. J. Majda, Dynamics stochastic superresolution of sparsely observed turbulent systems, *J. Comput. Phys.* 241 (2013) 333–363.