Gyres and Jets: Inversion of Tracer Data for Ocean Circulation Structure

RADU HERBEI
The Ohio State University, Columbus, Ohio

IAN W. MCKEAGUE
Columbia University, New York, New York

KEVIN G. SPEER
The Florida State University, Tallahassee, Florida

(Manuscript received 3 May 2007, in final form 31 October 2007)

ABSTRACT

This paper describes a quasi-3D Bayesian inversion of oceanographic tracer data from the South Atlantic Ocean. Initially, one *active* neutral-density layer is considered with an upper and lower boundary. The available hydrographic data are linked to model parameters (water velocities, diffusion coefficients) via a 3D advection–diffusion equation. A robust solution to the inverse problem can be obtained by introducing prior information about parameters and modeling the observation error. This approach estimates both horizontal and vertical flow as well as diffusion coefficients. A system of alternating zonal jets is found at the depths of the North Atlantic Deep Water, consistent with direct measurements of flow and concentration maps. A uniqueness analysis of the model is performed in terms of the oxygen consumption rate. The vertical mixing coefficient bears some relation to the bottom topography even though the authors do not incorporate topography into their model. The method is extended to a multilayer model, using thermal wind relations weakly in a local fashion (as opposed to integrating the entire water column) to connect layers vertically. Results suggest that the estimated deep zonal jets extend vertically, with a clear depth-dependent structure. The vertical structure of the flow field is modified by the tracer fields relative to the a priori flow field defined by thermal wind. The velocity estimates are consistent with independent observed flow at the depths of the Antarctic Intermediate Water; at still shallower depths, above the layers considered here, the subtropical gyre is a significant feature of the horizontal flow.

1. Introduction

Directly observed flow at the depths of the North Atlantic Deep Water in the South Atlantic Ocean shows a system of alternating zonal jets (Hogg and Owens 1999). Deep zonal flow has been explained, for example, using a coarse wind-driven circulation model in the Pacific Ocean (Nakano and Suginohara 2002). Several numerical models of varying resolution of the South Atlantic Ocean have been used to study the origin of the zonal flows, leading to the conclusion that wind is the primary force that generates zonal flow in the Brazil Basin (Treguier et al. 2003). Whether these zonal jets are wind driven or not, it has become clear that a better knowledge and understanding of their horizontal and vertical structure is important to an explanation of their origins and dynamics. Recently, Maximenko et al. (2005) and Richards et al. (2006) have used satellite data and higher-resolution numerical models to investigate the distribution and persistence of zonal flow structures in the Pacific and World Oceans.

The present paper takes an indirect path to the study of the horizontal and vertical structure of ocean circulation. Rather than using numerical models based on primitive equations (i.e., GCMs), we start with observed tracer concentrations and first-principle dynamics to find an estimate of the steady-state circulation. We attempt to distinguish the turbulent component (random mixing) from the advective component. We recognize that this distinction may be dependent on the particular inversion model used and its resolution, that is, that it depends on the method itself. However, we

Corresponding author address: Radu Herbei, Department of Statistics, The Ohio State University, 1958 Neil Avenue, Columbus, OH 43210.
E-mail: herbei@stat.osu.edu

DOI: 10.1175/2007JPO3835.1

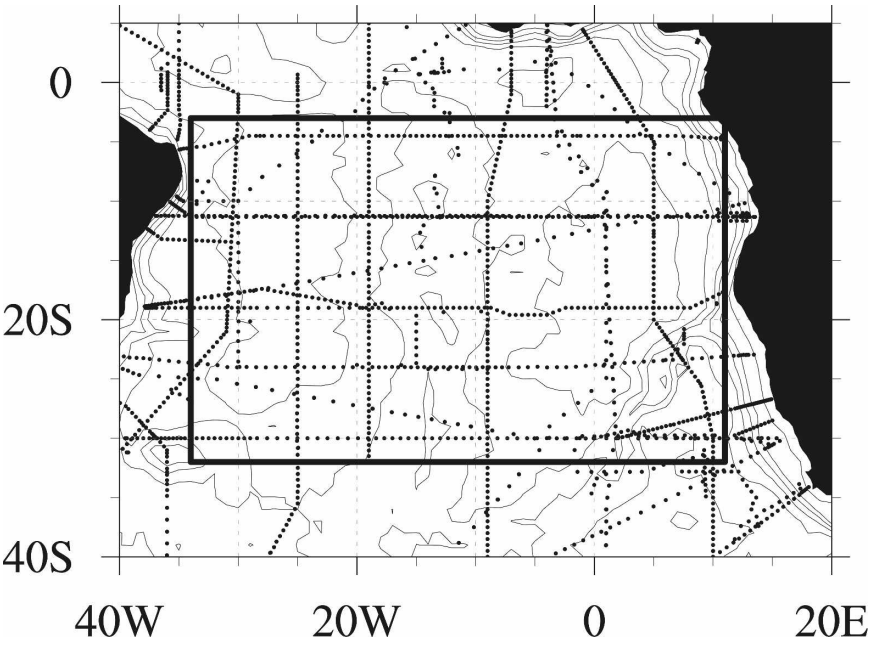

FIG. 1. Lateral boundaries of the inversion domain $\mathcal{S}$. Black dots indicate sites where data have been collected. Background shading represents bathymetry.

are able to explore this dependency and proceed on the principle that the observed basin-scale tracer distributions operationally define our scale separation. Our results will show that with this definition, a system of zonal jets appear with significant vertical structure, whose horizontal and vertical scales are controlled by the tracer distributions.

In an attempt to extract the signal associated with the mean circulation, oceanographers have found that spatial averaging over scales exceeding the mesoscale yields tracer concentration maps thought to be representative of the mean state. The scarcity of the data, though, usually requires some type of interpolation over the domain of interest. This gives only crude estimates for tracer concentrations in regions where no data are available, and the resulting maps are not necessarily consistent with physical principles.

McKeague et al. (2005) introduced a statistical approach to the estimation of quasi-horizontal flow in an abyssal neutral-density layer based on tracer concentration data. Throughout this paper we will refer to this approach as the *single-layer inversion method* or *2D inversion.* The estimated horizontal flow in this case was found to be primarily zonal, with alternating east–west jets, mainly in the western South Atlantic. The present work extends the single-layer methodology by adding vertical structure. We develop a quasi-3D Bayesian model to reconstruct vertical flow, as well as horizontal flow, and provide improved maps for tracer concentrations. A dynamical analysis of the model as well as uniqueness and accuracy issues are also discussed. The Bayesian approach allows us to construct a probability density (up to a normalizing constant) over the space of all possible solutions. This high-dimensional density provides information regarding the parameters to be estimated. Point estimates as well as uncertainty measures could be easily constructed (in principle) using this distribution. However, the normalizing constant is analytically intractable and numerical simulations are required to produce samples ("snapshots") from the posterior distribution. We present the posterior mean as the "solution" and posterior standard deviation as "uncertainty," although many other choices are possible (quantiles, modes, etc.).

Figure 1 presents the boundaries of our inversion domain, which we define as

$$\mathcal{S} = \{(x, y, z): x \in [34°\text{W}, 11°\text{E}], \quad y \in [32°\text{S}, 3°\text{S}], \quad \gamma(x, y, z) \in [28.02, 27.96]\},$$

where $\gamma(x, y, z)$ denotes the neutral density (McDougall 1987). The black dots indicate sites where tracer concentration measurements are available from ships. These bias-corrected hydrographic data [corrected for cruise-dependent standards and offsets following Thurnherr and Speer (2004)] are processed in a similar way to what is described in the 2D inversion method,

but this time we are considering several neutral-density layers. Initially our main focus is on the $\gamma = 28.00$ isopycnal (which we will call the "active" layer) and we use several layers above and below the active layer to specify the vertical structure. We will refer to this approach as a "quasi-3D" inversion since results will be available for a single isopycnal, but we show results for several different choices of targeted layers (e.g., Fig. 19). We then extend the methodology to a multilayer inversion model.

The domain $\mathcal{S}$ is discretized using a regular lattice $\mathcal{S}_G$ consisting of three 37×19 grids $\mathcal{L}_G$ stacked on top of one another, with each layer representing an isopycnal. The three neutral-density layers represent one above ($\gamma = 27.96$) and one below ($\gamma = 28.02$) the active layer ($\gamma = 28.00$), roughly corresponding to depths of 1750, 2000, and 2250 m, respectively. Figure 2 displays oxygen and salinity concentrations obtained by interpolating (using the MATLAB function griddata with option "nearest") the data onto the grid $\mathcal{L}_G$ for each layer. There is significant variation in the spatial distribution of the tracer concentration among these layers, indicating that adding vertical structure to the single-layer method would yield a better understanding of the abyssal circulation in this region. Objective mapping could be used to obtain more reliable maps than those in Fig. 2, but our aim is to relate estimates of tracer concentrations to basic physical principles as well. In Fig. 3, vertical profiles for oxygen and salinity measurements at two different locations in the inversion domain are displayed. The vertical structure around 2000-m depth varies among different tracers and across the domain as well, with local extremes in various places and significant concentration gradients (of the first and second order) in other locations, suggesting that vertical mixing terms may be significant. In our inversion we also include silica (SiO_2), potential temperature, and large-scale potential vorticity tracers.

Our findings reveal horizontal flow with zonal jets extending across the entire domain, a feature that is not clearly apparent using the single-layer inversion method, and that we attribute to intensified mixing over the Mid-Atlantic Ridge. This allows the zonal jets to maintain tracer balance across the ridge and to persist in the eastern basin. By taking the analysis further with a multilayer model, with each layer corresponding to an isopycnal, the vertical link between layers will be constructed using thermal wind relations. Rather than integrating thermal wind vertically over the entire water column, we impose it weakly and locally. An estimate of the vertical structure of these zonal flows is then available and we compare it to the flow simply derived by integrating thermal wind relations by using quasi-3D estimates as initial values.

The paper is organized as follows: in section 2 we develop a quasi-3D model involving a single active layer with vertical structure determined by boundary conditions above and below the active layer, along with the proposed Bayesian inversion technique and the Markov chain Monte Carlo (MCMC) computational procedure. Inversion results and their analysis are presented in section 3, along with a discussion of the uniqueness and accuracy of the inversion. In section 4, the proposed approach is extended to multiple active layers. Concluding remarks are presented in section 5 and technical details regarding the sampler used to perform the numerical simulations are in the appendix.

2. Quasi-3D inversion

a. Geostrophic flow model

In this section we build the quasi-3D flow model starting from a single horizontal active layer onto which we add vertical advection and diffusion, with boundary conditions for the tracers in layers above and below the layer of interest.

We begin by recalling the single-layer dynamics, which completely ignore vertical structure. The following standard notation is used: 3D water velocity $\underline{\mathbf{u}} = (u, v, w)$, in situ density ρ, pressure p, f denotes the Coriolis parameter, and $\beta = \partial f/\partial y$. The geostrophic flow model in the 2D inversion is based on exact geostrophic balance

$$-\rho f v = -\frac{\partial p}{\partial x}, \quad \rho f u = -\frac{\partial p}{\partial y}.$$

Cross-differentiating and subtracting these two equations, we get

$$\frac{\partial \rho f v}{\partial y} + \frac{\partial \rho f u}{\partial x} = 0.$$

Neglecting density variations (using the Boussinesq approximation) and following the reasoning from the single-layer inversion, under the assumption that zonal flow vanishes along the eastern boundary [$u(x_e, y) = 0$)] we can express u as

$$u(x, y) = \int_x^{x_e} \left(\frac{\partial v}{\partial y} + \frac{\beta}{f} v \right) dx', \tag{1}$$

where x_e is the longitude of the eastern boundary of the inversion domain.

For the quasi-3D flow model, vertical structure is specified through an advection–diffusion equation for the neutral density γ, obtained by neglecting the non-

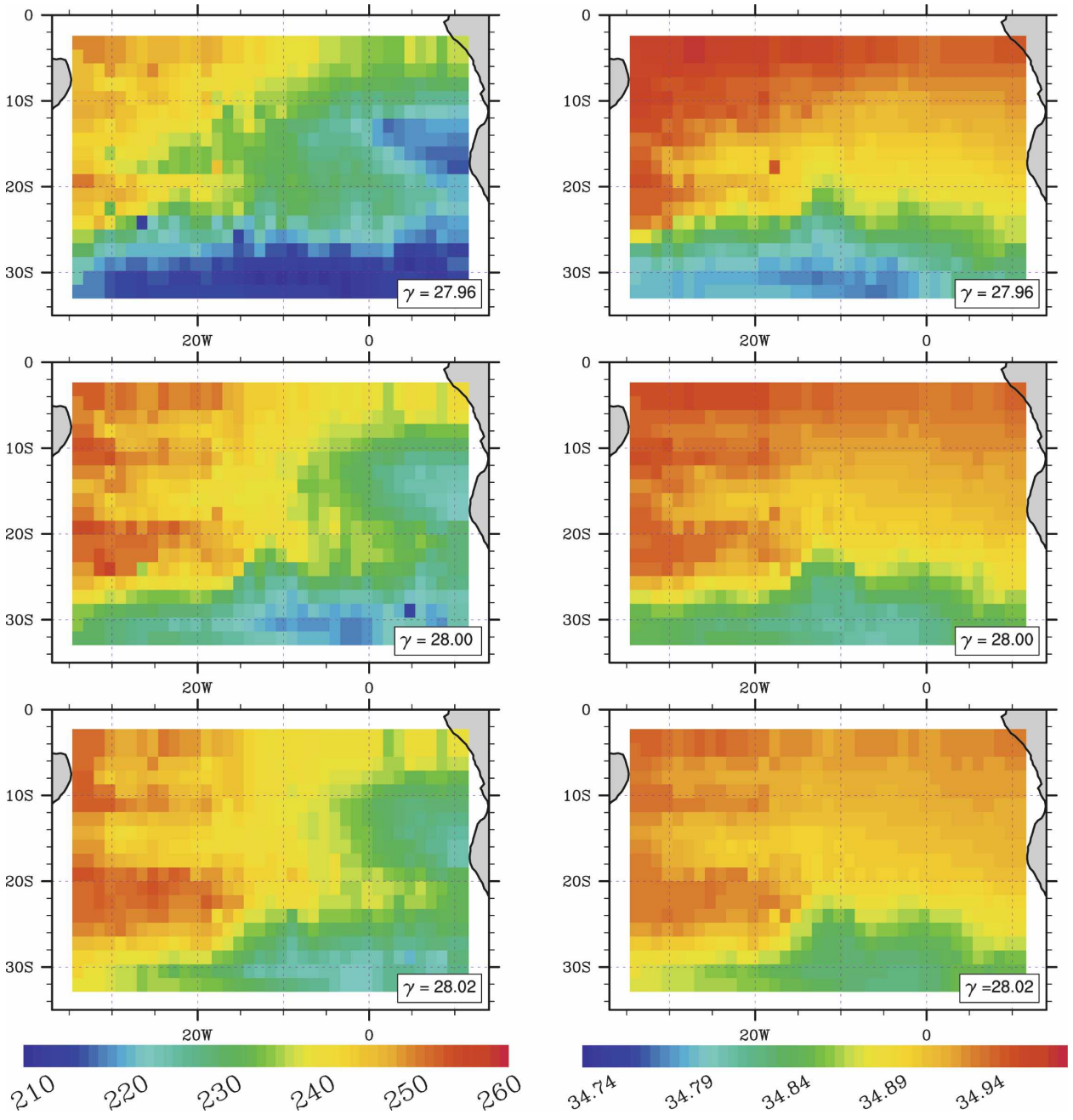

FIG. 2. (left) Interpolated oxygen and (right) salinity data for three neutral-density layers with one above (γ = 27.96) and one below (γ = 28.02) the layer of interest (γ = 28.00).

linear terms (cabbeling and thermobaricity) in Eq. (12) from McDougall (1991),

$$\left[w - \frac{\partial \kappa^{(z)}}{\partial z} \right] \frac{\partial \gamma}{\partial z} = \kappa^{(z)} \frac{\partial^2 \gamma}{\partial z^2}, \qquad (2)$$

where $\kappa^{(z)}$ denotes the vertical diffusivity. McDougall (1991) notes that cabbeling and thermobaricity are likely to be important in the Antarctic Circumpolar Current. Our domain does not include polar regions; thus, we neglect these terms and restrict the model to the mechanical mixing framework. Vertical variations of $\kappa^{(z)}$ have been hypothesized to be important within 500–1000 m of the ridge flank (Toole et al. 1994; Polzin et al. 1997). At middepth, however, vertical variations of $\kappa^{(z)}$ are thought to be negligible, so we set the term

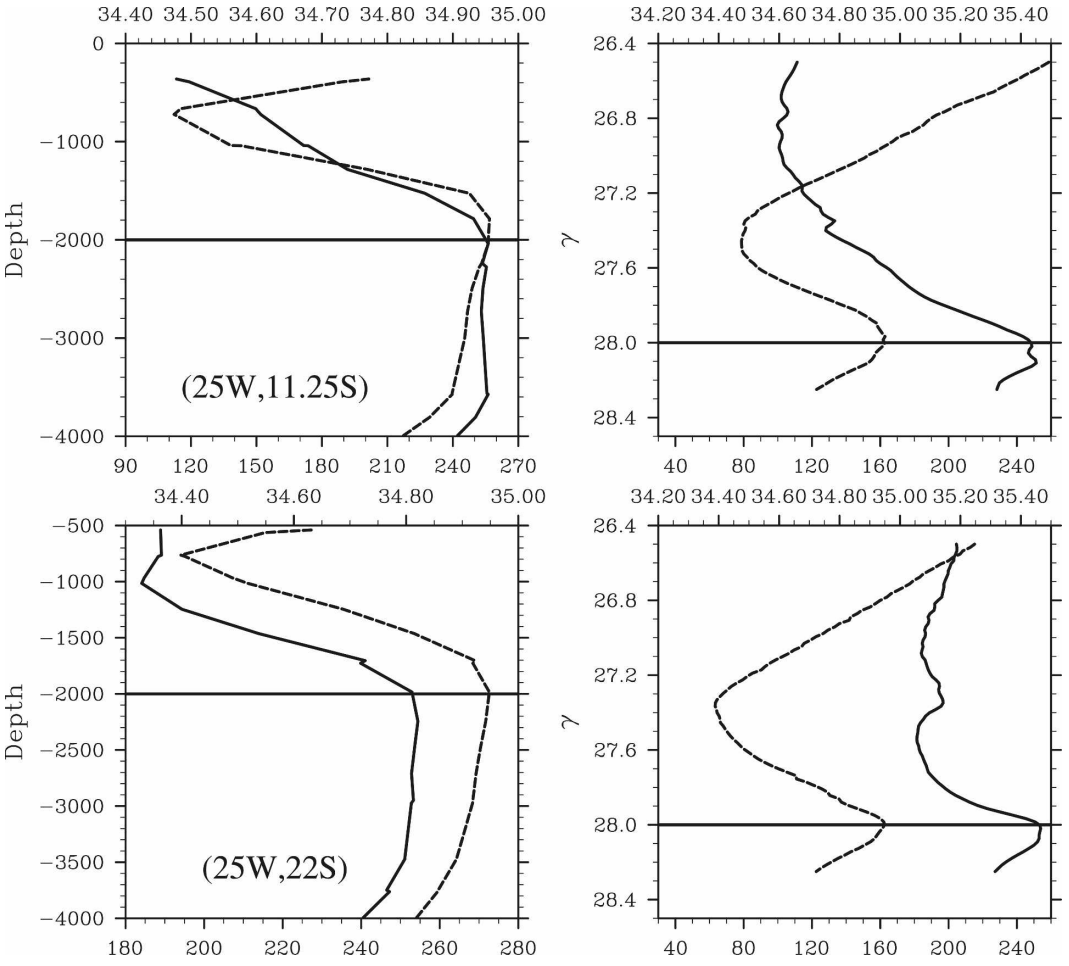

FIG. 3. Vertical profiles for oxygen (solid, bottom scale) and salinity (dashed, top scale) concentrations for two hydrographic stations, plotted against (left) depth and (right) neutral density.

$\partial\kappa^{(z)}/\partial z$ in (2) equal to zero. Note that although we include vertical structure, we ignore vertical variations in u, v, w, and $\kappa^{(z)}$ because our inversion has only one active layer and two boundary layers. Later, we relax this assumption and investigate a topographically driven $\kappa^{(z)}$ as well as a model with multiple active layers.

The neutral density γ is computed using the equations of state for seawater. Vertical gradients of γ are approximated using centered differences. Figure 4 presents the first and second vertical derivatives of the neutral density $\gamma_z \equiv \partial\gamma/\partial z$, $\gamma_{zz} \equiv \partial^2\gamma/\partial z^2$ at $z_0 = -2000$ m, and the ratio $r(x, y) = \gamma_{zz}/\gamma_z$, which measures the vertical variation of the abyssal neutral density field and is referred to as the *inverse density scale.* At least five layers (two above and two below the active layer) of data are needed to approximate these vertical gradients. Ignoring vertical variations in $\kappa^{(z)}$ and w, we can express the vertical velocity

$$w = w(x, y) = \kappa^{(z)}(x,y)r(x, y), \tag{3}$$

and thus, given $\kappa^{(z)}$, see that w is also determined. The inverse density scale $r(x, y)$ presented in the lower panel of Fig. 4 plays a very important role in the inversion. Because $\kappa^{(z)}$ is positive, w is forced to have the same sign as $r(x, y)$. Moreover, if the cross-isopycnal mixing field is fairly smooth, then the vertical velocity field will resemble $r(x, y)$. Although we assume that r is known, it is an inherently noisy quantity (being the ratio of two derivatives). To investigate whether perturbations in r significantly affect the results of the inversion, Eq. (3) could be extended to $w(x, y) = \kappa^{(z)}(x, y)[r(x, y) + \zeta(x, y)]$, where ζ is a zero-mean 2D Gaussian Markov random field as described in the next section. In the present paper, however, we assume that ζ is negligible.

The connection between velocities $\underline{\mathbf{u}}(x, y, z) = (u, v, w)$, diffusion coefficients $K = \text{diag}[\kappa^{(x)}, \kappa^{(y)}, \kappa^{(z)}]$, and

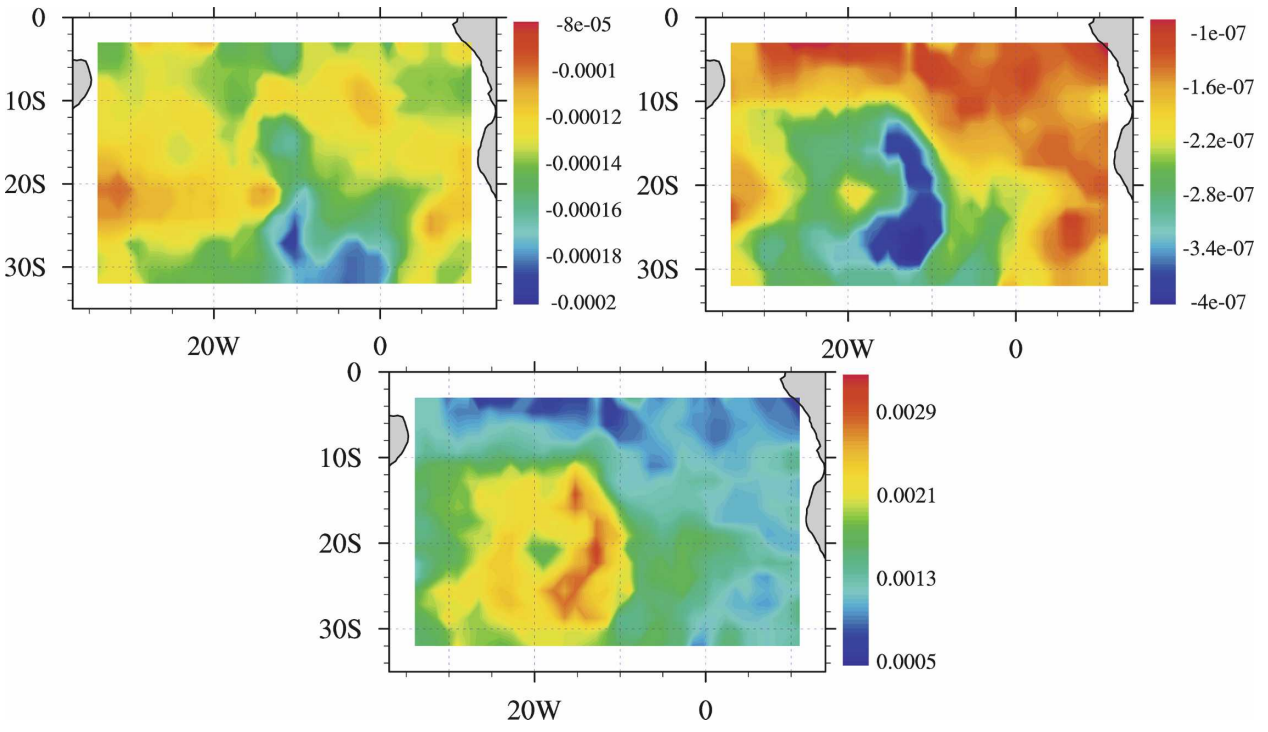

FIG. 4. Vertical gradients of the neutral density: (a) γ_z, (b) γ_{zz}, and (c) the inverse density scale $r(x, y)$ at $z_0 = -2000$ m.

tracer concentrations $C = C(x, y, z)$ is made via a steady-state 3D advection–diffusion equation

$$\underline{\mathbf{u}} \cdot \nabla C = \nabla \cdot (K \nabla C) + Q_C, (x, y, z) \in \mathcal{S}, \qquad (4)$$

with Dirichlet boundary conditions $C(x, y, z) = C_{\partial\mathcal{S}}$ for $(x, y, z) \in \partial\mathcal{S}$. The source term is $Q_C = -\lambda C$ for oxygen and $Q_C = 0$ for conservative tracers; we specify $\lambda = 10^{-10}\ s^{-1}$. We refer the reader to section 3c for an in-depth discussion regarding our choice for λ. The vertical diffusion coefficient $\kappa^{(z)}$ is equivalent to the one used in (2), since neutral density diffuses with the same coefficient as temperature and salinity. We aim in this study to arrive at a solution with no eddy advection terms involving gradients of $\kappa^{(x)}$ and $\kappa^{(y)}$. Thus, the horizontal diffusion coefficients are assumed to be spatially invariant. Variations in mixing can be investigated in our framework, but this extension is left for future work. Given a set of parameters $\Phi = (\underline{\mathbf{u}}, K, C_{\partial\mathcal{S}})$, the advection–diffusion equation may be solved numerically to reconstruct the tracer concentration (the forward problem).

b. Bayesian inversion

Beyond the physical model, we need the observational model to connect the observed tracer concentrations to the underlying steady-state tracer fields. We assume that the measured concentrations $C \equiv [C^{(j)}_{\text{obs,i}}]$ satisfy $C^{(j)}_{\text{obs,i}} \sim N(C^{(j)}_i, \sigma^2_{(j)})$, where $i = 1, \ldots, n_D$ indexes a site on the lattice $\mathcal{L}_G$ where data are available; $j = 1, \ldots, n_C$ indexes a tracer; and $C^{(j)}_i$ denotes the unknown, true, steady-state tracer concentration. The observations $C^{(j)}_{\text{obs,i}}$ are assumed to be independent random variables with constant (and prespecified) variance $\sigma^2_{(j)}$ for each tracer j (see the end of section 3d for further details). It is important to keep in mind that the underlying tracer concentration $C^{(j)}_i(x, y, z) = C^{(j)}_i(x, y, z|\Phi)$ is obtained as the solution of the advection–diffusion Eq. (4), given the parameters $\Phi = (\underline{\mathbf{u}}, K, C_{\partial\mathcal{S}})$. This solution is numerically approximated using a 3D multigrid iteration routine available in MUDPACK (Adams 1991). In the solver, we use a finer grid created by adding extra grid points between all neighboring points of $\mathcal{S}_G$ to improve the accuracy of the solution to the forward problem. The boundary values on the finer grid are determined by linearly interpolating their values on $\mathcal{S}_G$.

We pursue a Bayesian approach, choosing a prior model, which we specify by the unnormalized density function $\pi(\Phi)$, for the parameters to be estimated. Using the data model described above, the Bayes formula gives the probability density function (pdf) of Φ conditionally on the tracer data, up to a normalizing constant

$$\pi(\Phi|C) \propto L(C|\Phi)\pi(\Phi), \qquad (5)$$

where L denotes the likelihood function,

TABLE 1. Values of the tuning parameters δ_1 and δ_2 that appear in the prior distributions of υ, $\kappa^{(z)}$, and $C_{\partial S}^{(j)}$.

	Salinity	Oxygen	SiO_2	Potential vorticity	Potential temperature	υ	$\kappa^{(z)}$
δ_1	8.6×10^4	0.09	0.17	0.5×10^2	3.3×10^3	2×10^4	10^{10}
δ_2	10^{-6}	10^{-7}	10^{-7}	10^{-4}	10^{-4}	0	10^8

$$L(C|\Phi) \propto \prod_{i=1}^{n_D}\prod_{j=1}^{n_C} \exp\left\{ -\frac{[C_{\text{obs},i}^{(j)} - C_i^{(j)}]^2}{2\sigma_{(j)}^2} \right\}.$$

We specify υ, $\kappa^{(x)}$, $\kappa^{(y)}$, $\kappa^{(z)}$, and $C_{\partial S}$ to be independent under the prior distribution. The remaining components u and w are determined via Eqs. (1) and (3). We choose to model υ using a 2D Gaussian Markov random field (GMRF), having pdf

$$\pi(\upsilon) \propto \exp\left\{ -\delta_1 \sum_{s\sim s'} (\upsilon_s - \upsilon_{s'})^2 - \delta_2 \sum_{s\in \mathcal{L}_G} \upsilon_s^2 \right\},$$

where the first sum runs over all neighboring sites ($s \sim s'$ if s and s' are adjacent sites on the lattice $\mathcal{L}_G$). The positive tuning parameters δ_1 and δ_2 control the neighbor interaction and the variance in the prior distributions. We specify these parameters in Table 1. The values presented are selected based on inspection of several simulations from the prior distribution, leading to reasonable prior information. We find that random draws from this prior distribution display the features that we expect (see Fig. 15, upper left). As in the single-layer approach, exponential distributions (with mean 1000 $m^2\,s^{-1}$) are selected for the horizontal diffusion coefficients $\kappa^{(x)}$ and $\kappa^{(y)}$. As discussed before, we ignore vertical variations in $\kappa^{(z)}$ and restrict it to positive values; hence, the absolute value of a 2D GMRF is an appropriate choice as a prior distribution for the vertical diffusion coefficient. For each tracer $C^{(j)}$, the boundary conditions $C_{\partial S}^{(j)}$ are now represented by two 2D GMRFs (for the layer above and below the layer of interest) and a 1D GMRF (for the edge of the middle layer), thereby allowing for interaction between adjacent sites on the lateral surfaces of the 3D inversion domain.

EXTRACTING INFORMATION FROM THE POSTERIOR DISTRIBUTION

The posterior distribution

$$\pi(\Phi|C) \propto \left\langle \prod_{i=1}^{n_D}\prod_{j=1}^{n_C} \exp\left\{ -\frac{[C_{\text{obs},i}^{(j)} - C_i^{(j)}]^2}{2\sigma_{(j)}^2} \right\} \right\rangle \times \pi(\upsilon)\pi[\kappa^{(x)}]\pi[\kappa^{(y)}]\pi[\kappa^{(z)}]\prod_{j=1}^{n_C}\pi[C_{\partial S}^{(j)}]$$

may be thought of as "the solution to the inverse problem." The output of our approach is an entire collection of possible solutions, onto which we have built a probability model allowing us to quantify whether one particular solution is more probable than another. We summarize this distribution by estimating its mean and standard deviation. Formally describing the high-dimension posterior distribution $\pi(\Phi|C)$ is not possible because of the intractable normalizing constant. Using MCMC techniques, we are able to explore the posterior state space in a manner that allows us to draw samples from it. This exploration is performed by moving from one state (or possible solution) to another such that the amount of time the Markov chain spends in a particular region of the state space (quantified as the number of states from that region visited by the chain) is proportional to the probability mass the posterior distribution puts on that region. Having drawn samples from the posterior distribution, the mean and standard deviations are then estimated using the sample statistics. For more background on MCMC we refer the reader to Liu (2001). Following discussions from the single-layer inversion method, the posterior state space needs to be explored adequately in order to assess the posterior uncertainty accurately. Standard random walk Metropolis moves and column moves as described in McKeague et al. (2005) are used. In addition, we introduce new moves (scaling moves) to improve the mixing of the chain and achieve fast decorrelation in the posterior sample. These moves are described and discussed in the appendix. A complete iteration in the MCMC scheme we use consists of 1000 cycles; each cycle consists of (i) a scaling move; (ii) a random walk move on $\kappa^{(x)}$, $\kappa^{(y)}$ (one chosen at random); (iii) a random walk move on one randomly selected component of υ; (iv) a random walk move on one randomly selected component of $\kappa^{(z)}$; (v) a random walk move on one randomly selected component of $C_{\partial S}$; and (vi) a column move.

3. Inversion results

a. Flow estimates

The posterior mean horizontal flow shows predominant zonal flow with alternating jets associated with tracer tongues, narrower in the vicinity to the equator and broader farther south (Fig. 5). Quasi-3D inversion

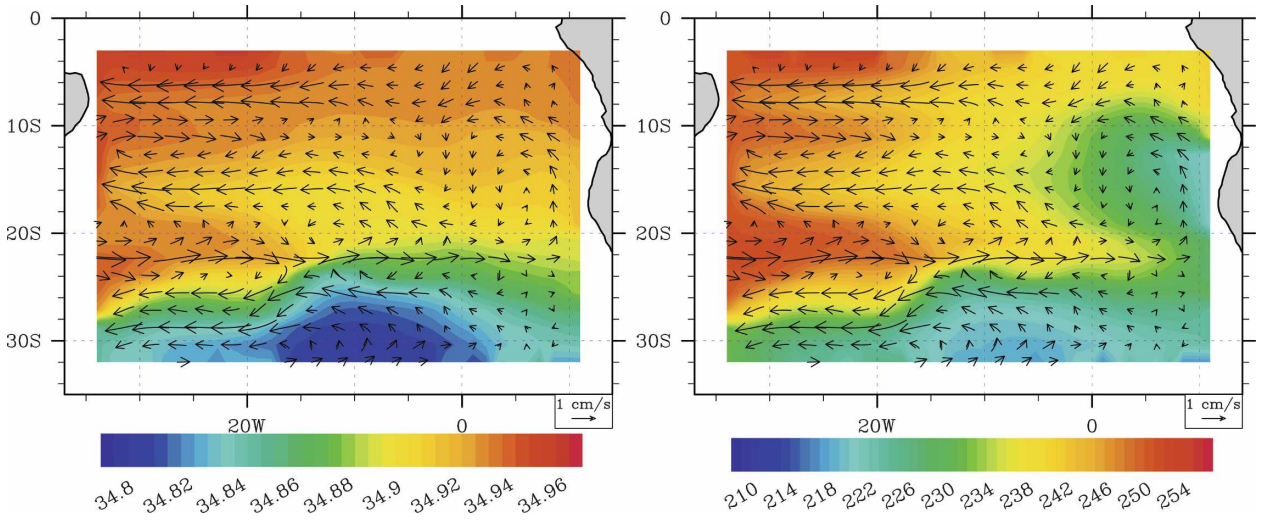

FIG. 5. Posterior mean flow from the quasi-3D inversion superimposed on the (a) posterior mean salinity and (b) oxygen concentration fields, on $\gamma = 28.00$.

velocities significantly different from zero by at least one and two posterior standard deviations are displayed in the left panel of Fig. 6. In comparison to the solution of the single-layer inversion method (Fig. 6, right panel), we observe more significant zonal flow and less meridional flow with only a few significant values along the east side of the Mid-Atlantic Ridge in the southern part of the domain. There is evidence of alternating zonal jets that extend across the entire domain. To get an adequate comparison of the widths of the alternating zonal jets to the Rossby radius, we would need to obtain sufficient dynamic range, which would require a larger inversion region.

Our solution includes a large-scale potential vorticity (PV = f/H) estimate (Fig. 7) because PV is one of the tracers. This estimate enables gradients of PV to be calculated consistently with the other tracers and with the velocity field (this does not mean that the relative vorticity is important). We approximate the meridional gradient of PV and display it in the right panel of Fig. 7, with the mean flow field superimposed. We note the zonally banded structure of $\partial(\mathrm{PV})/\partial y$ and observe that westward zonal jets are generally associated with high positive meridional gradients. The meridional scale of the jets is about 2°–4°, similar to the findings of Treguier et al. (2003).

In regions with low vertical diffusion (i.e., southwest or northeast of the inversion domain; see section 3b for details) we observe flow along the isopleths of the (conservative) tracers, a characteristic of a purely advective solution. This is no longer valid for the region above the Mid-Atlantic Ridge, where more intense cross-

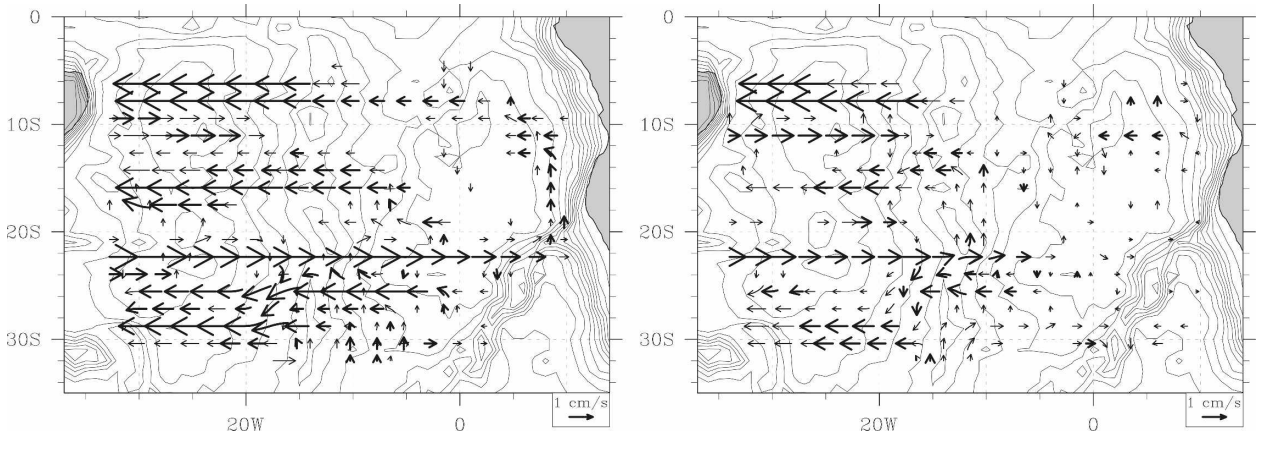

FIG. 6. Posterior mean velocities that differ from zero by at least one (light) and two (dark) posterior std devs for the (a) quasi-3D approach and (b) single-layer method; contours of the bottom topography (500-m interval).

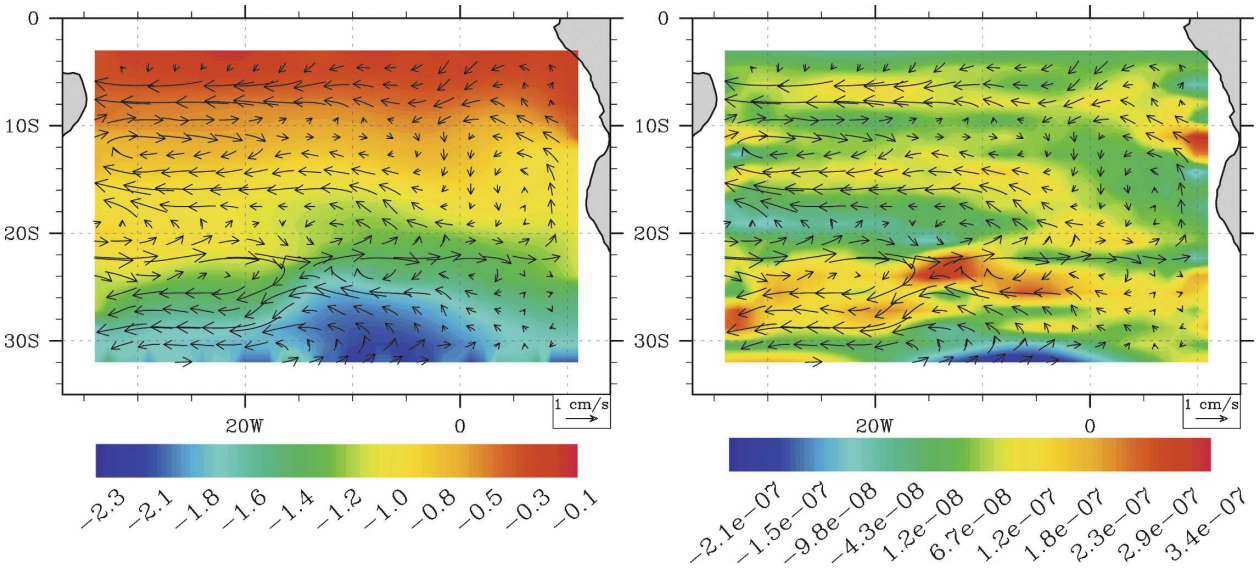

FIG. 7. (a) Posterior mean flow superimposed on the posterior mean PV (normalized scale) and (b) the PV_y.

isopycnal mixing is estimated. For a better understanding of the main balances in the advection–diffusion equation, we quantify the magnitude of the advective and diffusive terms along each axis in Eq. (4) using a spatial root-mean-square (RMS). For example, the RMS of the advective term along the x axis is

$$\mathrm{RMS} = \sqrt{\frac{\sum_{s\in L_G} [u(s)C_x(s)]^2}{19\times 37}}. \qquad (6)$$

In Fig. 8 we detail this analysis for oxygen, showing the RMS for each term as a function of iteration in an MCMC simulation. We find that our solution is primarily advective with vertical components that have an overall small influence, presumably due to the coarse vertical resolution in our model. The vertical terms $\kappa^{(z)}O_{zz}$ and wO_z are one order of magnitude smaller than the horizontal ones and exhibit variation that is too small to be observed on the scale of the plot.

b. Vertical structure

Our approach provides estimates and uncertainties of the vertical diffusion coefficient $\kappa^{(z)}$ and velocity w (Fig. 9). We observe upwelling throughout the domain.

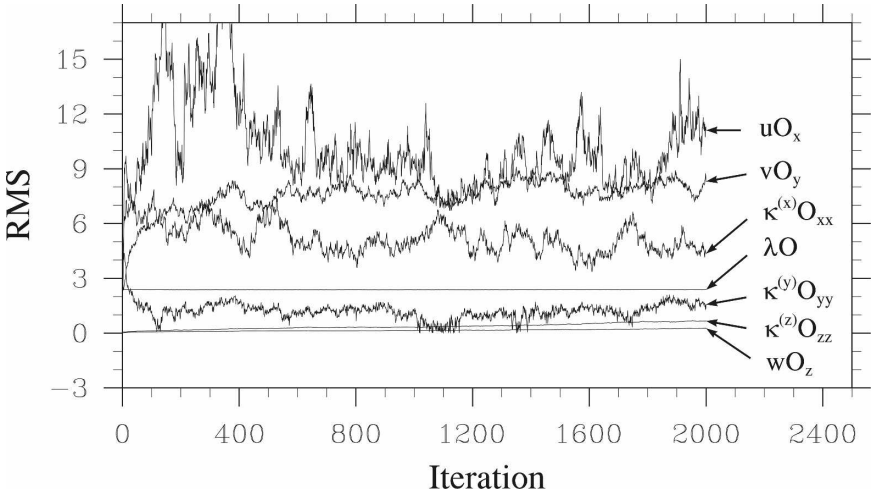

FIG. 8. Advective–diffusive balances: trace plots of RMS for each term in the oxygen advection–diffusion equation (10^{-8}). The approximately horizontal lines correspond to the vertical diffusion [$\kappa^{(z)}\,O_{zz}$], vertical advection (wO_z), and the "sink" term (λO) whose variations are too small for the scale of this plot.

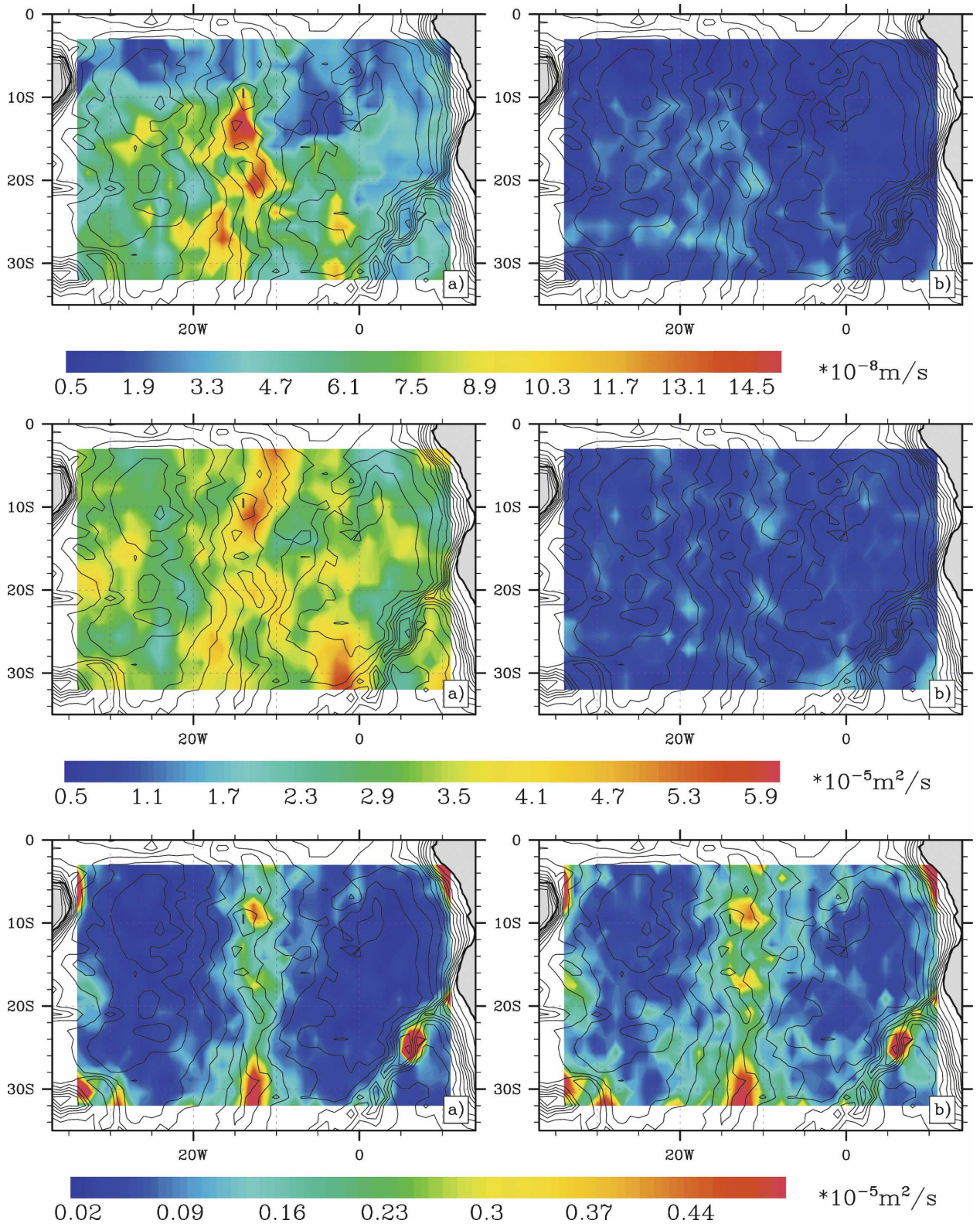

FIG. 9. (a) Posterior mean vertical velocity w and (b) std dev; (c) posterior mean vertical mixing coefficient $\kappa^{(z)}$ and (d) posterior std dev; (e) posterior mean vertical mixing coefficient $\kappa^{(z)}$ and (f) posterior std dev obtained using the model (7); contours of the bottom topography (500-m interval).

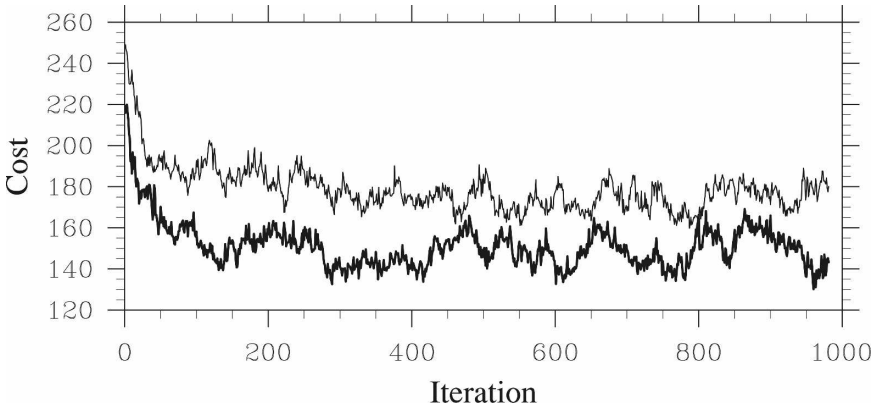

FIG. 10. Oxygen cost comparison for the two vertical mixing models: 2D GMRF (thick line) and topography driven (thin line). Trace plots of Err(Φ).

This is mainly a result of the simplified model and the constraints imposed from Eq. (3) [note that $\kappa^{(z)}$ and w must have the same sign because the inverse density scale $r(x, y)$ is positive throughout the active layer].

The center panels of Fig. 9 present the posterior mean (left) and standard deviations (right) for the vertical mixing coefficient. Increased vertical mixing is usually associated with higher uncertainty in the posterior distribution. Although the average field is fairly smooth across the entire domain, it is interesting, even though we do not incorporate topography into our model, to notice an association between $\kappa^{(z)}$ and the bottom topography, with higher values over the midocean ridge.

We further explore the relationship between vertical mixing and bottom topography by imposing the following idealistic model:

$$\kappa^{(z)}(x, y) = L(x, y) \exp\left\{-\frac{z(x, y) - h(x, y)}{\alpha}\right\}, \quad (7)$$

where $z(x, y)$ denotes the depth of the layer of interest, $h(x, y)$ is the depth of the ocean, α denotes an exponential decay scale, and $L(x, y)$ is an arbitrary (smooth) level function. We treat h and z as fixed (nonrandom) functions of location and L and α as random positive quantities. Given an absence of prior knowledge about the level function, a 2D GMRF is suitable as a prior distribution for L. For the decay scale, we use an exponential prior with mean 500 m. Huang and Jin (2002) specify the vertical diffusivity similarly, using a quadratic exponential decay to study deep circulation induced by bottom intensified mixing. At middepths shallower than the tip of the midocean ridge, the appropriateness of such a model is not clear.

Our results from this approach (Fig. 9, lower panels) show higher posterior variability and overall reduced vertical diffusivity values, becoming insignificant in various places in the eastern and western basin. The posterior mean exponential decay scale of 800 m has very high uncertainty (posterior standard deviation ~500 m), suggesting that the inversion is trying to introduce more variability than model (7) can accommodate.

Does the use of model (7) produce a better fit to the data than the model described in section 2b? To make such a comparison, we define a cost function as follows: for each visited state Φ, the distance between available tracer observations $C^{(j)}_{\text{obs},i}$ and the tracer concentration $C(\Phi)$ determined via the advection–diffusion equations is quantified using a cost function (or equivalently, minus the log-likelihood), which is defined as the error sum of squares normalized by twice the error variance and takes the form

$$\text{Err}(\Phi) = \sum_{i=1}^{n_D} \sum_{j=1}^{n_C} \frac{[C^{(j)}_{\text{obs},i} - C^{(j)}_i(\Phi)]^2}{2\sigma^2_{(j)}}.$$

We use this cost function to compare the fit to the oxygen data produced by our model with topographic mixing (7) and our earlier model [GMRF directly on $\kappa^{(z)}$; see Fig. 10]. Note that there is a higher error associated with topographic mixing, which we attribute to a less realistic horizontal velocity field. This is due to a strong control exerted over $\partial w/\partial z$, hence over v and u as well.

c. Oxygen consumption rate and uniqueness

The consumption rate λ appearing in the oxygen equation plays an important role in our setting. Although it varies near regions with high productivity, λ is thought to be nearly constant over the broad interior of the ocean (Feely et al. 2004). For simplicity we set λ to be a constant; this provides a fixed clock rate in the

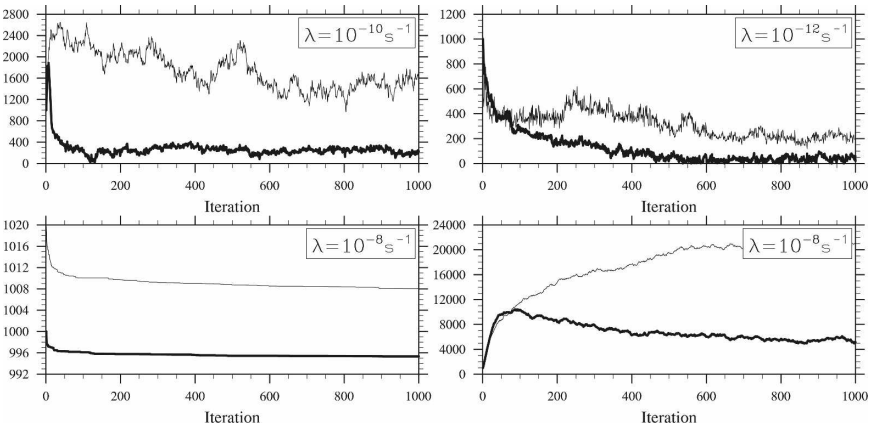

FIG. 11. Trace plots for $\kappa^{(x)}$ (thin line) and $\kappa^{(y)}$ (thick line) in $m^2\ s^{-1}$. Simulation runs for different values of λ, showing the behavior of the chain (a): $\lambda = 10^{-10}\ s^{-1}$, (b) $\lambda = 10^{-12}\ s^{-1}$, (bottom) $\lambda = 10^{-8}\ s^{-1}$ with (c) simulation using all tracers, and (d) with the simulation using oxygen only.

advection–diffusion equation and avoids the ill-posedness of the inverse problem. The magnitude of λ can be estimated from a simple analysis of the terms in the advection–diffusion equation for oxygen,

$$\underline{\mathbf{u}} \cdot \nabla \mathrm{O} = K\nabla^2 \mathrm{O} - \lambda \mathrm{O}.$$

A typical deep ocean is characterized in the interior (outside the western boundary current) by length scales of order $L \sim 10^6$ m, velocities of about $10^{-2}\ m\ s^{-1}$, and diffusion coefficients of $10^3\ m^2\ s^{-1}$. Thus, the advection term in the equation above is estimated to be of the order $10^{-2}\ m\ s^{-1} \times 10^{-5}\ \mu mol\ kg^{-1}\ m^{-1} = 10^{-7} \mu mol\ kg^{-1}\ s^{-1}$. Similarly, the diffusion term is roughly $10^3\ m^2\ s^{-1} \times 10^{-11}\ \mu mol\ kg^{-1}\ m^{-2} = 10^{-8}\ \mu mol\ kg^{-1}\ s^{-1}$. This leads to $\lambda O \sim 10^{-7} \mu mol\ kg^{-1}\ s^{-1}$, or equivalently, $\lambda \sim 5 \times 10^{-10}\ s^{-1}$. A similar value can be found by a different scaling argument. The time scale of our system is of the order 10^8 s (3 yr); thus, changes in the oxygen concentration are $\Delta O \sim \lambda \times 10^8 \times O$, hence $\lambda \sim 5 \times 10^{-10}\ s^{-1}$. Also, Feely et al. (2004) report an oxygen utilization rate (λO) of $0.1 \times 10^{-9}\ \mu mol\ kg^{-1}\ yr^{-1}$ which leads to $\lambda \sim 2 \times 10^{-11}\ s^{-1}$. Thus, we settled on the average value for $\lambda = 10^{-10}\ s^{-1}$.

Higher values for λ would increase the "sink" term in the oxygen equation, requiring greater diffusion near the equator where advection is small. Indeed, this is shown by trace plots for $\kappa^{(x)}$ and $\kappa^{(y)}$ in a simulation run using only oxygen and $\lambda = 10^{-8}\ s^{-1}$ (Fig. 11, lower right panel). The lateral diffusion coefficients are of order $10^5\ m^2\ s^{-1}$. This would throw off the balances in the advection–diffusion equation for the other tracers (the diffusion term dominates), which is physically unrealistic. It would also cause the Markov chain to get stuck. To illustrate this, for the same value of λ, a simulation run using all the tracers was performed. Examining the trace plots for the lateral diffusion coefficients (Fig. 11, lower left panel) we find that these parameters do not "move" (nearly flat lines) for many iterations.

Lower values of λ are not acceptable either. Reducing the oxygen consumption rate by two orders of magnitude, to $10^{-12}\ s^{-1}$, lowers the velocities and diffusion coefficients by the same factor (in order to maintain balance). In particular $\kappa^{(y)}$ becomes almost negligible (about $10\ m^2\ s^{-1}$; see Fig. 11, upper right panel), even though the Markov chain is not getting stuck in this case. The choice of $\lambda = 10^{-10}\ s^{-1}$ appears to be satisfactory (Fig. 11, upper left panel) in that it produces plausible diffusion coefficients and adequate convergence of the Markov chain.

We further examine the behavior of the chain by characterizing each accepted solution by a pair of nondimensional quantities ($\tilde{\lambda}$, Pe), where $\tilde{\lambda} = \lambda L/U$. The velocity component U is estimated using the spatial RMS described above and varies as the chain is exploring the posterior distribution. The Peclet number is computed as $\mathrm{Pe} = UL/\kappa^{(x)}$. The length scale L and the clock rate λ are held constants: $L = 10^6$ m, $\lambda = 10^{-10}\ s^{-1}$. In Fig. 12 we display visited states in the (Pe, $\tilde{\lambda}$) space. We notice that the Markov chain does converge (the gray dots tend to stay in a relatively small region of this space). Once the chain has reached stationarity, the possible solutions are characterized by Pe ~ 6. The same value for the Peclet number can also be achieved if diffusion coefficients and velocities are simultaneously very large or very small, which could happen under the prior model, yet in that case $\tilde{\lambda}$ would be very small or very large. The sink term in the oxygen advection–diffusion equation is thus selecting the solution.

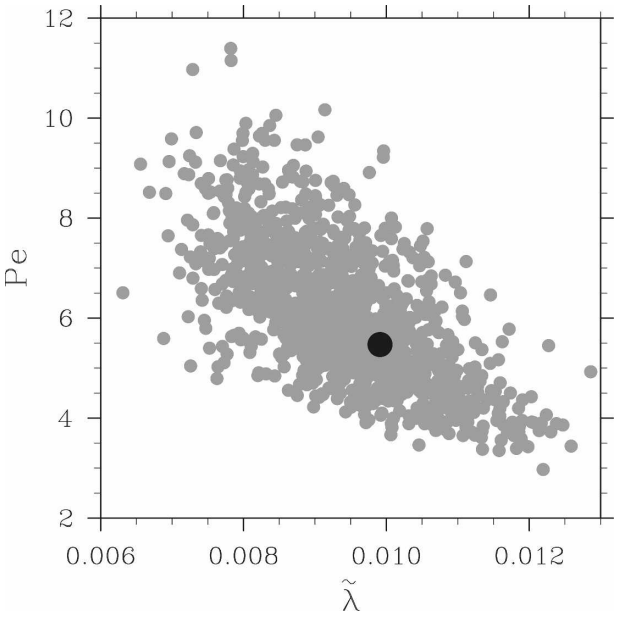

FIG. 12. Probable solution distribution in the ($\tilde{\lambda}$, Pe) space. We display 1000 visited solutions (gray dots) after the Markov chain has reached stationarity. The black dot represents the posterior mean solution.

Tonguelike distributions of tracer concentration data are usually interpreted as advective effects. As Armi and Haidvogel (1982) point out, for a steady-state model, variable and anisotropic diffusion coefficients can generate similar distributions. Floats deployed in the Brazil Basin (Hogg and Owens 1999) have shown a flow dominated by zonal motions. Our results show a significant anisotropy of the diffusion coefficients (see Fig. 13), but we find that the tracer concentration tongues are dominantly advective (see, e.g., Fig. 8). We estimate a mean meridional component of roughly 220 $m^2\ s^{-1}$, consistent with low eddy kinetic energy levels, of the order 1 $cm^2\ s^{-2}$ with a time scale of 20–30 days. Based on the balance between zonal advection and meridional eddy diffusion and on estimates of tracer gradients, Thurnherr and Speer (2004) come to a similar value for $\kappa^{(y)}$. Our estimates are also consistent with the large-scale eddy diffusivities $[\kappa^{(x)}, \kappa^{(y)}] = (1.5 \pm 0.7, 0.7 \pm 0.4) \times 10^3\ m^2\ s^{-1}$ estimated using float and tracer observations from the North Atlantic Tracer Release Experiment (Sundermeyer and Price 1998). Simulations done with imposed isotropic diffusivities have shown significantly higher cost functions, as displayed in Fig. 14.

d. Accuracy of the inversion: Simulation example

In this subsection we examine the question of whether our approach provides an accurate solution in the setting of a synthetic inversion problem. We create a simple example with simulated data and known target parameters to assess the performance of the inversion method proposed in sections 2a and 2b.

We prespecify a "realistic" meridional velocity field v and vertical diffusion coefficient field $\kappa^{(z)}$, representing the main targets of the inversion (see the first panels of Figs. 15 and 17, respectively). Each field is a random draw from its prior distribution described in section 2b. The horizontal diffusivities are assumed to be $\kappa^{(x)} = 1000\ m^2\ s^{-1}$ and $\kappa^{(y)} = 500\ m^2\ s^{-1}$. For simplicity we only use a single tracer (oxygen) with constant boundary conditions on the upper and lower layers, and linear interpolation to the boundary of the middle layer. To generate the oxygen tracer data we numerically solved the forward problem (4) yielding the tracer concentration over the entire grid $\mathcal{S}_G$, discard all but the values at observation sites, and add independent N(0, 1) measurement error. To compensate for only using a single tracer we have reduced the measurement error variance from 2.46 to 1.0. We use the same priors as before.

The results are displayed in Figs. 15–17. The posterior mean flow field (Fig. 15, top right panel) is in close agreement with its target (top left panel) in that the main features are reproduced (e.g., strong zonal flow in the northwest part of the domain). Note, however, that the posterior mean flow is considerably smoother than its target, which is to be expected because it does not reflect any variability. To get a sense of the variability in the inversion, it is helpful to inspect draws of the posterior flow field (Fig. 15, bottom panels); these also exhibit the same features as the target. Posterior histograms of the horizontal diffusion coefficients (Fig. 16)

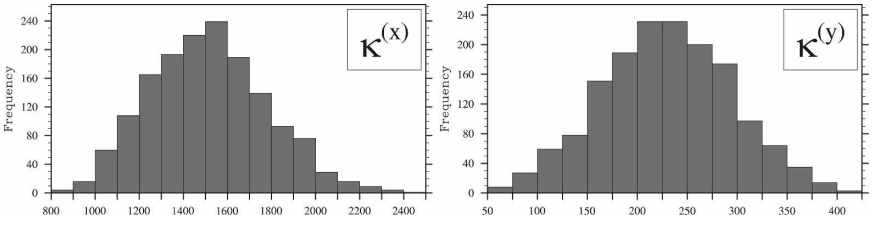

FIG. 13. Posterior histograms for the lateral diffusivities (left) $\kappa^{(x)}$ and (right) $\kappa^{(y)}$ ($m^2\ s^{-1}$).

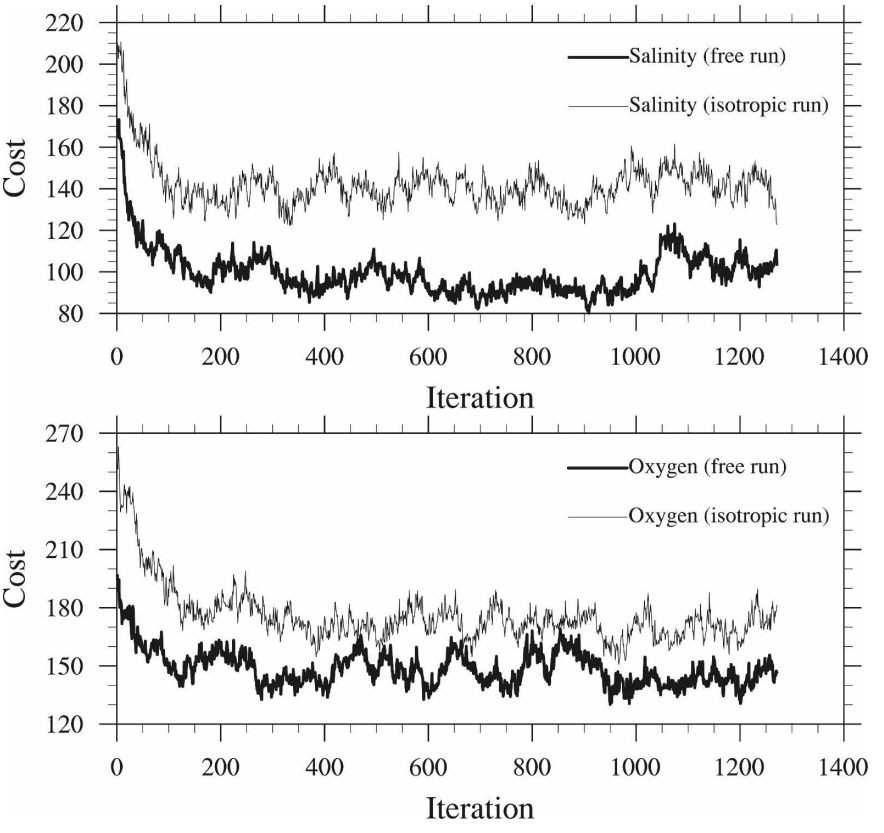

FIG. 14. Comparison between the cost functions for (top) salinity and (bottom) oxygen concentrations in two simulation runs: imposed isotropic lateral diffusion coefficients (thin lines) and a free run (thick lines).

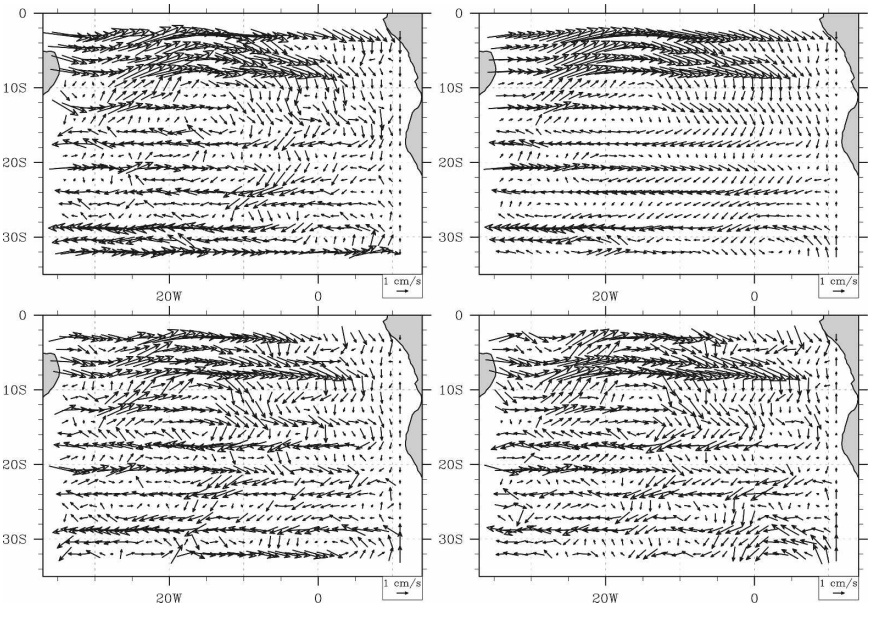

FIG. 15. Simulation example: (a) assumed true flow field, (b) posterior mean flow field, and (c) and (d) two random draws from the posterior distribution.

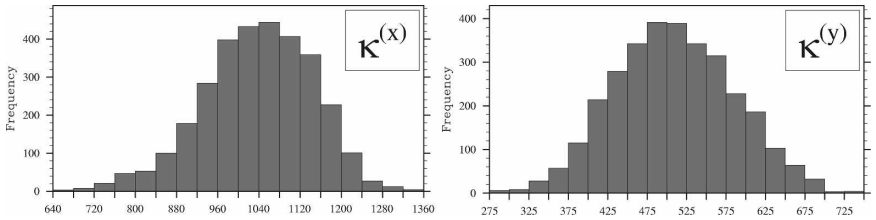

FIG. 16. Simulation example: posterior histograms for (a) $\kappa^{(x)}$ and (b) $\kappa^{(y)}$.

are quite tightly concentrated around the target values, with posterior means of 1031 $m^2\,s^{-1}$ for $\kappa^{(x)}$ and 508 $m^2\,s^{-1}$ for $\kappa^{(y)}$. The posterior standard deviations are 95 and 69 $m^2\,s^{-1}$ respectively. These posterior distributions are approximately Gaussian and are significantly different from their exponential priors, showing that the data have had a large influence. The posterior mean of the vertical diffusivity $\kappa^{(z)}$ (Fig. 17, middle panel) again reproduces the main features of its target (first panel), at least up to a degree of variability given by the posterior standard deviation (last panel).

Two other important factors are influencing the posterior distribution: the error variance and the number of data sites. There are 239 data locations (in both the real and artificial problems) on a 37 × 19 lattice. Our experiments have shown that we can reduce the number of data points by as much as 10% without significantly altering the resulting posterior mean flow and mixing values. It would require separate (time consuming) simulation runs to determine the influence of each data point individually. Both the error variance and the number of data locations are controlling the "power of the data," by either strengthening or relaxing the likelihood. Correspondingly, varying these parameters will either reduce or enlarge the posterior variance, assuming the prior does not change. In the real problem, the number of data locations is determined by the grid size and the original hydrographic stations where data were collected. The error variance $\sigma^2_{(j)}$ for each tracer is estimated by computing a local sample variance from tracer measurements in a small box around an observation site, and then taking a weighted average of these sample variances, with the weights being reciprocals of the local sample sizes. This is the procedure used in McKeague et al. (2005), resulting in the values given in Table 1 of that paper.

4. Thermal wind and a multilayer inversion

A natural question is what happens when our model is applied at other depths? We can easily perform the quasi-3D statistical inversion at any depth we care to specify. In Figs. 18 and 19 the upper panels present the posterior mean flow and tracer concentrations for the $\gamma = 27.63$ and $\gamma = 27.15$ isopycnals. One of the deficiencies of our model is that these estimated velocities and tracer concentrations are not related in any way to each other or to the $\gamma = 28.00$ layer. A natural way to link these layers vertically would be to use thermal wind relations:

$$\frac{\partial u}{\partial z} = \frac{g}{f\rho_0}\frac{\partial \rho}{\partial y}, \quad \frac{\partial v}{\partial z} = -\frac{g}{f\rho_0}\frac{\partial \rho}{\partial x}.$$

We note that since tracers are advected by the total velocity, our approach will estimate the total field, barotropic plus baroclinic. Using velocity estimates from the $\gamma = 28.00$ layer, one can integrate thermal wind equations vertically and obtain velocities across

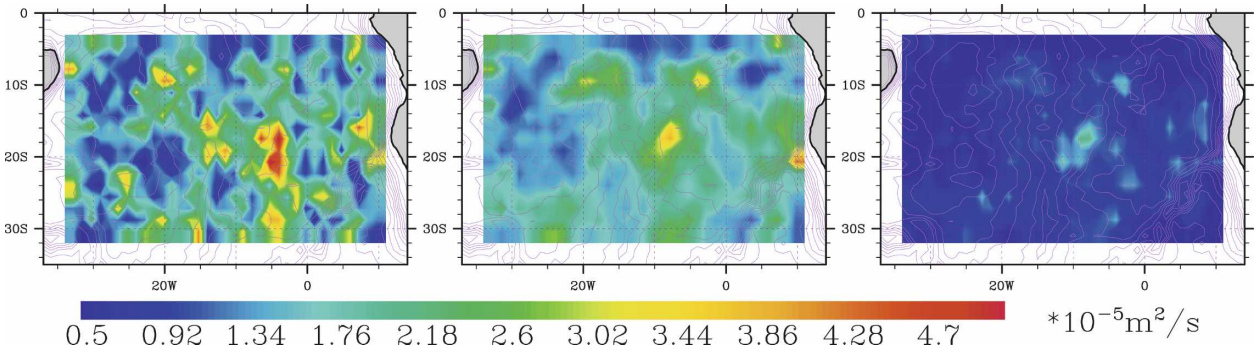

FIG. 17. Simulation example: (a) assumed vertical diffusion, (b) posterior mean, and (c) std dev fields for the vertical mixing coefficient $\kappa^{(z)}$.

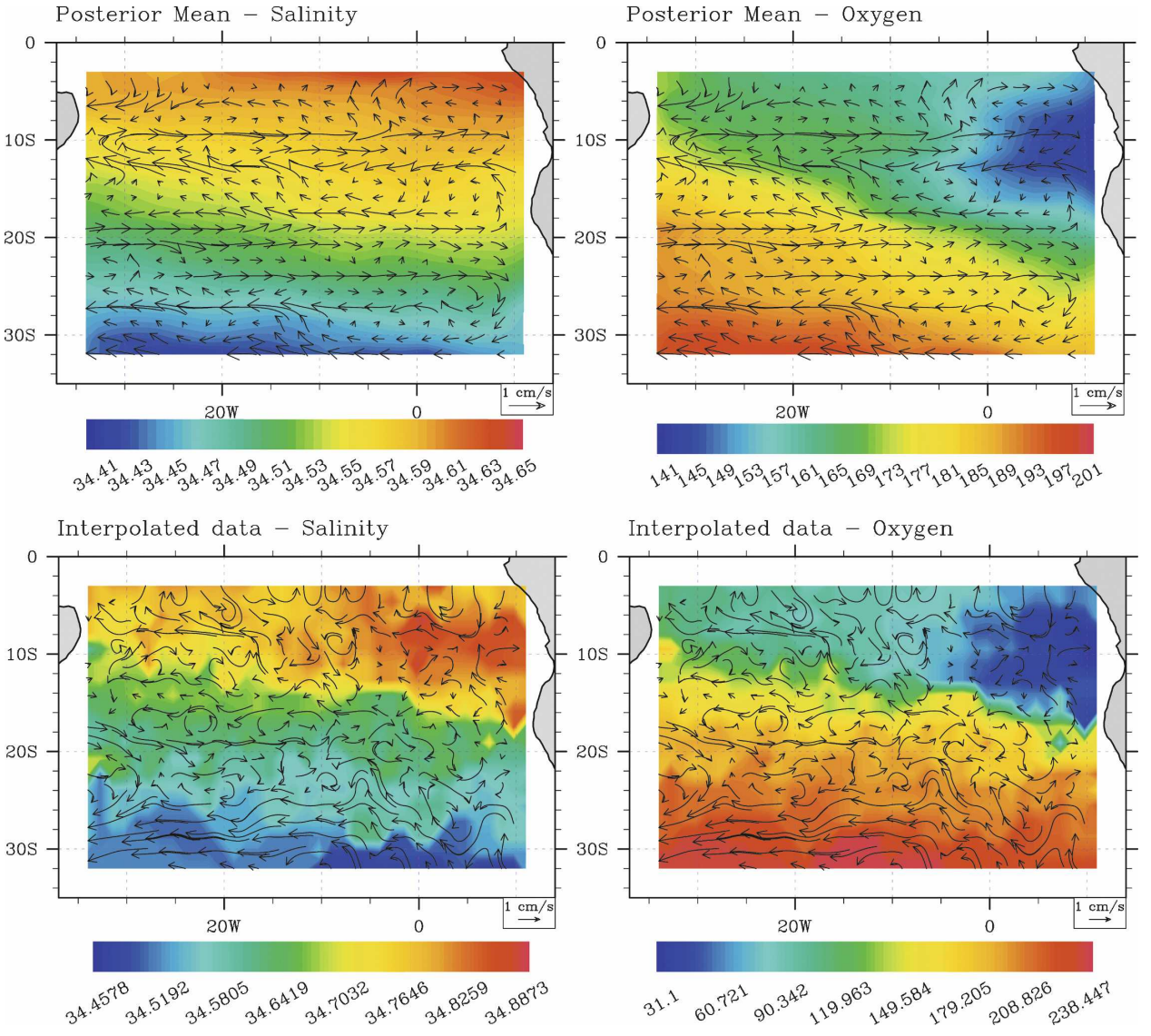

FIG. 18. Posterior mean velocities and reconstructed tracer concentrations of (a) salinity and (b) oxygen for the $\gamma = 27.63$ neutral-density layer (~1000-m depth) using a quasi-3D model. (c) Salinity, and (d) oxygen interpolated tracer concentration data and velocities obtained by vertically integrating thermal wind relations for the same isopycnal.

the entire water column. The lower panels in Figs. 18 and 19 present velocities obtained in this fashion on top of interpolated tracer concentration data. Neither of these approaches produces satisfactory results. The quasi-3D methodology lacks a vertical link between layers. Flow derived by integrating thermal wind relations is turbulent, includes eddies, and may not be consistent with observed tracer concentrations. We choose to combine the two methods in one unified model that uses thermal wind to connect layers and the advection–diffusion equation to link tracer concentration data to velocities and diffusion coefficients.

The inversion domain $\mathcal{S}$ in this situation will have the same lateral boundaries as in the quasi-3D setup but the vertical boundaries are now set to $\gamma = 28.00$ and $\gamma = 27.15$. We discretize $\mathcal{S}$ using seven 37×19 grids $\mathcal{L}_G$ stacked on top of one another, with each layer representing an isopycnal ($\gamma = 28.00$, 27.96, 27.89, 27.78, 27.63, 27.45, and 27.15, roughly corresponding to depths of 2000 to 500 m in steps of 250 m). Using hydrographic data, we compute the in situ density ρ and approximate its horizontal gradients $\partial\rho/\partial x$, $\partial\rho/\partial y$. With $k = 1, \ldots, 7$ indexing the layers, let

$$G^{(k)}_{u,\mathrm{obs}} = \frac{g}{f\rho_0}\frac{\partial \rho^{(k)}}{\partial y}, \quad G^{(k)}_{v,\mathrm{obs}} = -\frac{g}{f\rho_0}\frac{\partial \rho^{(k)}}{\partial x}.$$

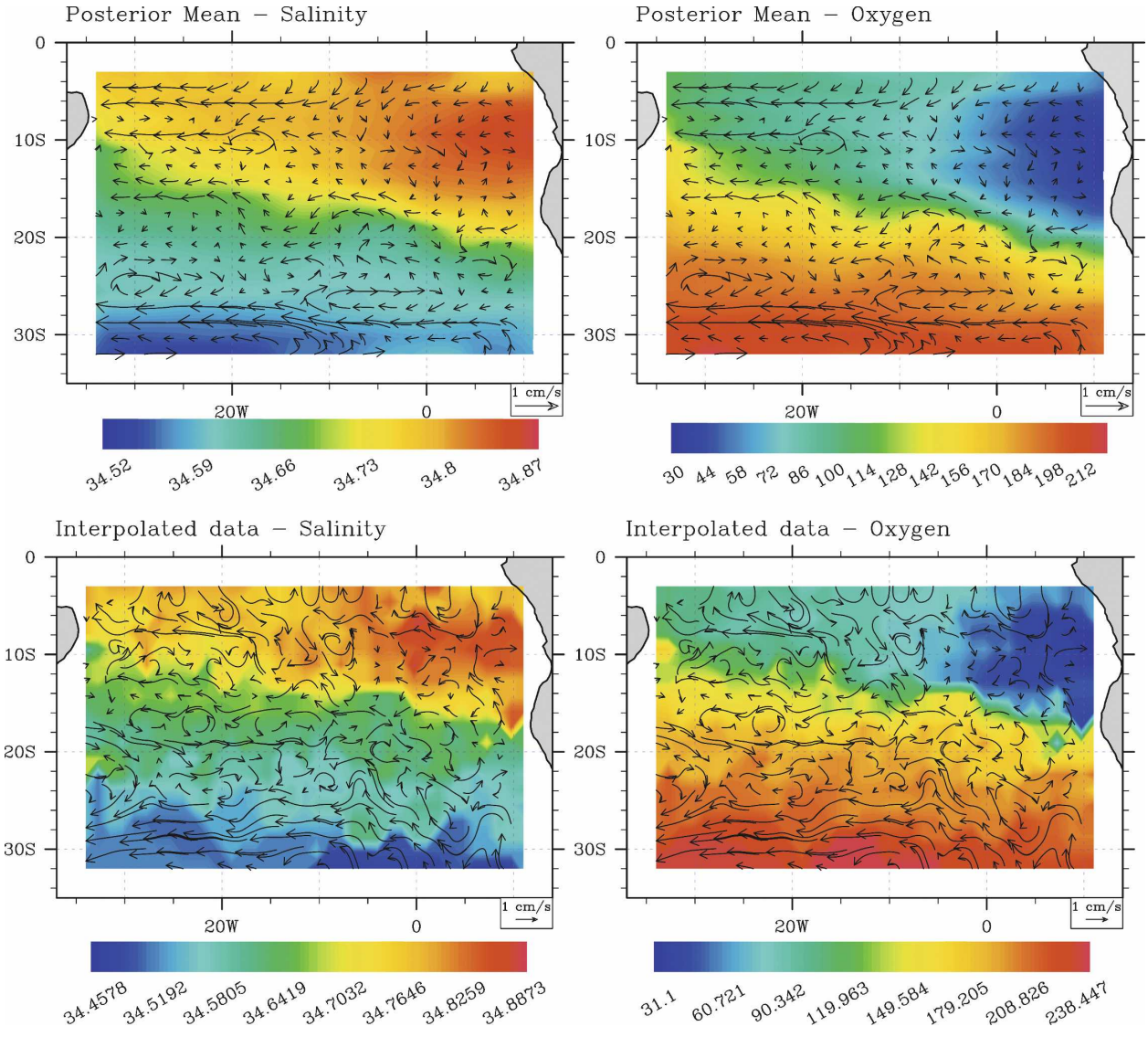

FIG. 19. Same as Fig. 18 but for the $\gamma = 27.15$ neutral-density layer (~500-m depth).

We view $G \equiv [G^{(k)}_{u,\mathrm{obs}}, G^{(k)}_{v,\mathrm{obs}}]$ as additional data, supplementing observed tracer concentrations. Together with the observational model described before, we also assume that for each layer k and each grid site $s \in \mathcal{S}_G$, $G^{(k)}_{u,\mathrm{obs}}$, (s) and $G^{(k)}_{v,\mathrm{obs}}(s)$ are independent Gaussian variables having means $(\partial u/\partial z)^{(k)}$ (s) and $(\partial v/\partial z)^{(k)}$ (s), respectively, and constant variances σ_u^2 and σ_v^2. For every hydrographic station where two ship tracks intersect at a right angle, the zonal gradient of ρ will be estimated using data from one cruise while the meridional gradient will be estimated using data from the other cruise. As a result, the two gradients can be assumed to be statistically independent. This argument can be extended to every grid site $s \in \mathcal{S}_G$. Following the Bayesian approach, the likelihood function then becomes

$$L(C,G_u,G_v|\Phi) \propto \prod_{i=1}^{n'_D}\prod_{j=1}^{n_C} \exp\left\{-\frac{[C^{(j)}_{\mathrm{obs},i} - C^{(j)}_i]^2}{2\sigma^2_{(j)}}\right\}$$
$$\times \prod_{k=1}^{7}\prod_{s\in\mathcal{L}_G} \exp\left\{-\frac{\left(G^{(k)}_{u,\mathrm{obs}}(s) - \dfrac{\partial u^{(k)}}{\partial z}(s)\right)^2}{2\sigma_u^2}\right\}$$
$$\times \prod_{k=1}^{7}\prod_{s\in\mathcal{L}_G} \exp\left\{-\frac{\left(G^{(k)}_{v,\mathrm{obs}}(s) - \dfrac{\partial v^{(k)}}{\partial z}(s)\right)^2}{2\sigma_v^2}\right\},$$

where $i = 1, \ldots, n'_D$ now indexes a site on one of the five interior lattices where data are available. The parameters to be estimated are now v, w, $\kappa^{(x)}$, $\kappa^{(y)}$, $\kappa^{(z)}$, and $C_{\partial S}$, which are again specified to be independent

under the prior distribution. The prior distributions for v and w in this case are constructed by stacking seven independent 2D GMRFs on top of each other; the distribution for $\kappa^{(z)}$ is constructed similarly except that we use the absolute value of each GMRF. The zonal velocity u is then determined via Eq. (1), which we now write for every layer. Horizontal diffusivities are again assumed to have prior exponential distributions (with a mean of 1000 $m^2 s^{-1}$), and for each tracer, the same prior (as in the quasi-3D case) is selected for values on the boundary of the inversion domain. We have also experimented with specifying a 2D prior distribution for the meridional velocity on the deepest layer and a 3D prior distribution for the vertical gradient $\partial v/\partial z$, which will consequently determine v on every layer. Our simulations in this case have shown an unrealistic flow field and higher cost functions when compared to those resulting from the 3D prior described above. Using Bayes's formula, we derive the conditional distribution of $\Phi \equiv (v, w, \kappa^{(x)}, \kappa^{(y)}, \kappa^{(z)}, C_{\partial S})$, given tracer data and thermal wind data. The vertical gradients $\partial u/\partial z$ and $\partial v/\partial z$ are approximated using centered differences for the interior layers and forward differences for the upper and lower boundaries. Markov chain Monte Carlo methods are again used to sample from this unnormalized distribution.

For each neutral-density layer specified above, the posterior mean flow and oxygen concentration are displayed in Fig. 20. A few interesting features of the 3D velocity field are to be observed: deep layers exhibit flow that is mostly zonal as opposed to shallower layers where meridional flow becomes significant. In comparison, we also derive flow by integrating thermal wind relations with respect to z (Fig. 21) and using quasi-3D results as reference values. Analyzing each layer individually, we observe that the horizontal structure of the two fields differs, especially in the central and northern regions of our domain. However, in every layer, the westward flow near 28°S is present in both fields. In addition, floats deployed to study circulation of the South Atlantic Antarctic Intermediate Water (Núñez-Riboni et al. 2005) show the same westward flow between 20° and 30°S being part of the northern branch of the South Atlantic subtropical gyre. A quantitative comparison of this flow shows that flow derived from float data has a mean speed of 4.7 ± 3.3 cm s^{-1} while that estimated from the inversion has a posterior mean speed of 3.3 cm s^{-1} and a posterior standard deviation of 0.6 cm s^{-1}.

Using data compiled from Núñez-Riboni et al. (2005), we converted the mean zonal transport across the South Atlantic Ocean into averaged mean velocities for the $\gamma = 27.45$ isopycnal and compared them to multilayer posterior mean zonally averaged x velocities (Fig. 22) to find consistent agreement between the two. A significant disagreement appears near 11°S where posterior mean flow is eastward and near 13°S where posterior mean flow is westward. Float sampling was inadequate in this region to observe this flow, and the comparison is invalid there. These are the latitudes of the South Equatorial Countercurrent (SECC) and the South Equatorial Current, and our results are consistent qualitatively with the two currents. The SECC is thought to feed a small gyre near the eastern boundary called the Angola Dome. This gyre is not evident in the estimated mean flow because it lies mainly above the upper boundary of our inversion domain at roughly 500-m depth.

Another interesting feature of the estimated flow field is that zonal jets extend vertically across several layers. Figure 23 displays the vertical structure of the zonal velocity u for the 20°W meridian. In the left panel, flow estimated using the multilayer model is used, whereas in the right panel velocities are determined by integrating thermal wind equations. In comparison to the thermal wind flow, we find that the multilayer estimate shows a stronger depth-dependent structure. We also notice smaller horizontal-scale zonal jets at lower latitudes versus broader jets farther south.

The eastward jet at 20°S that we find significant at 2000-m depth extends up to 1000-m depth, while the westward flow at 30°S is present across the entire water column. A relatively strong westward flow at 16°S present in the thermal wind flow does not appear in the multilayer estimate. The solution does not extend close enough to the equator to compare to the much smaller vertical scales found there by Treguier et al. (2003).

5. Concluding remarks

We develop a Bayesian approach to inverting tracer concentration measurements. Following the single-layer approach we add vertical structure, initially building a quasi-3D model and furthering this to a full multilayer setup. What sets this method apart from various other approaches is that the term "solution to the inverse problem" (which we define as the collection of velocities, diffusion coefficients, and tracer boundary values to be estimated) may be ambiguous in this view. Traditional inverse methods in data assimilation provide a "unique" or "best" solution, generally computed by minimizing some objective function. Measures of uncertainty in this solution are also provided. We obtain more than that. Consider the set of all possible solutions, given the data measurements and the physical constraints required by the dynamical model. Some

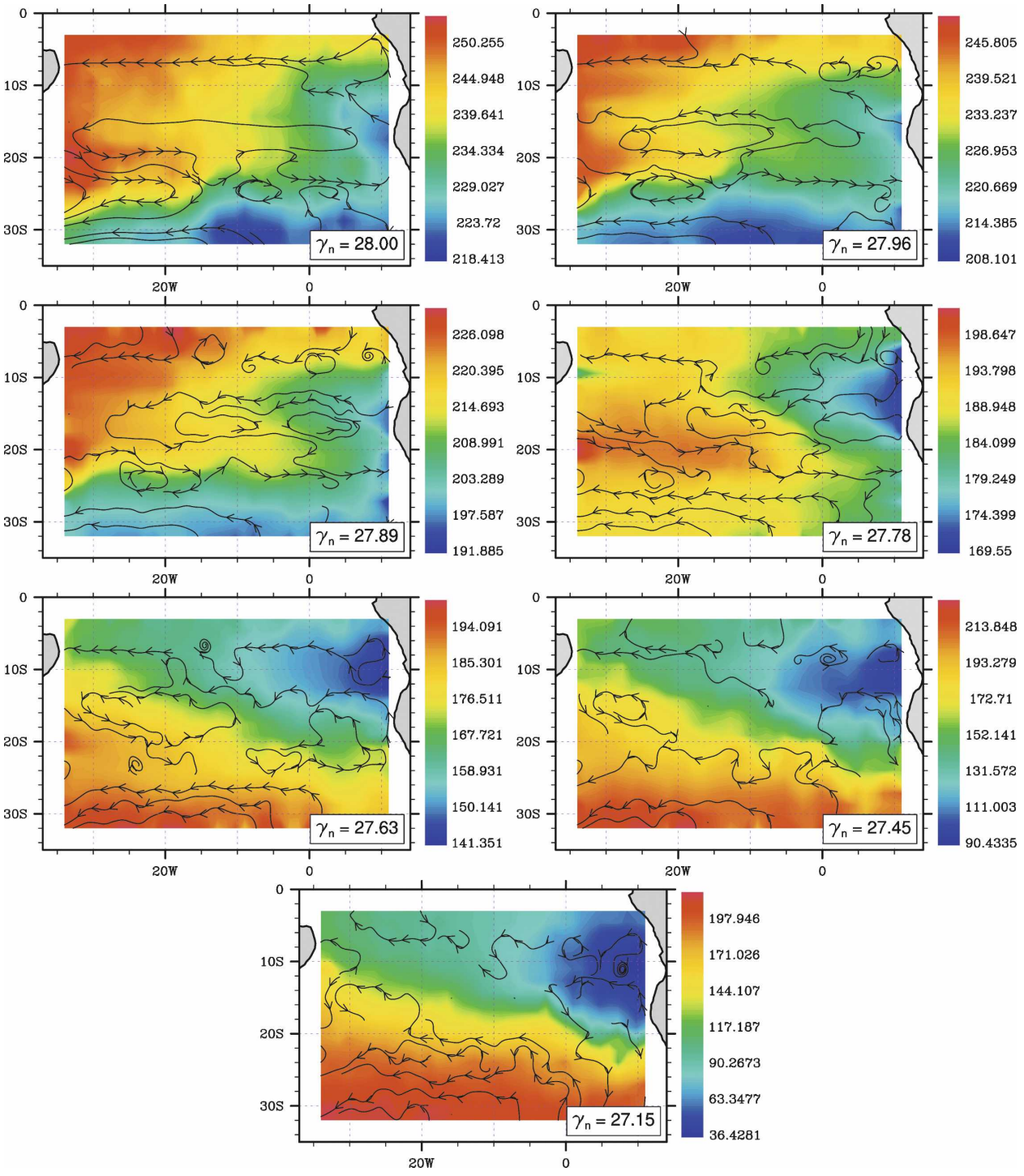

FIG. 20. Posterior mean flow plotted in streamfunction form with reduced contours, set for better visualization (no smoothing) and reconstructed oxygen tracer concentration for seven neutral-density layers.

of these solutions are more probable than others. Loosely speaking, "probable" solutions can be characterized by a realistic flow field resulting in concentration maps that are close to the observed ones. Our Bayesian approach provides us with the necessary tools to quantify "probable." A full probability distribution is obtained to characterize the state space of all solutions to the inverse problem [i.e., the parameters u, v, w, $\kappa^{(x)}$,

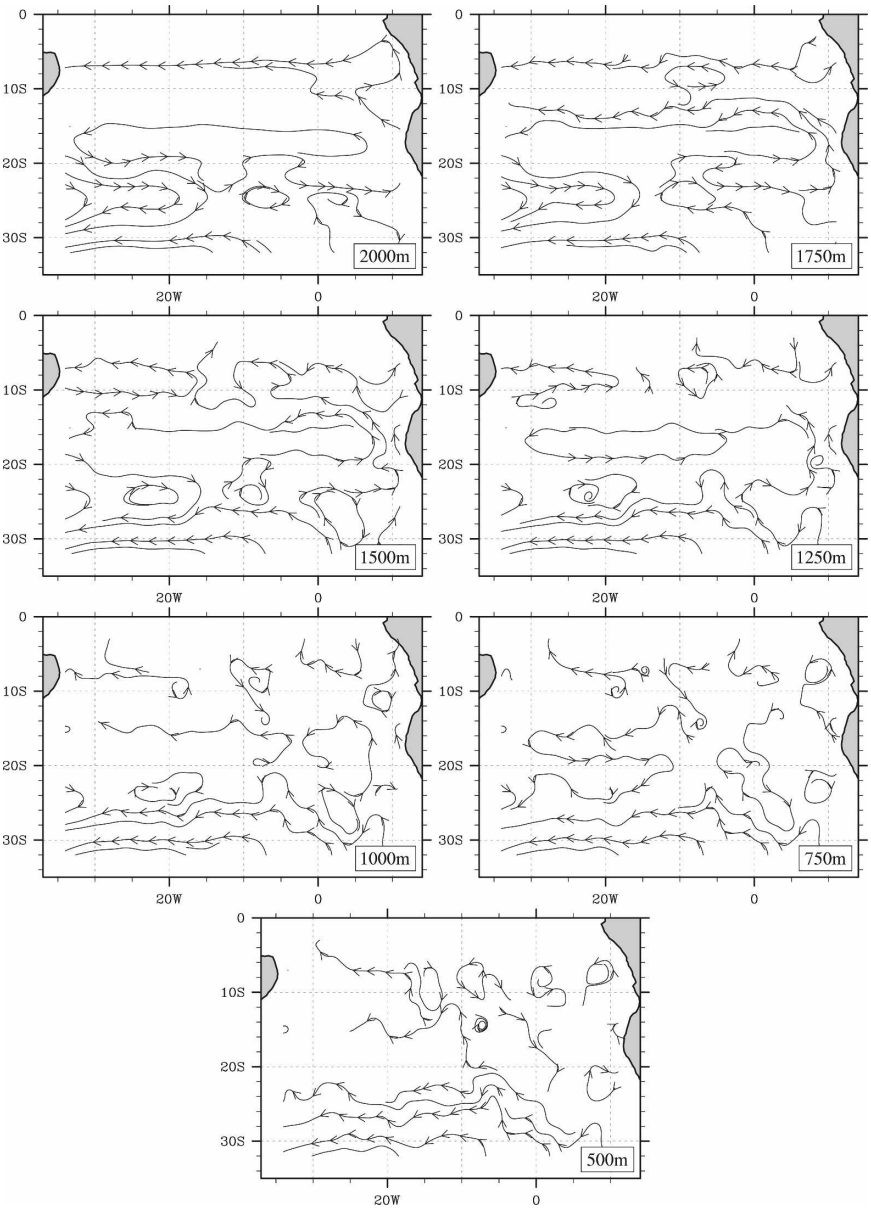

FIG. 21. Flow derived by vertically integrating thermal wind equations from 2000- to 500-m depth, plotted in streamfunction form with reduced contours, using quasi-3D results as reference level values.

$\kappa^{(y)}$, $\kappa^{(z)}$, and $C_{\partial S}$]. Hence, when inquiring about a particular variable, instead of providing a value, we are providing a distribution (including mean, variances, and so on), which is a balance between information in the data and information in the prior. There are three key ingredients in our approach: prior, dynamics, and data. All play an important role and it is impossible to uniquely determine the "source" of information. Our prior assumptions are generally mild. We only impose a certain degree of smoothness for the 2D fields and restrict diffusion parameters to positive values. We have experimented (in a limited way) with different choices for the hyperparameters that specify the prior distribution and found only minor changes in the results. When comparing posterior distributions to the prior ones, we do notice that they are quite different (many velocities now have a posterior mean that is different from zero, and horizontal diffusion coefficients appear to be Gaussian rather than exponential). We conclude that the data have swamped the prior in terms of information. Markov chain Monte Carlo techniques make our approach computationally feasible, allowing us to

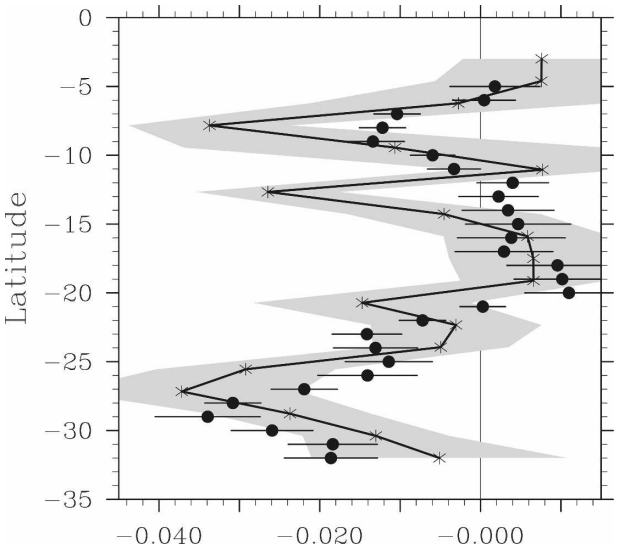

FIG. 22. Zonally averaged posterior mean zonal velocities (solid curve) in the western basin ± two posterior std devs (shaded area) and estimated zonal mean velocities from float data (dots) in the intermediate water, for $\gamma = 27.45$ neutral-density layer.

sample this high-dimensional space. Summary statistics (sample averages, most likely values, and standard deviations) may be further used to select a representative solution and to quantify uncertainty. We are able to reconstruct tracer concentration maps by numerically solving the advection–diffusion equation for a selected set of inputs (coefficients and boundary values). All the computations were carried out on a dual Power Mac G5 workstation. All the results were based on five million cycles of the MCMC sampler as described in section 2b.

We estimate abyssal horizontal flow that is dominated by alternating zonal jets consistent with observed tracer concentration maps and direct measurements of the flow at the depths of the North Atlantic Deep Water. Our results show intensified diapycnal mixing over the Mid-Atlantic Ridge (as observed by Polzin et al. 1997), which allows jets to extend over the entire domain; however, vertical diffusion driven directly by bottom topography results in a less realistic flow field. We have chosen a model configuration with a given resolution of about 1°, amounting to an operational definition of the separation of scales between subgrid-scale diffusion and mean advection. Moreover, by choosing K to be uniform over the inversion domain, we neglect its likely spatial variability, which can enter the tracer equation like advection (tracer attracted to high-energy high-K regions). We also make the choice to model tracer advection and diffusion with a mean flow and K. There is a missing nonlocal effect (not representable by K) due to the tracer being bodily translated by large eddies and deposited somewhere else. This might be represented by random eddy flux vectors but it is not implemented in the current configuration. Thus, there are several open questions regarding parameterization of tracer transport in the ocean (or atmosphere, or many other systems) that might be investigated. Our study provides the answer to a simple version of the problem, and a basis for further investigation.

We extend the approach and use thermal wind relations to construct a multilayer model. Integrating thermal wind equations with respect to z over the entire water column produces a flow field that contains eddies and is not necessarily consistent with observed tracer maps. We avoid these problems by imposing thermal wind weakly in a local fashion. As a consequence, the vertical structure of the flow departs from the one de-

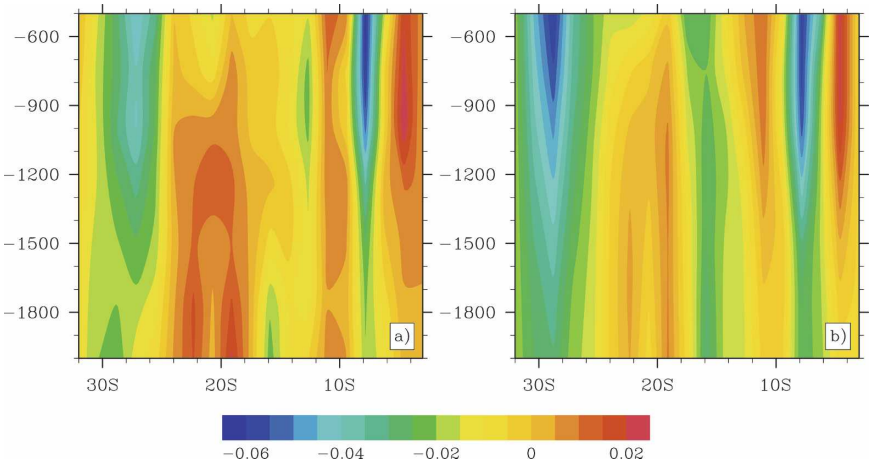

FIG. 23. Vertical structure for the zonal velocity u at 20°W. (a) The posterior mean. (b) Here, u is determined by integrating thermal wind equations.

rived by integrating thermal wind relations, revealing a strongly depth-dependent flow field. Alternating zonal jets extending vertically over 1000 m or more are still the main feature of the flow field. We understand that we make a subjective choice to impose thermal wind relations weakly to provide the vertical structure of the flow. We do account for thermal wind error in this manner, but in our approach we assume that the depths of each isopycnal are known. A further stage in our analysis will relax this assumption, building thermal wind into the dynamical model and working with an updated density field.

Acknowledgments. This research was supported in part by NSF Grants ATM-0222244 and OCE-0336697.

APPENDIX

Bayesian Approach, MCMC Moves, and Acceptance Probabilities

The aim of our paper is to estimate $\Phi = (\underline{\mathbf{u}}, K, C_{\partial S})$, given measurements of tracer concentration C. We are pursuing a Bayesian approach that consists of two major steps. The first one is the model specification, which has two components: a prior model for the parameters Φ and a model for the data C (conditional on Φ). The data model (or observation model) has the general form Data = $f(\Phi, \text{error})$. Assumptions of distributions of errors lead to the likelihood function $L(C|\Phi)$. The prior model describes all the (subjective) information the user has about the parameters (prior knowledge). Formally, this amounts to specifying the prior density $\pi(\Phi)$. Acquisition of information (the data) will modify the prior model into the posterior $\pi(\Phi|C)$, according to the Bayes Eq. (5). The posterior density is the proposed solution: a probability measure on the space of the parameters Φ. It allows us to quantify which sets of parameters are more or less likely and to compute a mean value, a variance, or other measures of uncertainty. The drawback is that $\pi(\Phi|C)$ is not normalized, and in most cases, such as the one presented in this paper, the normalizing constant is intractable. Thus, formally summarizing $\pi(\Phi|C)$ is impossible; however, using MCMC techniques, we can numerically draw samples from this distribution.

The second step in the Bayesian approach consists in constructing a Markov chain, which explores the state space of Φ in a manner consistent with $\pi(\Phi|C)$. A very popular method to do so is a Metropolis random walk. The chain is started in a state Φ_0, selected arbitrarily. From the current state Φ_j the chain moves to $\Phi_{j+1} = \Phi_j + r$ with probability

$$\alpha(\Phi_{j+1}|\Phi_j) = \min\left[1, \frac{\pi(\Phi_{j+1}|C)}{\pi(\Phi_j|C)}\right],$$

or stays put, $\Phi_{j+1} = \Phi_j$, with probability $1 - \alpha(\Phi_{j+1}|\Phi_j)$. The sequence Φ_j, $j = 0, 1, 2, \ldots$ constitutes a realization of the Markov chain. Here r is a uniform (zero mean) variable, and $\alpha(\Phi_{j+1}|\Phi_j)$ is called the acceptance probability. Standard Markov chain theory (Tierney 1994) provides sufficient conditions for the random walk Metropolis algorithm to converge to the target distribution $\pi(\Phi|C)$. Once the chain has converged, which we determine by examining trace plots (as in Fig. 11), subsequent output is used to obtain inversion results. In practice, however, several other types of moves (rather than simple random walk) are designed to improve the quality of the posterior sample. We describe all these moves formally as reversible jumps (Green 1995). From the current state Φ we propose to move to $\Phi' = F(\Phi, r)$, F being a nonrandom operator and r having a given density q. The new state Φ' is accepted with probability

$$\alpha(\Phi'|\Phi) = \min\left(1, \frac{\pi(\Phi'|\text{Data})q(r')}{\pi(\Phi|\text{Data})q(r)}\left|\frac{\partial(\Phi', r')}{\partial(\Phi, r)}\right|\right),$$

where r' is the unique value such that $\Phi = F(\Phi', r')$ and Data is a generic notation for tracer data C in the quasi-3D setup or tracer data combined with thermal wind data G for the multilayer setup. We combine classical random walk–type moves with several new scaling moves such as

$$[\upsilon', \kappa^{(z)'}] = [r\upsilon, \kappa^{(z)}/r], \quad [\kappa^{(x)'}, \kappa^{(y)'}] = [r\kappa^{(x)}, r\kappa^{(y)}],$$

$$[\kappa^{(x)'}, \kappa^{(y)'}] = [r\kappa^{(x)}, \kappa^{(y)}/r].$$

For such moves, the Peclet number will vary strongly. These moves allow the Markov chain to jump quickly between states for which the advective–diffusive balances are different.

REFERENCES

Adams, J., 1991: Multigrid software for elliptic partial differential equations: MUDPACK. NCAR Tech. Note NCAR/TN-357+STR, 51 pp.

Armi, L., and D. B. Haidvogel, 1982: Effects of variable and anisotropic diffusivities in a steady-state diffusion model. *J. Phys. Oceanogr.,* **12,** 785–794.

Feely, R. A., C. L. Sabine, R. Schlitzer, J. L. Bullister, S. Mecking, and D. Greeley, 2004: Oxygen utilization and organic carbon remineralization in the upper water column of the Pacific Ocean. *J. Oceanogr.,* **60,** 45–52.

Green, P., 1995: Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika,* **82,** 711–732.

Hogg, N. G., and W. B. Owens, 1999: Direct measurement of the deep circulation within the Brazil Basin. *Deep-Sea Res. II,* **46,** 335–353.

Huang, R. X., and X. Jin, 2002: Deep circulation in the South Atlantic induced by bottom-intensified mixing over the mid-ocean ridge. *J. Phys. Oceanogr.,* **32,** 1150–1164.

Liu, J. S., 2001: *Monte Carlo Strategies in Scientific Computing.* Springer, 343 pp.

Maximenko, N. A., G. Bang, and H. Sasaki, 2005: Observational evidence of alternating zonal jets in the world ocean. *Geophys. Res. Lett.,* **32,** L12607, doi:10.1029/2005GL022728.

McDougall, T. J., 1987: Neutral surfaces. *J. Phys. Oceanogr.,* **17,** 1950–1964.

——, 1991: Parameterizing mixing in inverse models. *Dynamics of Oceanic Internal Gravity Waves: Proc. 'Aha Huliko'a Hawaiian Winter Workshop,* Honolulu, HI, University of Hawaii at Manoa, 355–386.

McKeague, I. W., G. K. Nicholls, K. G. Speer, and R. Herbei, 2005: Statistical inversion of South Atlantic circulation in an abyssal neutral density layer. *J. Mar. Res.,* **63,** 683–704.

Nakano, H., and N. Suginohara, 2002: A series of mid-depth zonal flows in the Pacific driven by winds. *J. Phys. Oceanogr.,* **32,** 161–176.

Núñez-Riboni, I., O. Boebel, M. Ollitrault, Y. You, P. Richardson, and R. Davis, 2005: Lagrangian circulation of Antarctic Intermediate Water in the subtropical South Atlantic. *Deep-Sea Res. II,* **52,** 545–564.

Polzin, K. L., J. M. Toole, J. R. Ledwell, and R. W. Schmitt, 1997: Spatial variability of turbulent mixing in the abyssal ocean. *Science,* **276,** 93–96.

Richards, K. J., N. A. Maximenko, F. O. Bryan, and H. Sasaki, 2006: Zonal jets in the Pacific Ocean. *Geophys. Res. Lett.,* **33,** L03605, doi:10.1029/2005GL024645.

Sundermeyer, M. A., and J. F. Price, 1998: Lateral mixing and the North Atlantic Tracer Release Experiment: Observations and numerical simulations of Lagrangian particles and a passive tracer. *J. Geophys. Res.,* **103** (C10), 21 481–21 497.

Thurnherr, A. M., and K. G. Speer, 2004: Representativeness of meridional hydrographic sections in the western South Atlantic. *J. Mar. Res.,* **62,** 37–65.

Tierney, L., 1994: Markov chains for exploring posterior distributions. *Ann. Stat.,* **22,** 1701–1728.

Toole, J. M., K. L. Polzin, and R. W. Schmitt, 1994: Estimates of diapycnal mixing in the abyssal ocean. *Science,* **264,** 1120–1123.

Treguier, A. M., N. G. Hogg, M. Maltrud, K. G. Speer, and V. Thierry, 2003: The origin of deep zonal flows in the Brazil Basin. *J. Phys. Oceanogr.,* **33,** 580–599.