A Thermohaline Inverse Method for Estimating Diathermohaline Circulation and Mixing

Sjoerd Groeskamp

*CSIRO Wealth from Oceans National Research Flagship, and Institute for Marine and Antarctic Studies,
University of Tasmania, Hobart, Tasmania, Australia*

Jan D. Zika

University of Southampton, National Oceanography Centre, Southampton, United Kingdom

Bernadette M. Sloyan

*CSIRO Wealth from Oceans National Research Flagship, and Centre for Australian Weather and Climate Research,
CSIRO Marine and Atmospheric Research, Hobart, Tasmania, Australia*

Trevor J. McDougall

School of Mathematics and Statistics, University of New South Wales, Sydney, New South Wales, Australia

Peter C. McIntosh

Centre for Australian Weather and Climate Research, CSIRO Marine and Atmospheric Research, Hobart, Tasmania, Australia

ABSTRACT

The thermohaline inverse method (THIM) is presented that provides estimates of the diathermohaline streamfunction $\Psi_{S_A\Theta}^{\mathrm{dia}}$, the downgradient along-isopycnal diffusion coefficient K, and the isotropic downgradient turbulent diffusion coefficient D of small-scale mixing processes. This is accomplished by using the water mass transformation framework in two tracer dimensions: here in Absolute Salinity S_A and Conservative Temperature Θ coordinates. The authors show that a diathermal volume transport down a Conservative Temperature gradient is related to surface heating and cooling and mixing, and a diahaline volume transport down an Absolute Salinity gradient is related to surface freshwater fluxes and mixing. Both the diahaline and diathermal flows can be calculated using readily observed parameters that are used to produce climatologies, surface flux products, and mixing parameterizations for K and D. Conservation statements for volume, salt, and heat in (S_A, Θ) coordinates, using the diahaline and diathermal volume transport expressed as surface freshwater and heat fluxes and mixing, allow for the formulation of a system of equations that is solved by an inverse method that can estimate the unknown diathermohaline streamfunction $\Psi_{S_A\Theta}^{\mathrm{dia}}$ and the diffusion coefficients K and D. The inverse solution provides an accurate estimate of $\Psi_{S_A\Theta}^{\mathrm{dia}}$, K, and D when tested against a numerical climate model for which all these parameters are known.

1. Introduction

Mixing in the ocean influences Earth's climate through its ability to alter the ocean's circulation and uptake and distribution of tracers such as heat, oxygen, and carbon. An increased understanding of ocean mixing, both observationally and its representation in models, is necessary to better understand and model the ocean's influence on the climate system.

Currently climate ocean models parameterize interior ocean mixing as downgradient epineutral diffusion (along-isopycnal diffusion), with diffusion coefficient K, and dianeutral downgradient turbulent diffusion due to small-scale mixing, with diffusion coefficient D (Redi 1982; Griffies 2004). The spatial- and temporal-varying magnitudes of K and D are not easily obtained from theory. Therefore, we require empirically (observationally)

based estimates to improve our understanding of K and D.

Munk (1966) provided such an estimate using an approximate balance between (vertical) ocean advection and mixing along with tracer observations to obtain a steady-state estimate of D. Munk's study demonstrated that observed estimates of tracers can be used to obtain estimates of ocean circulation and mixing.

To obtain estimates of the structure and magnitude of the ocean circulation from observations, Stommel and Schott (1977) and Wunsch (1978) introduced inverse methods into the field of oceanography. Ever since, many inverse studies have provided observationally based estimates of circulation (Schott and Stommel 1978; Killworth 1986; Cunningham 2000; Sloyan and Rintoul 2000, 2001) of which some used the diapycnal fluxes to provide estimates for D (Ganachaud and Wunsch 2000; Lumpkin and Speer 2007). Zhang and Hogg (1992) and Zika et al. (2010a) developed an inverse method that, simultaneously, solves for the circulation and both K and D.

Ocean circulation is often represented by a volumetric streamfunction, simplifying the three-dimensional time-varying (global) ocean circulation into a two-dimensional time-averaged circulation. The streamfunction has been defined using different combinations of coordinates, both geographic and thermodynamic (Bryan et al. 1985; Döös and Webb 1994; Hirst et al. 1996; Hirst and McDougall 1998; Nycander et al. 2007). Based on an averaging technique developed by Nurser and Lee (2004), Ferrari and Ferreira (2011) have defined an advective meridional streamfunction $\Psi_{C_1 y}$ for an instantaneous velocity field v and arbitrary tracer C_1. Zika et al. (2012) generalized this technique to compute a mean advective streamfunction $\Psi_{C_1 C_2}$ of any two tracers, C_1 and C_2, for an instantaneous velocity field, $\mathbf{u} = (u, v, w)$. Here $\Psi_{C_1 C_2}$ represents the advective transport of fluid parcels in (C_1, C_2) coordinates. Zika et al. (2012) applied this to salinity S and temperature T coordinates to obtain the advective thermohaline streamfunction Ψ_{ST}^{adv}. In independent work, Döös et al. (2012) also developed Ψ_{ST}^{adv}.

Recently, Groeskamp et al. (2014) showed how the total circulation in (S, T) coordinates is driven by thermohaline forcing, that is, surface freshwater and heat fluxes and salt and heat fluxes by diffusive mixing. They showed that the total circulation in (S, T) coordinates is in fact a summation of the advective thermohaline streamfunction Ψ_{ST}^{adv} (as calculated in recent literature), the local thermohaline streamfunction Ψ_{ST}^{loc}, and the thermohaline trend. The sum of Ψ_{ST}^{adv} and Ψ_{ST}^{loc} represents the total nondivergent diathermohaline ocean circulation in (S, T) coordinates and is quantified by the diathermohaline streamfunction Ψ_{ST}^{dia}.

Only $\Psi_{ST}^{dia} = \Psi_{ST}^{adv} + \Psi_{ST}^{loc}$ can be directly related to thermohaline forcing. Zika et al. (2012), Döös et al. (2012), and Groeskamp et al. (2014) showed Ψ_{ST}^{dia}, Ψ_{ST}^{adv}, and Ψ_{ST}^{loc} are all useful model diagnostics to understand the relation between thermohaline forcing and ocean circulation.

Calculation of Ψ_{ST}^{dia}, as in Groeskamp et al. (2014), requires accurate three-dimensional velocity data and an ocean hydrography. To avoid the use of velocity data to calculate a diathermal advection, Walin (1982) showed that the area-integrated surface heat flux between two outcropping isotherms, combined with a diffusive heat flux across the bounding isotherms can be used to calculate the diathermal advection in a steady-state ocean (in which the volume enclosed between two isotherms remains constant). He thus framed the relation between surface and interior heat fluxes and the cross-isothermal ocean circulation.

To obtain the diathermohaline streamfunction as defined by Groeskamp et al. (2014), we require both diathermal and diahaline transport. To obtain both transports we will, like Speer and Tziperman (1992) and Hieronymus et al. (2014), use salt and heat fluxes in combination with Walin's framework to estimate water mass transformation rates in (S, T) coordinates. Here a water mass transformation is a change of the S and T properties of a water mass, which can result in along- and cross-isopycnal transport (Speer 1993; IOC et al. 2010). Speer (1993) suggested that in a steady-state ocean, the net divergence of the water mass transformation due to surface heat and freshwater fluxes, projected in (S, T) coordinates, should be balanced by mixing, thus providing constrains on mixing estimates.

In the present paper, we will merge both the Munk (1966) and Walin (1982) frameworks in (S, T) coordinates, obtaining a balance between surface forcing, mixing, and circulation. This is obtained using a two-dimensional extension into (S, T) coordinates by applying Walin's framework to a volume bounded by a pair of isotherms *and* a pair of isohalines (Fig. 1). We can then provide an estimate of the diathermohaline circulation using boundary salt and heat fluxes and diffusive mixing. Representing the diathermohaline circulation by a diathermohaline streamfunction we develop the thermohaline inverse method (THIM) that can be applied to observationally based ocean hydrography and surface heat and freshwater fluxes to simultaneously obtain estimates of both Ψ_{ST}^{dia} and the tracer diffusion coefficients K and D. We test the THIM by calculating Ψ_{ST}^{dia} directly using a model's hydrographic and velocity output as described by Groeskamp et al. (2014) and compare this with an inverse estimate of Ψ_{ST}^{dia} based on the model's boundary fluxes of salt and heat in combination with its hydrography.
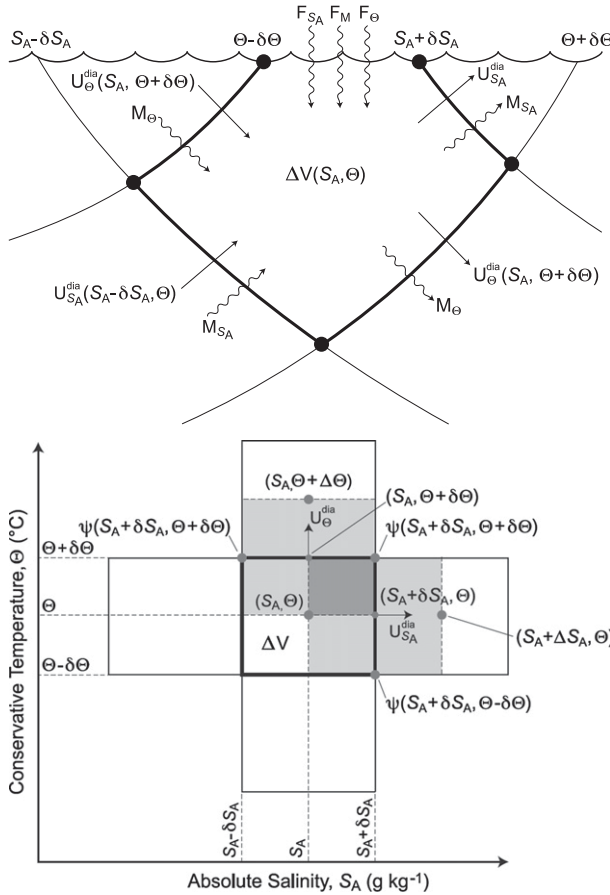
FIG. 1. (top) The two-dimensional (S_A, Θ) (thermohaline) version of the Walin (1982) framework in Cartesian coordinates. Advection through the pair of isotherms and isohalines enclosing ΔV (bold black lines); requires a change in the S_A and Θ values of a part of ΔV. For such a change, the divergence of salt or heat fluxes is required. These are provided by the thermohaline forcing terms, which are the diffusion of salt and heat through the isotherms and isohalines (M_{S_A} and M_Θ, respectively) and the surface fluxes of mass, salt, and heat (F_{S_A}, F_Θ, and F_m, respectively). (bottom) The same ΔV represented in (S_A, Θ) coordinates (bold black lines). In these coordinates we have also defined the streamfunction, such that the diathermohaline streamfunction difference is equal to the net flux through the isotherm or isohaline, which itself is provided by salt and heat fluxes gridded in the grid spanning $(S_A \pm \delta S_A, \Theta \pm \delta\Theta)$, where (S_A, Θ) is the location at which $\mathbf{U}^{\text{dia}}_{S_A\Theta}$ is evaluated.

2. Diathermohaline streamfunction

This section is a summary of a derivation by Groeskamp et al. (2014) leading to an expression for the diathermohaline streamfunction $\Psi^{\text{dia}}_{S_A\Theta}$, which represents ocean circulation in Absolute Salinity S_A and Conservative Temperature Θ coordinates. Here Conservative Temperature is proportional to potential enthalpy (by the constant heat capacity factor c_p^0), which represents the heat content per unit mass of seawater (McDougall 2003; Graham and McDougall 2013). Absolute Salinity

is measured on the Reference Composition Salinity Scale (Millero et al. 2008) and represents the mass fraction of dissolved material in seawater (g kg^{-1}) (IOC et al. 2010; McDougall et al. 2012).

The superscript or prefix "dia" indicates a transport through a surface. Hence, we define $u_C^{\text{dia}}(\mathbf{x}, t)$ as the velocity of a fluid parcel through a surface of constant, conserved tracer $C = C(\mathbf{x}, t)$, where $\mathbf{x} = (x, y, z)$. Hence, $u_C^{\text{dia}}(\mathbf{x}, t)$ can be obtained from the material derivative of C (Griffies 2004; Groeskamp et al. 2014):

$$
\begin{aligned}
u_C^{\text{dia}} &= \frac{1}{|\boldsymbol{\nabla} C|} \frac{DC}{Dt} \\
&= \frac{1}{|\boldsymbol{\nabla} C|} \frac{\partial C}{\partial t} + \mathbf{u} \cdot \frac{\boldsymbol{\nabla} C}{|\boldsymbol{\nabla} C|} \\
&= \frac{1}{|\boldsymbol{\nabla} C|} (f_C + m_C).
\end{aligned} \tag{1}
$$

Here $\mathbf{u} = \mathbf{u}(\mathbf{x}, t) = [u(\mathbf{x}, t), v(\mathbf{x}, t), w(\mathbf{x}, t)]$ and the forcing terms $f_C = f_C(\mathbf{x}, t)$ and $m_C = m_C(\mathbf{x}, t)$ (both in $C \, \text{s}^{-1}$) are flux divergences of C due to boundary fluxes and diffusive mixing processes, respectively. Equation (1) shows that $u_C^{\text{dia}} = u_C^{\text{dia}}(\mathbf{x}, t)$ is the difference between the velocity of the fluid parcel in the direction normal to the surface of constant C $[\mathbf{u} \cdot (\boldsymbol{\nabla} C/|\boldsymbol{\nabla} C|)]$ and the movement of the surface itself $[(1/|\boldsymbol{\nabla} C|)(\partial C/\partial t)]$.

A trend in C over time period Δt leads to a net shift of the geographical position of the surface of constant C. This shift can be expressed as a net velocity of the surface of constant C over Δt, given by

$$
u_C^{\text{tr}} = \frac{1}{|\boldsymbol{\nabla} C|} \frac{1}{\Delta t} \int_t^{t+\Delta t} \frac{\partial C}{\partial t} \, dt. \tag{2}
$$

Here $1/\Delta t \int_t^{t+\Delta t} (\dots) \, dt = \overline{(\dots)}$ is time averaged (later denoted by an overbar), and $u_C^{\text{tr}} = u_C^{\text{tr}}(\mathbf{x}, t)$.

Inserting $C = S_A(\mathbf{x}, t)$ and $C = \Theta(\mathbf{x}, t)$ in Eq. (1), we can define the diathermohaline velocity vector $\mathbf{u}^{\text{dia}}_{S_A\Theta} = (u_{S_A}^{\text{dia}}, u_\Theta^{\text{dia}})$ as the velocity with which the fluid parcel crosses isohalines and isotherms. Inserting $C = S_A(\mathbf{x}, t)$ and $C = \Theta(\mathbf{x}, t)$ in Eq. (2), we can define the diathermohaline trend $\mathbf{u}^{\text{tr}}_{S_A\Theta} = (u_{S_A}^{\text{tr}}, u_\Theta^{\text{tr}})$ as the net velocity of a shift of the geographical position of the isohalines and isotherms due to a trend in time in the local changes of S_A and Θ. Following Groeskamp et al. (2014), for a Boussinesq ocean in which $\boldsymbol{\nabla} \cdot \mathbf{u} = 0$, the diathermohaline streamfunction $\Psi^{\text{dia}}_{S_A\Theta}$ is then given by

$$
\begin{aligned}
\Psi^{\text{dia}}_{S_A\Theta}(S_A, \Theta) &= \int_t^{t+\Delta t} \int_{\Theta' \le \Theta|_{S_A}} u_{S_A}^{\text{dia}} - u_{S_A}^{\text{tr}} \, dA \, dt \\
&= -\int_t^{t+\Delta t} \int_{S_A' \le S_A|_\Theta} u_\Theta^{\text{dia}} - u_\Theta^{\text{tr}} \, dA \, dt.
\end{aligned} \tag{3}
$$

Here $\int_{\Theta' \leq \Theta|_{S_A}} dA$ is the area integral over all Θ', smaller than or equal to Θ on a surface of constant S_A, and $\int_{S'_A \leq S_A|_\Theta} dA$ is the area integral over all S'_A, smaller than or equal to S_A on a surface of constant Θ. For a statistically steady ocean, $\mathbf{u}^{\text{dia}}_{S_A\Theta} = 0$. Hence, $\Psi^{\text{dia}}_{S_A\Theta}$ represents the nondivergent component of the ocean circulation in (S_A, Θ) coordinates, while the diathermohaline trend represents the divergent component of this circulation.

From Eqs. (1) and (3) it is clear that $\Psi^{\text{dia}}_{S_A\Theta}$ can only be calculated if $\mathbf{u}(\mathbf{x}, t)$ is known. An expression that provides $\Psi^{\text{dia}}_{S_A\Theta}$ from commonly observed variables (ocean hydrography and boundary salt and heat fluxes) is obtained when inserting the last line of Eq. (1) into Eq. (3), thereby directly relating $\Psi^{\text{dia}}_{S_A\Theta}$ to thermohaline forcing:

$$\Psi^{\text{dia}}_{S_A\Theta}(S_A, \Theta) = \int_t^{t+\Delta t} \int_{\Theta' \leq \Theta|_{S_A}} \frac{f_{S_A} + m_{S_A}}{|\nabla S_A|} - u^{\text{tr}}_{S_A} \, dA \, dt$$
$$= -\int_t^{t+\Delta t} \int_{S'_A \leq S_A|_\Theta} \frac{f_\Theta + m_\Theta}{|\nabla \Theta|} - u^{\text{tr}}_\Theta \, dA \, dt. \tag{4}$$

However, this expression requires knowledge of the ocean's diffusive salt and heat fluxes everywhere. Unfortunately the ocean's spatial- and temporal-varying diffusive salt and heat fluxes are not well known, requiring us to develop a different method to derive $\Psi^{\text{dia}}_{S_A\Theta}$ from observations, which is what we do in this paper.

3. The diathermohaline volume transport

The diathermohaline velocity $\mathbf{u}^{\text{dia}}_{S_A\Theta}$ integrated over an area results in a diathermohaline volume transport. In this section, we will use the conservation of volume, salt, and heat to derive an expression for the diathermohaline volume transport as a function of thermohaline forcing. In section 4, we relate the diathermohaline volume transport with $\Psi^{\text{dia}}_{S_A\Theta}$ to obtain an expression for $\Psi^{\text{dia}}_{S_A\Theta}$ as a function of the thermohaline forcing. This expression allows us to construct an inverse method that will provide an estimate for both the ocean's tracer diffusion coefficients and the diathermohaline streamfunction.

To express the conservation equations we consider a volume ΔV, bounded by a pair of isotherms that are separated by $\Delta\Theta$ ($= 2\delta\Theta$) and a pair of isohalines that are separated by ΔS_A ($= 2\delta S_A$). The volume's Θ ranges between $\Theta \pm \delta\Theta$ and S_A ranges between $S_A \pm \delta S_A$. As a result, $\Delta V = \Delta V(S_A \pm \delta S_A, \Theta \pm \delta\Theta, t)$ may have any shape in (x, y, z) coordinates (Fig. 1a), but it covers a square grid in (S_A, Θ) coordinates (Fig. 1b).

At the center of ΔV, at coordinates (S_A, Θ), we define a diathermohaline volume transport vector. The diahaline

volume transport, in the positive S_A direction through the area of the surface of constant S_A, for the Θ range of $\Theta \pm \delta\Theta$ is given by

$$U^{\text{dia}}_{|S_A}(S_A, \Theta \pm \delta\Theta, t) = \int_{\Theta \pm \delta\Theta|_{S_A}} u^{\text{dia}}_{S_A} \, dA. \tag{5}$$

The diathermal volume transport, in the positive Θ direction through the area of the surface of constant Θ, for the S_A range of $S_A \pm \delta S_A$, is given by

$$U^{\text{dia}}_{|\Theta}(S_A \pm \delta S_A, \Theta, t) = \int_{S_A \pm \delta S_A|_\Theta} u^{\text{dia}}_\Theta \, dA. \tag{6}$$

This derivation results in the diathermohaline volume transport vector $\mathbf{U}^{\text{dia}}_{S_A\Theta} = (U^{\text{dia}}_{|S_A}, U^{\text{dia}}_{|\Theta})$.

Using $\mathbf{U}^{\text{dia}}_{S_A\Theta}$, we can construct the conservation equations for volume, salt, and heat for ΔV. It would be more accurate to use conservation of mass in a non-Boussinesq ocean, but we leave this for future work. As $\Delta V = \Delta V(S_A \pm \delta S_A, \Theta \pm \delta\Theta, t)$, this results in $S_A = S_A(t)$ and $\Theta = \Theta(t)$, where S_A and Θ can only vary in time within the range $S_A \pm \delta S_A$ and $\Theta \pm \delta\Theta$, respectively. Using this and applying the Boussinesq approximation, the conservation of volume, salt, and heat for ΔV is given by

$$\frac{\partial \Delta V}{\partial t} + \nabla_{S_A\Theta} \cdot \mathbf{U}^{\text{dia}}_{S_A\Theta} = \frac{1}{\rho_0} F_m, \tag{7}$$

$$\frac{\partial \Delta V S_A}{\partial t} + \nabla_{S_A\Theta} \cdot (S_A \mathbf{U}^{\text{dia}}_{S_A\Theta}) = \frac{1}{\rho_0} M_{S_A}, \quad \text{and} \tag{8}$$

$$\frac{\partial \Delta V \Theta}{\partial t} + \nabla_{S_A\Theta} \cdot (\Theta \mathbf{U}^{\text{dia}}_{S_A\Theta}) = \frac{1}{\rho_0 c_p^0} (F_\Theta + M_\Theta). \tag{9}$$

Here we have used that the derivative of $U^{\text{dia}}_{|S_A}$ ($U^{\text{dia}}_{|\Theta}$), with respect to S_A (Θ), is constant over the interval $S_A \pm \delta S_A$ ($\Theta \pm \delta\Theta$), such that we can define the thermohaline divergence operator

$$\nabla_{S_A\Theta} = \left(\Delta S_A \frac{\partial}{\partial S_A}, \Delta\Theta \frac{\partial}{\partial \Theta} \right). \tag{10}$$

The thermohaline divergence operator [Eq. (10)] is a short-hand notation for taking the difference of the outflow and inflow of volume and accompanied tracers, through the pair of isohalines that enclose ΔV that are separated exactly by ΔS_A, and isotherms separated exactly by $\Delta\Theta$.

In Eq. (7), F_m is the boundary mass flux into ΔV due to evaporation E, precipitation P, ice melt and formation, and river runoff R. We assumed that the boundary salt flux is zero, that is, neglecting the formation of sea spray,

the interchange of salt with sea ice and salt entering from the ocean boundaries. Although the total amount of salt in the ocean remains constant, the ocean's salinity is modified by F_m (Huang 1993; Griffies 2004). The term F_Θ is the net convergence of heat into ΔV due to boundary fluxes. The values F_m and F_Θ can be obtained from surface freshwater and heat flux products. The terms M_{S_A} and M_Θ are the net convergence of salt and heat into ΔV due to *all* interior diffusive processes and can be obtained from an ocean hydrography in combination with a mixing parameterization.

After expanding the lhs of Eqs. (8) and (9), we insert Eq. (7) into Eqs. (8) and (9), and we use that by definition in these coordinates $\partial S_A / \partial \Theta = \partial \Theta / \partial S_A = 0$. Taking the time average, this results in the following expression for $\overline{\mathbf{U}}_{S_A\Theta}^{\mathrm{dia}} = (\overline{U}_{|S_A}^{\mathrm{dia}}, \overline{U}_{|\Theta}^{\mathrm{dia}})$:

$$\overline{U}_{|S_A}^{\mathrm{dia}}(S_A, \Theta \pm \delta\Theta)$$
$$= \frac{1}{\rho_0 \Delta S_A}\left(\overline{M_{S_A}} - \overline{S_A F_m} - \rho_0 \overline{\Delta V \frac{\partial S_A}{\partial t}}\right), \quad (11)$$

$$\overline{U}_{|\Theta}^{\mathrm{dia}}(S_A \pm \delta S_A, \Theta)$$
$$= \frac{1}{\rho_0 c_p^0 \Delta\Theta}\left(\overline{F_\Theta} + \overline{M_\Theta} - \overline{\Theta F_m} - \rho_0 c_p^0 \overline{\Delta V \frac{\partial\Theta}{\partial t}}\right). \quad (12)$$

Here we have obtained an expression for the diathermohaline volume transport out of ΔV due to water mass transformation as a result of a convergence of salt or heat into ΔV. This is equivalent to an extension of Walin (1982)'s framework in (S_A, Θ) coordinates. The vector $\overline{\mathbf{U}}_{S_A\Theta}^{\mathrm{dia}} = (\overline{U}_{|S_A}^{\mathrm{dia}}, \overline{U}_{|\Theta}^{\mathrm{dia}})$ is equivalent to the **J** vector as defined by Hieronymus et al. (2014), except for the last terms on the rhs of Eqs. (11) and (12).

We will now calculate the terms on the right-hand side of Eqs. (11) and (12) in detail. In section 4, we relate $\overline{\mathbf{U}}_{S_A\Theta}^{\mathrm{dia}}$ with $\Psi_{S_A\Theta}^{\mathrm{dia}}$ such that we can estimate $\Psi_{S_A\Theta}^{\mathrm{dia}}$, from observationally based products, using the thermohaline inverse method.

a. Boundary salt and heat fluxes

The boundary mass flux into the ocean $\overline{F}_m$ (kg s^{-1}) is given by the integral of $E - P - R$ over area A_b, bounded by a pair of isohalines and isotherms:

$$\overline{F}_m = -\frac{1}{\Delta t}\int_t^{t+\Delta t}\int_{A_b}(E - P - R)\,dA\,dt. \quad (13)$$

Modification of the ocean's salinity through $\overline{F}_m$ can be calculated as

$$\overline{S_A F_m} = -\frac{1}{\Delta t}\int_t^{t+\Delta t} S_A\int_{A_b}(E - P - R)\,dA\,dt$$
$$= -\overline{S}_A\overline{F}_m - \overline{S_A' F_m'}. \quad (14)$$

Here we have applied Reynolds decomposition and averaging. Note that S_A' may only change within $S_A \pm \delta S_A$; hence, S_A' approaches 0 when δS_A approaches 0. A simple scale analysis shows that the heat flux equivalent $\overline{\Theta F_m}$ in the heat conservation Eq. (6) is negligible compared to the other terms.

The boundary heat flux into the ocean F_Θ (J s^{-1}) is the integral of the surface heat flux $f_h(x, y)$ due to longwave and shortwave radiation, and the latent and sensible heat flux, and geothermal heating over area A_b:

$$\overline{F}_\Theta = \frac{1}{\Delta t}\int_t^{t+\Delta t}\int_{A_b} f_h\,dA\,dt. \quad (15)$$

Note that, if required, one can include effects of solar penetration below the surface (Iudicone et al. 2008).

b. Diffusive salt and heat fluxes

Generally, tracer diffusion is parameterized according to two different physical processes: 1) epineutral downgradient tracer diffusion due to mesoscale eddies by means of an eddy diffusion coefficient K and 2) small-scale isotropic downgradient turbulent diffusion by means of a turbulent diffusion coefficient D. The isotropic nature of D is discussed in McDougall et al. (2014), but has previously been regarded to be diapycnal (Redi 1982; Griffies 2004; or vertical in the small-slope approximation). To represent epineutral diffusion in Cartesian coordinates, one makes use of a rotated tensor. The component of the diffusive tracer flux (C m^3 s^{-1}), through a general surface of constant φ, with surface area A_φ due to both eddy and turbulent diffusion, projected into Cartesian coordinates is given by

$$m_C^\varphi = -\int_{A_\varphi} \mathbf{K}\nabla C \cdot \frac{\nabla\varphi}{|\nabla\varphi|}\,dA. \quad (16)$$

This represents diffusion of any conserved tracer C through any surface φ. Here **K** is

$$\mathbf{K} = \frac{K}{1 + S^2}\begin{pmatrix} 1 + \epsilon + S_y^2 & -S_x S_y & S_x \\ -S_x S_y & 1 + \epsilon + S_x^2 & S_y \\ S_x & S_y & \epsilon + S^2 \end{pmatrix}, \quad (17)$$

where $\epsilon = D(1 + S^2)/K$, and

$$\mathbf{S} = (S_x, S_y) = -\left(\frac{\frac{\partial\gamma^n}{\partial x}}{\frac{\partial\gamma^n}{\partial z}}, \frac{\frac{\partial\gamma^n}{\partial y}}{\frac{\partial\gamma^n}{\partial z}}\right). \quad (18)$$

Using neutral density γ^n (McDougall 1987; Jackett and McDougall 1997), **S** provides the neutral direction, and $S^2 = S_x^2 + S_y^2$.

1) DIFFUSIVE HEAT FLUX

The downgradient diathermal transport of heat m_Θ^Θ, through a surface of constant Θ, between the S_A range of $[S_A \pm \delta S_A]$, can be expressed as

$$m_\Theta^\Theta = -\rho_0 c_p^0 \int_{S_A \pm \delta S_A | \Theta} \mathbf{K}\nabla\Theta \cdot \frac{\nabla\Theta}{|\nabla\Theta|} dA. \qquad (19)$$

As isotherms and isohalines are not necessarily orthogonal in Cartesian coordinates, there will also be a downgradient diahaline diffusive transport of heat ($m_\Theta^{S_A}$), through a surface of constant S_A, between the Θ range of $[\Theta \pm \delta\Theta]$, given by

$$m_\Theta^{S_A} = -\rho_0 c_p^0 \int_{\Theta \pm \delta\Theta | S_A} \mathbf{K}\nabla\Theta \cdot \frac{\nabla S_A}{|\nabla S_A|} dA \qquad (20)$$

Using Eqs. (19) and (20) we construct $\mathbf{m_\Theta} = (m_\Theta^{S_A}, m_\Theta^\Theta)$ and use this to calculate the net convergence of heat into ΔV given by

$$\overline{M}_\Theta = \frac{1}{\Delta t} \int_t^{t+\Delta t} (\nabla_{S_A\Theta} \cdot \mathbf{m}_\Theta)\, dt. \qquad (21)$$

2) DIFFUSIVE SALT FLUX

The convergence of salt into ΔV is given by

$$\overline{M}_{S_A} = \frac{1}{\Delta t} \int_t^{t+\Delta t} (\nabla_{S_A\Theta} \cdot \mathbf{m}_{S_A})\, dt, \qquad (22)$$

where $\mathbf{m_{S_A}} = (m_{S_A}^{S_A}, m_{S_A}^\Theta)$, and $m_{S_A}^{S_A}$ is the downgradient diahaline diffusive transport of salt, through a surface of constant S_A, between the Θ range of $[\Theta \pm \delta\Theta]$, given by

$$m_{S_A}^{S_A} = -\rho_0 \int_{\Theta \pm \delta\Theta | S_A} \mathbf{K}\nabla S_A \cdot \frac{\nabla S_A}{|\nabla S_A|} dA, \qquad (23)$$

and $m_{S_A}^\Theta$ is the downgradient diathermal diffusive transport of salt, through a surface of constant Θ, between the S_A range of $[S_A \pm \delta S_A]$, given by

$$m_{S_A}^\Theta = -\rho_0 \int_{S_A \pm \delta S_A | \Theta} \mathbf{K}\nabla S_A \cdot \frac{\nabla\Theta}{|\nabla\Theta|} dA. \qquad (24)$$

c. Local response

The local response is the last term on the rhs of Eqs. (11) and (12), given by $\overline{L}_{S_A} = \rho_0 \overline{\Delta V(\partial S_A/\partial t)}$ for salt and $\overline{L}_\Theta = \rho_0 c_p^0 \overline{\Delta V(\partial\Theta/\partial t)}$ for heat. The terms $\overline{L}_{S_A}$ and $\overline{L}_\Theta$ are the amount of salt and heat per unit time, needed to change the S_A and Θ properties of ΔV by an amount remaining within the defined (S_A, Θ) grid. As a result, the observed changes in salt and heat do not lead to a diathermohaline transport, but nonetheless reduces the amount of salt and heat available for water mass transformation. Applying Reynolds decomposition and averaging we obtain

$$\overline{L}_{S_A} = \rho_0 \overline{\Delta V' \frac{\partial S_A'}{\partial t}}, \quad \text{and} \qquad (25)$$

$$\overline{L}_\Theta = \rho_0 c_p^0 \overline{\Delta V' \frac{\partial\Theta'}{\partial t}}. \qquad (26)$$

Note that both S_A' and Θ' deviate from (S_A, Θ) only within the range defined by $S_A \pm \delta S_A$ and $\Theta \pm \delta\Theta$, respectively. Hence, S_A' and Θ' approach 0 when δS_A and $\delta\Theta$ approach 0.

4. The thermohaline inverse model

As we do not know the exact spatial and temporal distribution of K and D embedded in M_{S_A} and M_Θ in Eqs. (11) and (12), we formulate THIM, which uses an inverse technique that enables us to simultaneously estimate K, D, and $\Psi_{S_A\Theta}$.

We express the nondivergent component of $\overline{\mathbf{U}}_{S_A\Theta}^{\text{dia}}$ as a diathermohaline streamfunction difference in the S_A direction,

$$\begin{aligned}
\Psi_{S_A\Theta}^{\text{dia}}(S_A, \Theta + \delta\Theta) - \Psi_{S_A\Theta}^{\text{dia}}(S_A, \Theta - \delta\Theta) &= \int_t^{t+\Delta t} \int_{\Theta \pm \delta\Theta|_{S_A}} u_{S_A}^{\text{dia}} - u_{S_A}^{\text{tr}}\, dA\, dt \\
&= \overline{U}_{|S_A}^{\text{dia}}(S_A, \Theta \pm \delta\Theta) - \overline{U}_{|S_A}^{\text{tr}}(S_A, \Theta \pm \delta\Theta) \\
&= \frac{1}{\rho_0 \Delta S_A}(\overline{M}_{S_A} - \overline{L}_{S_A} - \overline{S_A F_m}) - \overline{U}_{|S_A}^{\text{tr}}(S_A, \Theta \pm \delta\Theta), \qquad (27)
\end{aligned}$$

and in the Θ direction,

$$-[\Psi_{S_A\Theta}^{\mathrm{dia}}(S_A + \delta S_A, \Theta) - \Psi_{S_A\Theta}^{\mathrm{dia}}(S_A - \delta S_A, \Theta)] = \int_t^{t+\Delta t}\int_{S_A \pm \delta S_A|_\Theta} u_\Theta^{\mathrm{dia}} - u_\Theta^{\mathrm{tr}}\, dA\, dt$$

$$= \overline{U}_{|\Theta}^{\mathrm{dia}}(S_A \pm \delta S_A, \Theta) - \overline{U}_{|\Theta}^{\mathrm{tr}}(S_A \pm \delta S_A, \Theta)$$

$$= \frac{1}{\rho_0 c_p^0 \Delta\Theta}(\overline{F_\Theta} + \overline{M_\Theta} - \overline{L_\Theta}) - \overline{U}_{|\Theta}^{\mathrm{tr}}(S_A \pm \delta S_A, \Theta). \qquad (28)$$

Here we used Eqs. (11) and (12) to obtain the second line of Eqs. (27) and (28), and

$$\overline{U}_{|S_A}^{\mathrm{tr}}(S_A, \Theta \pm \delta\Theta) = \int_t^{t+\Delta t}\int_{\Theta \pm \delta\Theta|_{S_A}} u_{S_A}^{\mathrm{tr}}\, dA\, dt \quad (29)$$

is the diahaline volume transport due to a trend in S_A (x, y, z, t), and

$$\overline{U}_{|\Theta}^{\mathrm{tr}}(S_A \pm \delta S_A, \Theta) = \int_t^{t+\Delta t}\int_{S_A \pm \delta S_A|_\Theta} u_\Theta^{\mathrm{tr}}\, dA\, dt \quad (30)$$

is the diathermal volume transport due to a trend in $\Theta(x, y, z, t)$. Groeskamp et al. (2014) showed that $\overline{U}_{|S_A}^{\mathrm{tr}}$ and $\overline{U}_{|\Theta}^{\mathrm{tr}}$ can be obtained from an ocean hydrography only.

For each ΔV, two unique equations can be constructed and combined in the form $\mathbf{A}\mathbf{x} = \mathbf{b}$. Here $\mathbf{x}$ is a $1 \times M$ vector of unknown $\Psi_{S_A\Theta}$ values and K and D coefficients, such that $M = (N_{S_A} + 1)(N_\Theta + 1) + N_K + N_D$. Here N_K and N_D are the number of unknown coefficients used to represent spatial and temporal variation of epineutral and turbulent diffusion and N_{S_A} and N_Θ are the number of ΔV in the S_A and Θ direction, respectively. The quantity $\mathbf{A}$ is a $N \times M$ matrix of their coefficients, with $N = 2N_{S_A}N_\Theta$, and $\mathbf{b}$ is a $1 \times N$ vector of the known forcing terms.

Both $\mathbf{A}$ and $\mathbf{b}$ are based on data that includes error, leading to unknown equation error. Hence, we have N unknown equation errors and M unknown variables, leading to $N + M$ unknowns and N equations. Linear dependencies may reduce the effective number of equations N, always resulting in an underdetermined set of equations with an infinite number of solutions. An estimate for $\mathbf{x}$ can be obtained using an inverse technique that minimizes χ^2, which is the sum of both the solution error and the equation error (Menke 1984; Wunsch 1996; McIntosh and Rintoul 1997):

$$\chi^2 = (\mathbf{x} - \mathbf{x}_0)^T \mathbf{W}_c^{-2}(\mathbf{x} - \mathbf{x}_0) + \mathbf{e}^T \mathbf{W}_r^2 \mathbf{e}. \qquad (31)$$

Rewriting the error as $\mathbf{e} = \mathbf{A}\mathbf{x} - \mathbf{b}$ and setting $\partial\chi^2/\partial\mathbf{x} = 0$, the solution can be written as

$$\mathbf{x} = \mathbf{x}_0 + \mathbf{W}_c^2 \mathbf{A}^T(\mathbf{A}\mathbf{W}_c^2\mathbf{A}^T + \mathbf{W}_r^{-2})^{-1}(\mathbf{b} - \mathbf{A}\mathbf{x}_0). \quad (32)$$

Here $\mathbf{x}_0$ is a prior estimate of $\mathbf{x}$, $\mathbf{W}_r$ is the row (equation) weighting matrix, and $\mathbf{W}_c$ is the column (variable) weighting matrix. If $\mathbf{x}$ and $\mathbf{e}$ are jointly normally distributed, the minimization of χ^2 is equivalent to finding the most probable solution of $\mathbf{x}$, with a standard deviation given by the square root of the diagonal of the posterior covariance matrix given by (Menke 1984)

$$\mathbf{C}_p = \mathbf{W}_c^2 - \mathbf{W}_c^2\mathbf{A}^T(\mathbf{A}\mathbf{W}_c^2\mathbf{A}^T + \mathbf{W}_r^{-2})^{-1}\mathbf{A}\mathbf{W}_c^2. \quad (33)$$

Although this solution may have a correct statistical interpretation and can guide the choices of the weights, it may not always be physically realistic, particularly if a good estimate of the covariance structure is not available. Therefore, we adopt a minimization process based on Eq. (31) that provides a solution with a physically realistic interpretation rather than a statistical interpretation. Let $\mathbf{W}_r$ and $\mathbf{W}_c$ be diagonal with elements $1/\sigma_{\mathbf{e}}$ and $\sigma_{\mathbf{x}}$, respectively, such that we can rewrite Eq. (31) as

$$\chi^2 = \chi_{\mathbf{x}}^2 + \chi_{\mathbf{e}}^2.$$

$$= \sum_{m=1}^M \frac{(x_m - x_{0,m})^2}{\sigma_{x_m}^2} + \sum_{n=1}^N \frac{e_n^2}{\sigma_{e_n}^2}. \qquad (34)$$

To allow for a similar influence of each variable and equation on the solution, we require all elements of χ^2, and therefore $\chi_{\mathbf{x}}^2$ and $\chi_{\mathbf{e}}^2$, to have a similar nondimensionalized order of magnitude. Hence, let $\mathbf{x}_0$ be our best estimate of $\mathbf{x}$, and let $\sigma_{\mathbf{x}}$ be our best estimate of the error between $\mathbf{x}_0$ and $\mathbf{x}$, such that $(x_m - x_{0,m})^2/\sigma_{x_m}^2$ become order one values and $\chi_{\mathbf{x}}^2 \approx M$. Equivalently, let $\sigma_{\mathbf{e}}$ be our best estimate of the equation error $\mathbf{e}$, such that $e_n^2/\sigma_{e_n}^2$ become order one values and $\chi_{\mathbf{e}}^2 \approx N$.

To illustrate why we want $\chi_{\mathbf{e}}^2 \approx N$ and $\chi_{\mathbf{x}}^2 \approx M$, imagine using large values of $\sigma_{\mathbf{e}}$ compared to $\mathbf{e}$, such that $\chi_{\mathbf{e}}^2 \ll N$. The solution then tends to $\mathbf{x} = \mathbf{x}_0$, by effectively minimizing $\chi_{\mathbf{x}}^2$, meaning that the equations have no information content. In contrast, when $\sigma_{\mathbf{x}}$ is large compared to $\mathbf{x} - \mathbf{x}_0$, then $\chi_{\mathbf{x}}^2 \ll M$, and the solution is obtained by minimizing $\chi_{\mathbf{e}}^2$. The solution then tends to satisfy $\mathbf{A}\mathbf{x} = \mathbf{b}$ as accurately as possible, regardless of how far $\mathbf{x}$ is from $\mathbf{x}_0$. To avoid fitting $\mathbf{x}$ toward either the equations or $\mathbf{x}_0$, we suggest that one should find

a physically realistic solution from a range of combinations for $\mathbf{x}_0$, $\sigma_\mathbf{x}$, and $\sigma_\mathbf{e}$, for which

$$\frac{N}{10} \leq \chi_\mathbf{e}^2 \leq 10N \quad \text{and} \quad \frac{M}{10} \leq \chi_\mathbf{x}^2 \leq 10M. \quad (35)$$

Note that, if prior statistics are not well known, the sensitivity of the solution to the choice of $\mathbf{x}_0$, $\sigma_\mathbf{x}$, and $\sigma_\mathbf{e}$ may be larger than the standard deviation for a particular solution obtained from $\mathbf{C}_p$. We refer to the combination of Eqs. (27) and (28) and the described inverse technique as the thermohaline inverse method.

5. The THIM applied to a numerical climate model

In this section, we apply the THIM to the hydrography and surface fluxes of an intermediate complexity numerical climate model's output, where the model's $\Psi_{S_A\Theta}^{dia}$, K, and D are known.

a. The University of Victoria Climate Model

We use the final 10 yr of a 3000-yr spinup simulation of the University of Victoria Climate Model (UVIC). This model is an intermediate complexity climate model with horizontal resolution of 1.8° latitude by 3.6° longitude grid spacing, 19 vertical levels, and a 2D energy balance atmosphere (Sijp et al. 2006; the case referred to as GM). The ocean model is the Geophysical Fluid Dynamics Laboratory Modular Ocean Model, version 2.2 (MOM2), using the Boussinesq approximation ($\rho \approx \rho_0 = 1035\,\text{kg}\,\text{m}^{-3}$) and a constant heat capacity ($c_p = 4000\,\text{J}\,\text{K}^{-1}\,\text{kg}^{-1}$) with the rigid-lid approximations applied, and the surface freshwater fluxes are modeled by way of an equivalent salt flux ($\text{kg}\,\text{m}^{-2}\,\text{s}^{-1}$) (Pacanowski 1996). The model conserves heat and salt by conserving potential temperature θ (°C) and Practical Salinity S_P. We use monthly averaged S_P and θ.

The vertical mixing coefficient increases with depth with a shape described by

$$D_\text{uvic} = 10^{-4}\{0.8 + \frac{1.05}{\pi}\arctan[4.5 \times 10^{-3}(z - 2500)]\}, \quad (36)$$

taking a value of $0.3 \times 10^{-4}\,\text{m}^2\,\text{s}^{-1}$ at the surface and increasing to $1.3 \times 10^{-4}\,\text{m}^2\,\text{s}^{-1}$ at the bottom. The model employs the eddy-induced advection parameterization of Gent et al. (1995) with a constant diffusion coefficient of $1000\,\text{m}^2\,\text{s}^{-1}$. Tracers are diffused in the isopycnal direction with a constant coefficient of $K_\text{uvic} = 1200\,\text{m}^2\,\text{s}^{-1}$. The model adopts epineutral diffusion everywhere and uses a slope maximum of $S_\text{max} = 1/100$; any slopes exceeding this limit are set to S_max.

b. Formulating the THIM for UVIC

At each (x, y, z) coordinate tracer location of UVIC, S_P and θ values are given, and one can define six interfaces that encloses a volume. To calculate diffusion, each interface is assumed to be a surface of constant S_P and θ. Adopting constant diffusion through a surface reduces the integral over the surface in Eq. (16) into a multiplication with the surface. Since the unit normal of the surfaces are exactly in the x, y, and z direction, we find the associated interfaces to be $A_{yz} = dydz$, $A_{xz} = dxdz$, and $A_{xy} = dxdy$. In the UVIC model the small-slope approximation has been applied, such that the convergence of salt and heat due to the sum of both epineutral and vertical turbulent downgradient diffusion is given by

$$M_{S_\text{P}} = \nabla_{S_\text{P}\theta} \cdot \left(\mathbf{K}_\text{small}\nabla S_\text{P}\mathbf{S} + D\frac{\partial S_\text{P}}{\partial z}A_{xy}\right), \quad \text{and} \quad (37)$$

$$M_\theta = \nabla_{S_\text{P}\theta} \cdot \left(\mathbf{K}_\text{small}\nabla\theta\mathbf{S} + D\frac{\partial\theta}{\partial z}A_{xy}\right). \quad (38)$$

Here $\nabla_{S_\text{P}\theta}$ is as Eq. (10), but for (S_P, θ) coordinates. We used the fact that $D/K \ll 1$, and $\mathbf{S}$ is

$$\mathbf{S} = \begin{pmatrix} A_{yz} & 0 & 0 \\ 0 & A_{xz} & 0 \\ 0 & 0 & A_{xy} \end{pmatrix}, \quad (39)$$

and $\mathbf{K}_\text{small}$ is

$$\mathbf{K}_\text{small} = K_\text{uvic}\begin{pmatrix} 1 & 0 & S_x \\ 0 & 1 & S_y \\ S_x & S_y & S^2 \end{pmatrix}. \quad (40)$$

For the inverse model we use one unknown for both K and D. To take into account the vertical structure of D, we multiply $(\partial S_\text{P}/\partial z A_{xy})|_z$ and $(\partial\theta/\partial z A_{xy})|_z$ by the vertical structure of D, given by $D_\text{uvic}(z) \times 10^4$. This reduces the variables for small-scale turbulent diffusion to a single parameter ($D = 1 \times 10^{-4}\,\text{m}^2\,\text{s}^{-1}$) and yet retains the vertical structure as given by Eq. (36). The isopycnal gradients are obtained by using the locally referenced potential density values (σ_n) to calculate density gradients. The gradients at $z = z_k$ are obtained by finding $\sigma_n = \rho(S_\text{P}, \theta, p_k)$ for the whole ocean according to McDougall et al. (2003), applied for all depth levels ranging from $k = 1$: N, where N is the total number of vertical layers.

To obtain the time-averaged net convergence of salt and heat in (S_P, θ) coordinates we (i) sum M_{S_P} and M_θ for each volume enclosed by the six interfaces in (S_P, θ)
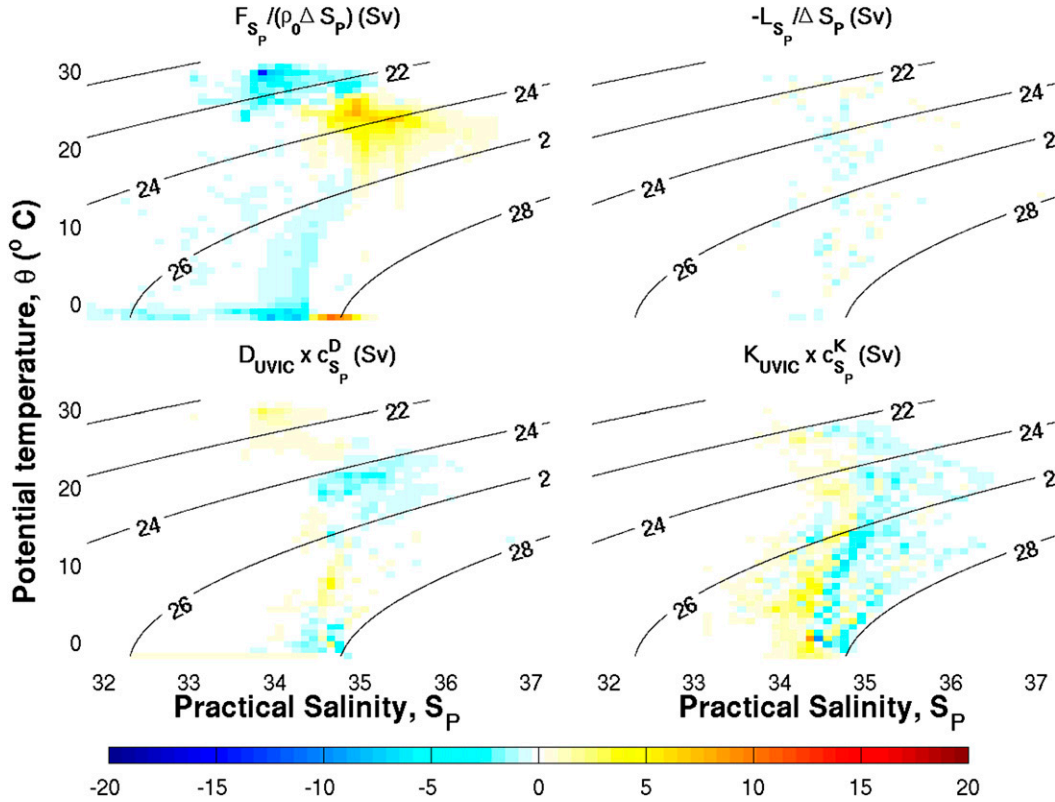
FIG. 2. The 10-yr-averaged, salt flux–induced, diahaline volume flux terms caused by (a) the (equivalent) surface salt flux $F_{S_P}/(\rho_0 \Delta S_P)$, (b) the local response term $-L_{S_P}/(\rho_0 \Delta S_P)$, (c) the turbulent diffusion term $D_{uvic}c_{S_P}^D$, and (d) isopycnal diffusion term $K_{uvic}c_{S_P}^K$. The grid sizes are $\Delta S_P = 0.1$ and $\Delta\theta = 0.75$. Black lines are contours of potential density (σ_0).

coordinates according to their tracer value on the (x, y, z) coordinate and then (ii) take the time average (Figs. 2, 3). We have chosen grid sizes in (S_P, θ) coordinates ($\Delta\theta = 0.75$ and $\Delta S_P = 0.1$) that distinguish the different water masses in the ocean's interior and provide an approximately equal number of equations in both S_P and θ directions.

The surface freshwater flux is provided as an equivalent salt flux, that is, $F_{S_P} = \overline{S_P F_m}$, as in Eq. (14), where we have neglected the prime–prime term. The term F_θ is calculated according to Eq. (15). The surface fluxes are gridded according to the surface grid S_P and θ values (Figs. 2, 3). The term $L_{S_P\theta}$ is calculated according to Eq. (25), and the time derivative is calculated using the values of two subsequent months and gridded according to S_P and θ of the first month (Figs. 2, 3). Using the above, (27) and (28) applied to UVIC are

$$\Psi_{S_P\theta}(S_P, \theta + \delta\theta) - \Psi_{S_P\theta}(S_P, \theta - \delta\theta)$$
$$- Kc_{S_P}^K - Dc_{S_P}^D = b_{S_P}, \tag{41}$$

and,

$$-[\Psi_{S_P\theta}(S_P + \delta S_P, \theta) - \Psi_{S_P\theta}(S_P - \delta S_P, \theta)]$$
$$- Kc_\theta^K - Dc_\theta^D = b_\theta. \tag{42}$$

Here $c_C^K = \overline{(K_{uvic}\Delta C)^{-1}\mathbf{V}_{S_P\theta} \cdot (\mathbf{K}_{small}\nabla C\mathbf{S})}$ and $c_C^D = \overline{(\Delta C)^{-1}\mathbf{V}_{S_P\theta} \cdot [(C_z A_{xy})_z D_{uvic}(z) \times 10^4]}$ contain the tracer–gradient divergence, with which the diffusion coefficients are multiplied, after replacing C with S_P and θ. Then $b_{S_P} = (\overline{F}_{S_P} - \rho_0\overline{L}_{S_P})/(\rho_0\Delta S_P) - \overline{U}_{S_P}^{tr}$ and $b_\theta = (\overline{F}_\theta - \rho_0 c_p^0 \overline{L}_\theta)/(\rho_0 c_p^0 \Delta\theta) - \overline{U}_\theta^{tr}$.

Writing Eqs. (41) and (42) for each grid leads to a set of equations that can be written in the form $\mathbf{Ax} = \mathbf{b}$. Here $\mathbf{x} = (\Psi_{S_P\theta}, K, D)$, and the coefficients with which we multiply $\mathbf{x}$ in $\mathbf{A}$ are 1 or -1 for $\Psi_{S_P\theta}, c_{S_P}^K$, and c_θ^K for K and $c_{S_P}^D$ and c_θ^D for D and $\mathbf{b} = (b_{S_P}, b_\theta)$.

c. The a priori constraints

To provide a physically realistic estimate of $\mathbf{x}$ using the THIM, we need to include boundary conditions and specify $\mathbf{x}_0$, $\sigma_{\mathbf{x}}$, and $\sigma_{\mathbf{e}}$. We have omitted equations for which both $|K_{uvic}c_{S_P}^K - D_{uvic}c_{S_P}^D|$ and $|b_{S_P}|$ are smaller than 0.1 Sverdrups (Sv; $1\,\mathrm{Sv} = 10^6\,\mathrm{m}^3\,\mathrm{s}^{-1}$), as these
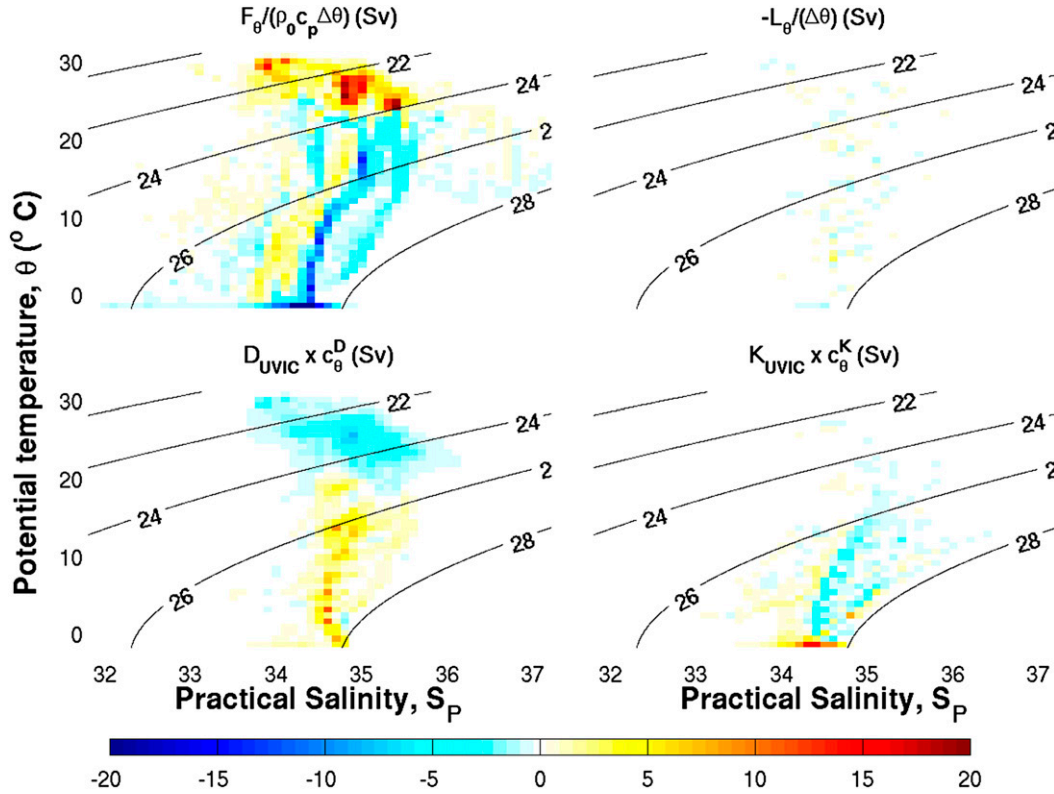
FIG. 3. As Fig. 2, but for the 10-yr-averaged, heat flux–induced, diathermal volume flux terms.

equations do not have a signal-to-noise ratio that adds information to the solution. We have also applied this to the heat equations. We have also imposed that transport into a ΔV that does not exist in the ocean is zero. That is, we have set $\Psi_{S_P\theta}^{\mathrm{dia}} = 0$ if two or more neighboring ΔV's do not exist in the ocean. The a priori expected magnitude for K and D are $x_0^K = K_{\mathrm{uvic}} = 1200\,\mathrm{m}^2\,\mathrm{s}^{-1}$ and $x_0^D = D_{\mathrm{uvic}} = 1 \times 10^{-4}\,\mathrm{m}^2\,\mathrm{s}^{-1}$. The a priori expected magnitudes for $\Psi_{S_P\theta}$ are unknown and therefore chosen to be $\mathbf{x}_0^\Psi = 0$. We obtain $N = 2709$ and $M = 1603$.

1) ROW WEIGHTING

Each equation represents a diathermohaline volume transport of a certain magnitude. An a priori estimate of the magnitude of the error for each equation is given by

$$\mathbf{e}_0 = |\mathbf{A}\mathbf{x}_0 - \mathbf{b}|. \tag{43}$$

Given that $\mathbf{x}_0^\Psi = 0$, this is equivalent to inserting $K = K_{\mathrm{uvic}}$ and $D = D_{\mathrm{uvic}}$ into the lhs of Eqs. (41) and (42) and taking the absolute value of the difference between both sides of the equations. Assuming that 1) the errors in the equations are proportional to, but not necessarily equal to, the size of the transports given by $\mathbf{e}_0$ and 2) that the salt and heat equations involve different physical processes and may therefore have a different

proportionality to $\mathbf{e}_0$, the expected errors of the diahaline and diathermal volume transport ($\sigma_{\mathbf{e}}^{S_P}$ and $\sigma_{\mathbf{e}}^\theta$, respectively) are then given by

$$\sigma_{\mathbf{e}}^{S_P} = \sigma_{\mathbf{e}}^{\min} + f_{S_P}\mathbf{e}_0^{S_P}, \tag{44}$$

$$\sigma_{\mathbf{e}}^{\boldsymbol{\theta}} = \sigma_{\mathbf{e}}^{\min} + f_\theta \mathbf{e}_0^{\boldsymbol{\theta}}.$$

Here $\mathbf{e}_0^{S_P}$ and $\mathbf{e}_0^\theta$ are the upper bounds of a priori estimates of error of the diahaline and diathermal volume transport, respectively [similar to Eq. (43)]. We have included a minimum value of $\sigma_{\mathbf{e}}^{\min} = 0.1\,\mathrm{Sv}$ to avoid terms becoming too small. We allow the proportionality factor for the salt and heat equations, f_{S_P} and f_θ, respectively, to vary between 0.01 and 1. This allows us to study the sensitivity of the solution to our choice of the equation error.

2) COLUMN WEIGHTING

We will allow for a standard error of 25% of the expected values for K and D. This leads to $\sigma_{\mathbf{x}}^K = 300\,\mathrm{m}^2\,\mathrm{s}^{-1}$ and $\sigma_{\mathbf{x}}^D = 2.5 \times 10^{-5}\,\mathrm{m}^2\,\mathrm{s}^{-1}$. The lack of information for $\mathbf{x}_0^\Psi$ can be compensated by a physically appropriate choice of $\sigma_{\mathbf{x}}^\Psi$. Assuming that $\mathbf{x} \approx \Psi_{S_P\theta}^{\mathrm{dia}}$, $\sigma_{\mathbf{x}}^\Psi$ should reflect $2|\Psi_{S_P\theta}^{\mathrm{dia}}|$ as this this takes into account the

structure of $|\Psi_{S_{\mathrm{P}}\theta}^{\mathrm{dia}}|$, while the factor 2 allows for an easy fit of $\Psi_{S_{\mathrm{P}}\theta}^{\mathrm{dia}}$ within the range allowed by $\sigma_{\mathbf{x}}^{\Psi}$.

To obtain an approximation of the structure of $|\Psi_{S_{\mathrm{P}}\theta}^{\mathrm{dia}}|$ without using prior knowledge of $\Psi_{S_{\mathrm{P}}\theta}^{\mathrm{dia}}$, we will assume that the size of $|\Psi_{S_{\mathrm{P}}\theta}^{\mathrm{dia}}|$ is proportional but not equal to the average of the magnitude of all diathermohaline volume transports [as given by Eq. (43)], in which a particular $|\Psi_{S_{\mathrm{P}}\theta}^{\mathrm{dia}}|$ is involved (maximal 4). This structure is then normalized with a maximum of 40 Sv to obtain $\sigma_{\mathbf{x}}^{\Psi} \approx 2|\Psi_{S_{\mathrm{P}}\theta}^{\mathrm{dia}}|$.

d. The solution

From the resulting range of solutions, we must choose the most physically realistic solution. We have chosen our solution to be the combination of $\sigma_{\mathbf{e}}^{S_{\mathrm{P}}}$ and $\sigma_{\mathbf{e}}^{\theta}$ at the ϵ_{rms} minimum within the range defined by Eq. (35). Here ϵ_{rms} is given by

$$\epsilon_{\mathrm{rms}} = \frac{\sqrt{\dfrac{1}{J}\displaystyle\sum_{j=1}^{J}(\Psi_{S_{\mathrm{P}}\theta,j}^{\mathrm{dia}} - \Psi_{S_{\mathrm{P}}\theta,j}^{\mathrm{dia,inv}})^2}}{\dfrac{1}{J}\displaystyle\sum_{j=1}^{J}|\Psi_{S_{\mathrm{P}}\theta,j}^{\mathrm{dia}}|}. \tag{45}$$

The term J is the number of unknown streamfunction variables, and ϵ_{rms} represents a weighted root-mean-square value of the difference between the diathermohaline streamfunction determined by the inverse method $\Psi_{S_{\mathrm{P}}\theta}^{\mathrm{dia,inv}}$ and the diathermohaline streamfunction calculated according to Groeskamp et al. (2014). Hence, if $\epsilon_{\mathrm{rms}} \geq 1$, our solution is no better than the solution given by our prior estimate $\Psi_{S_{\mathrm{P}}\theta}^{\mathrm{dia,inv}} = \mathbf{x}_0^{\Psi}$. If $\epsilon_{\mathrm{rms}} < 1$, the solution is more accurate than our prior estimate with a perfect solution (i.e., $\Psi_{S_{\mathrm{P}}\theta}^{\mathrm{dia,inv}} = \Psi_{S_{\mathrm{P}}\theta}^{\mathrm{dia}}$), if $\epsilon_{\mathrm{rms}} = 0$. The results of this solution are discussed in the next section.

6. Results and discussion

In this section, we discuss the skill of the THIM by comparing the UVIC model's variables with the inverse estimates. For a detailed physical interpretation of the circulation cells of $\Psi_{S_{\mathrm{P}}\theta}^{\mathrm{dia}}$, we refer to Zika et al. (2012), Döös et al. (2012), and Groeskamp et al. (2014).

a. The forcing terms

The surface salt flux binned in (S_{P}, θ) coordinates shows a diahaline transport in the direction of higher-salinity values for salty water and in the direction of lower-salinity values for freshwater (Fig. 2). A similar feature is observed for the surface heat flux binned in (S_{P}, θ) coordinates, which show diathermal transport in

the direction of higher temperatures for fluid parcels with high temperatures and in the direction of lower temperatures for fluid parcels with low temperatures (Fig. 3). Hence, both the surface salt and heat fluxes lead to divergence of volume in (S_{P}, θ) coordinates. The surface divergence is balanced by convergence of volume in (S_{P}, θ) coordinates due to both the eddy and turbulent diffusive transport terms for salt and heat (Figs. 2, 3). Note that the local term is very small compared to the surface and diffusive terms, and the trend term is statistically insignificant for this particular model within the 95% confidence level of the Student's t test (Groeskamp et al. 2014). Hence, the inverse method balances surface fluxes, mixing (of which the diffusion coefficients are estimated), and advection (represented by the diathermohaline streamfunction).

Sources of errors or variations in the inverse estimates, apart from weighting coefficients, are numerical diffusion and limits on the temporal and spatial resolution. The latter leads to averaging and rounding errors of the ocean's hydrography and surface fluxes and results in unresolved fluxes at the sea surface and unresolved flux divergence in the ocean interior. For example, unresolved fluxes with periods less than a month may occur because we have used monthly averaged values. Such fluxes are expected to have the largest influence on circulations that occur near the surface, as heat and salt fluxes are expected to vary at the surface on time scales shorter than a month. The numerical diffusion and unresolved fluxes lead to $\mathbf{V}_{S_{\mathrm{P}}\theta} \cdot (\overline{\mathbf{U}}_{S_{\mathrm{P}}\theta}^{\mathrm{dia}} - \overline{\mathbf{U}}_{S_{\mathrm{P}}\theta}^{\mathrm{tr}}) \neq 0$ (Fig. 4), that is, an imbalance between the diathermohaline circulation and fluxes of salt and heat. The larger this imbalance, the less accurate our inverse solution will be. Another result of this imbalance is that even when we know the exact spatial and temporal structure of the diffusion coefficents (K and D), the numerical noise will not allow us to calculate a streamfunction directly from Eq. (4). Hence, to obtain $\Psi_{S_{\mathrm{P}}\theta}$ from salt and heat fluxes will always require an (inverse) estimate.

b. The solution range

We first discuss the range of solutions and then discuss the results for the optimal solution, indicated by a black dot (Fig. 5). When $f_{S_{\mathrm{P}}}$ and f_θ (and therefore $\sigma_{\mathbf{e}}^{S_{\mathrm{P}}}$ and $\sigma_{\mathbf{e}}^{\theta}$, respectively) increase, $\chi_{\mathbf{e}}^2/N$ decreases as expected (Fig. 5c). Simultaneously the solution is pushed more toward $\mathbf{x} = \mathbf{x}_0$, and because $\mathbf{x}_0^{\Psi} = 0$, this leads to a decrease in $\chi_{\mathbf{x}}^2/M$ (Fig. 5c). For the optimal solution we find that $\chi_{\mathbf{x}}^2/M = 0.69$ and $\chi_{\mathbf{x}}^2/N = 9.94$; hence, the solution is obtained mainly satisfying the equations by using a relative small values for $\sigma_{\mathbf{e}}^{S_{\mathrm{P}}}$ and $\sigma_{\mathbf{e}}^{\theta}$ compared to $\mathbf{e}$. This is expected because we have limited knowledge of $\mathbf{x}_0^{\Psi}$.
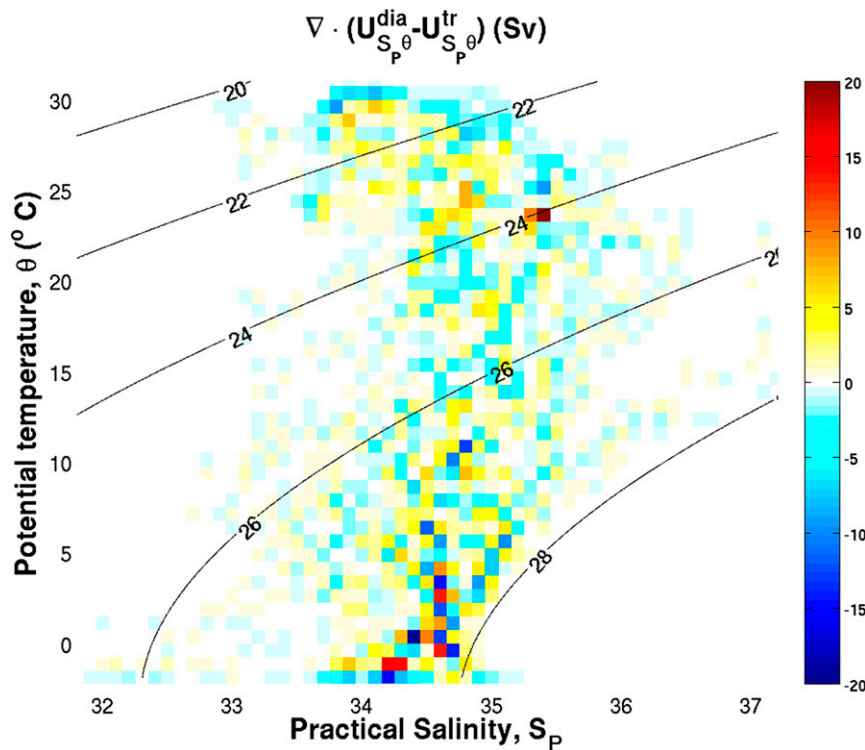
FIG. 4. Shows that $\mathbf{V}_{S_P\theta} \cdot (\overline{\mathbf{U}}_{S_P\theta}^{\text{dia}} - \overline{\mathbf{U}}_{S_P\theta}^{\text{tr}}) \neq 0$. Here $\overline{\mathbf{U}}_{S_P\theta}^{\text{dia}}$ is calculated using $K = K_{\text{uvic}}$ and $D = D_{\text{uvic}}$. The black lines are the contours for potential density (σ_0).

As $\mathbf{x}_0^{\Psi}$ is not known and the forcing terms are not in perfect balance in this model when using $K = K_{\text{uvic}}$ and $D = D_{\text{uvic}}$ (Figs. 2, 3), one will never obtain $\Psi_{S_P\theta}^{\text{dia,inv}} = \Psi_{S_P\theta}^{\text{dia}}$ and simultaneously obtain $K = K_{\text{uvic}}$ and $D = D_{\text{uvic}}$. As a result, an improved estimate of $\Psi_{S_P\theta}^{\text{dia,inv}}$ can only be obtained by satisfying the equations, with less emphasis on fitting to $\mathbf{x} = \mathbf{x}_0$, and thus moving away from $K = K_{\text{uvic}}$ and $D = D_{\text{uvic}}$, but toward $\Psi_{S_P\theta}^{\text{dia,inv}} = \Psi_{S_P\theta}^{\text{dia}}$. This is particularly the case for K. This is not due to an incorrect formulation of the inverse method, it is a problem imbedded in the model monthly means. The inverse solution behaves as expected, and for the whole range of solutions, both K and D are very reasonable approximations of K_{uvic} and D_{uvic} (Fig. 5c). This shows that the THIM is skilled in providing estimates of $\Psi_{S_P\theta}^{\text{dia}}$, K, and D.

The optimal solution selected according to section 5d gives $f_{S_P} = 0.02$ and $f_\theta = 0.24$, that is, errors with a magnitude of 2% and 24%, for the conservation of salt and heat, respectively. This suggests that most of the error in the equations is created by the models heat flux terms.

c. The inverse estimate

For the optimal solution, $\Psi_{S_P\theta}^{\text{dia,inv}}$ and $\Psi_{S_P\theta}^{\text{dia}}$ are very similar (Fig. 6), with $\epsilon_{\text{rms}} = 0.56$. The standard deviation

of $\Psi_{S_P\theta}^{\text{dia,inv}}$ obtained from $\mathbf{C}_p$ does not exceed 0.5 Sv (and is generally much smaller) and is very small compared to $|\Psi_{S_P\theta}^{\text{dia,inv}}|$ (not shown).

The difference $\Psi_{S_P\theta}^{\text{dia}} - \Psi_{S_P\theta}^{\text{dia,inv}}$ is generally about 1 Sv, with the exception of some areas in (S_P, θ) coordinates (Fig. 7). To explain the differences, we use that the dynamics of the area in (S_P, θ) for which $\theta > 20°C$ are dominated by processes that occur near the surface, while the area for $\theta < 10°$ and $S_P = 35 \pm 0.5\,\text{g\,kg}^{-1}$ are also strongly influenced by processes that occur in the ocean interior (Zika et al. 2012; Döös et al. 2012; Groeskamp et al. 2014; Hieronymus et al. 2014). The magnitudes of the turbulent diffusion terms, which can be scaled by changing D, are large throughout the whole S_P and θ domain (D terms in Figs. 2 and 3). The magnitude of the epineutral eddy diffusion terms, which can be scaled by changing K, is large only in the ocean interior (K terms in Figs. 2 and 3). As a result, the large surface fluxes of salt and especially heat of the area in (S_P, θ), for which $\theta > 20°$, can only be balanced by increasing the turbulent diffusion, increasing D. This will also result in an increased effect of turbulent diffusion in the oceans interior. The only way to compensate for this is by reducing the magnitude of the epineutral eddy diffusion, reducing K. This idea is supported by the
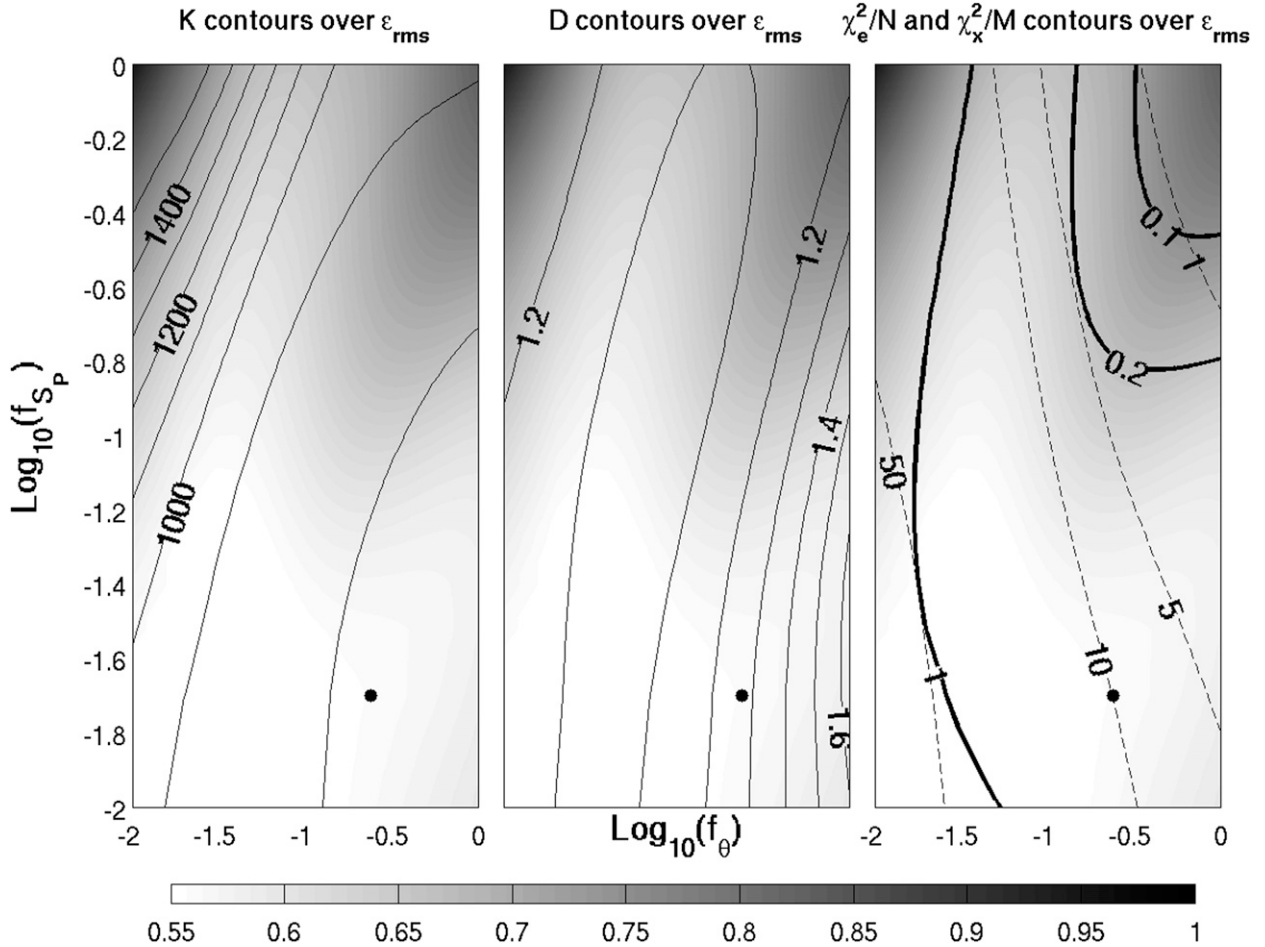
FIG. 5. Inverse estimates of (left) K and (middle) D and the (right) χ_e^2/N (dashed) and χ_x^2/M (solid) values, plotted on top of the associated ϵ_{rms} values (background color) for a range of the $\log_{10}$ of the row weighting values. With the variation of the weighting factor operating on the conservation of heat, given by f_θ, on the x axis and the variation of the weighting factor operating on the conservation of salt, given by f_{S_P}, on the y axis. Both f_θ and f_{S_P} range from 0.01 to 1. The lower the value of ϵ_{rms} (whiter), the closer the estimate of $\Psi_{S_P\theta}^{\text{dia,inv}}$ resembles $\Psi_{S_P\theta}^{\text{dia}}$. The black dot indicates the optimal solution defined by the minimum ϵ_{rms} value within the range given by Eq. (35).

estimates of the diffusion coefficients. The estimate of K for the optimal solution is $K = 929 \pm 7\,\text{m}^2\,\text{s}^{-1}$, which is close to $K_{\text{uvic}} = 1200\,\text{m}^2\,\text{s}^{-1}$, but an underestimation. The estimate of D is $D = 1.27 \pm 0.01 \times 10^{-4}\,\text{m}^2\,\text{s}^{-1}$, which is close to $D_{\text{uvic}} = 1 \times 10^{-4}\,\text{m}^2\,\text{s}^{-1}$, but an overestimation. We note that numerical diffusion may also increase the total diffusion in the model, possibly contributing to the fact that $D > D_{\text{uvic}}$. As a result of an increased D and reduced K, we then find general differences given by $\Psi_{S_P\theta}^{\text{dia}} - \Psi_{S_P\theta}^{\text{dia,inv}}$ (Fig. 7).

The fact that the standard deviation for $\mathbf{x}$ obtained from $\mathbf{C}_p$ is much smaller than the variation of the solution as a result of changing $\sigma_e^{\mathbf{S}_P}$ and σ_e^θ (Fig. 5) indicates that the accuracy of the solution is limited by our prior knowledge of the formal statistical structure of the equation and solution errors, rather than the information

content of the equations. This justifies our decision to choose the weights based on physical principles.

The similarity between $\Psi_{S_P\theta}^{\text{dia}}$ and $\Psi_{S_P\theta}^{\text{dia,inv}}$ is also evident when considering the diapycnal salt $[F_{\text{Salt}}(\rho_r)]$ and heat $[F_{\text{Heat}}(\rho_r)]$ transports for reference potential density ρ_r. Following (Zika et al. 2012) this is given by

$$F_{\text{Heat}}(\rho_r) = \int_{\rho_r} \rho c_p^0 \theta \, d\Psi_{S_P\theta}^{\text{dia}}, \qquad (46)$$

$$F_{\text{Salt}}(\rho_r) = \int_{\rho_r} S_P \, d\Psi_{S_P\theta}^{\text{dia}}.$$

We have applied an integration along lines of constant reference potential density ρ_r. The salt flux can be expressed as an equivalent freshwater flux by dividing it by
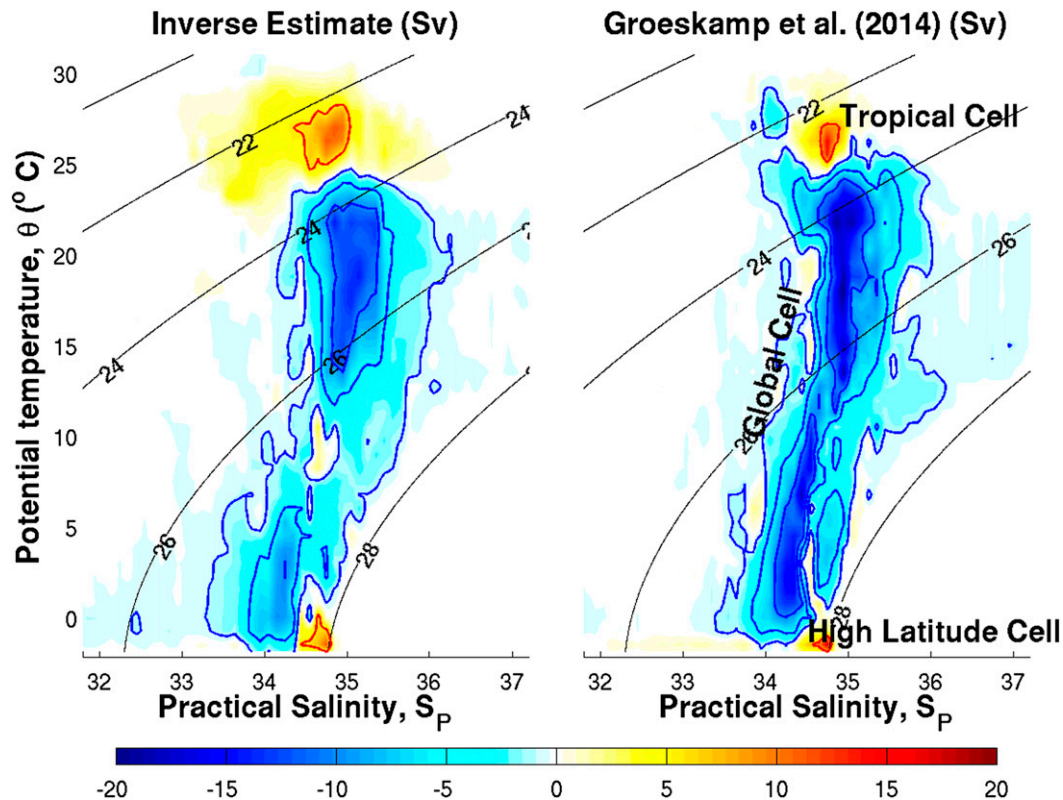
FIG. 6. (left) The inverse estimate $\Psi_{S_P\theta}^{\mathrm{dia,inv}}$ for the optimal solution, corresponding to the column weights given by the black dot in Fig. 5, and (right) $\Psi_{S_P\theta}^{\mathrm{dia}}$ calculated using the Groeskamp et al. (2014) method. The blue lines show the $[-15, -10, -1]$ Sv contours, the red lines show the 1-Sv contour, and the black lines show the contours for potential density σ_0. The cell names are adopted from Groeskamp et al. (2014).

a reference salinity $[F_{\mathrm{FW}}(\rho_r) = F_{\mathrm{Salt}}(\rho_r)/S_{\mathrm{ref}}, S_{\mathrm{ref}} = 35]$. At densities that cross the tropical cell $21\,\mathrm{kg\,m}^{-3} < \sigma_0 < 24\,\mathrm{kg\,m}^{-3}$, we have an overestimation of both the diapycnal freshwater and heat transport due to an increased magnitude of the tropical cell of $\Psi_{S_P\theta}^{\mathrm{dia,inv}}$ compared to $\Psi_{S_P\theta}^{\mathrm{dia}}$ (Fig. 8). Because of an underestimation of the global cell of $\Psi_{S_P\theta}^{\mathrm{dia,inv}}$ compared to $\Psi_{S_P\theta}^{\mathrm{dia}}$, we have a general underestimation of the magnitude of the diapycnal salt and heat transport for $24\,\mathrm{kg\,m}^{-3} < \sigma_0 < 27\,\mathrm{kg\,m}^{-3}$ (Fig. 8). However, the similarity of the shapes of the diapycnal transports show the THIM is skilled in capturing $\Psi_{S_P\theta}^{\mathrm{dia,inv}}$, regardless of the errors discussed previously.

d. Putting the THIM in perspective

Most inverse (box) models are designed with a focus on estimating the absolute velocity vector only. Currently inverse methods are one of few methods by which one is able to provide an estimate of mixing from observations. Inverse box methods that also estimate D often contain unknowns at the boundaries that require both dynamical constrains and conservation statements,

increasing the complexity of the system and sensitivity to prior estimates [Sloyan and Rintoul (2000, 2001), among many others]. The THIM uses boxes bounded using two pairs of tracer surfaces, analyzed in tracer coordinates, rather than Cartesian coordinates. This reduces the complexity of the system to a set of simple tracer conservation equations. The global application of the THIM leads to strong constrains on the solution, as confirmed by the small error calculated using $\mathbf{C}_P$ and small variation of the solution for a wide range of choices of the row weighting. There are inverse models that provide global estimates of D (Ganachaud and Wunsch 2000) or local estimates of both K and D (Zika et al. 2010b). However, to our knowledge the THIM is the only inverse estimates that can provide globally constrained estimates of both K and D from observations.

Groeskamp et al. (2014) showed that $\Psi_{S_A\Theta}^{\mathrm{dia}} = \Psi_{S_A\Theta}^{\mathrm{adv}} + \Psi_{S_A\Theta}^{\mathrm{loc}}$. Here the local thermohaline streamfunction $\Psi_{S_A\Theta}^{\mathrm{loc}}$ represents the circulation in (S_A, Θ) coordinates due to cyclic changes of the ocean's volume distribution in (S_A, Θ) coordinates, without motion in geographical space and can be calculated from an ocean
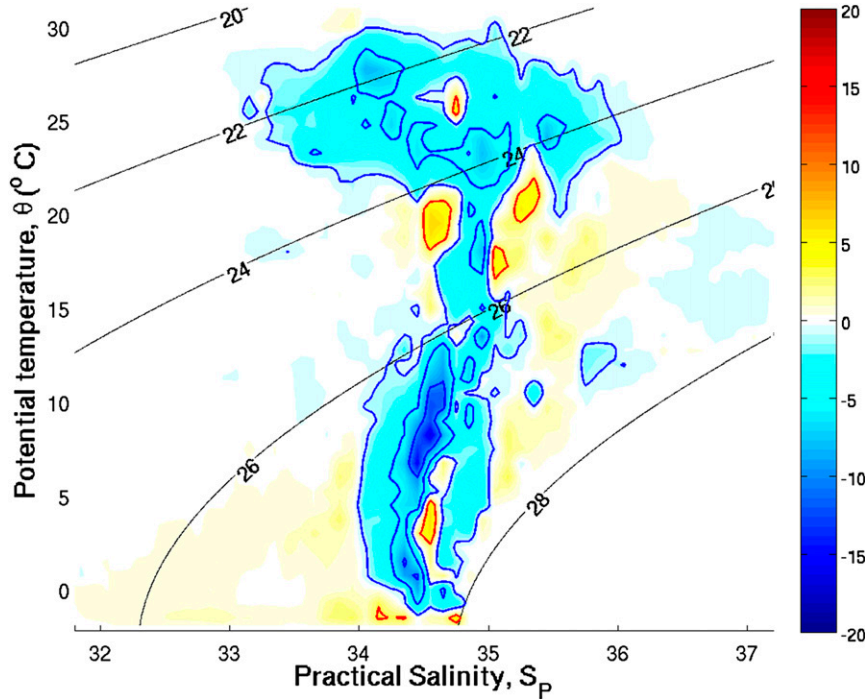
FIG. 7. The difference $\Psi_{S_\mathrm{P}\theta}^\mathrm{dia} - \Psi_{S_\mathrm{P}\theta}^\mathrm{dia,inv}$. Contours are as in Fig. 6.

hydrography. The advective thermohaline stream-function $\Psi_{S_\mathrm{A}\Theta}^\mathrm{adv}$ represents the nondivergent component of the geographical ocean circulation in the direction normal to the isohalines and isotherms, in (S_A, Θ) coordinates, and requires global observations of **u**. However, using the THIM and the calculation of $\Psi_{S_\mathrm{A}\Theta}^\mathrm{loc}$, as presented by Groeskamp et al. (2014), we can calculate $\Psi_{S_\mathrm{A}\Theta}^\mathrm{adv} = \Psi_{S_\mathrm{A}\Theta}^\mathrm{dia} - \Psi_{S_\mathrm{A}\Theta}^\mathrm{loc}$, from the more readily observed salinity and temperature.

This study has shown that premultiplying the turbulent diffusive terms with a structure function is an appropriate method to reduce the number of unknown diffusion coefficients, while allowing for its spatial variation. When applying the THIM to observations, such structure functions can be applied to reduce the number of unknowns required to capture the spatial and temporal variation of K and D. When the THIM is applied to observations, choosing S_A and Θ as tracer coordinates utilizes the extensive observational coverage of these tracers, reducing uncertainties in the solution. We therefore believe that the THIM has the potential to obtain well constrained global estimates of spatially- and temporally-varying values of K and D.

7. Conclusions

We have presented the thermohaline inverse method (THIM), which estimates the diathermohaline streamfunction $\Psi_{S_\mathrm{A}\Theta}^\mathrm{dia}$, epineutral eddy K, and isotropic turbulent D diffusion coefficients. The THIM uses a balance between advection and water mass transformation due to thermohaline forcing in (S_A, Θ) coordinates. The thermohaline forcing, that is, the surface freshwater and heat fluxes and diffusive salt and heat fluxes, can be obtained from ocean hydrography and surface flux products.

We have tested the THIM using a model's hydrography and surface fluxes and compared the inverse estimate of K, D, and $\Psi_{S_\mathrm{A}\Theta}^\mathrm{dia}$ to the model's imbedded diffusion coefficients and diathermohaline streamfunction. The latter was calculated independently as described in Groeskamp et al. (2014), using the model's velocity output and hydrography. The results showed that the THIM is skilled in estimating K, D, and $\Psi_{S_\mathrm{A}\Theta}^\mathrm{dia}$ and that the differences between the inverse estimate and the model are explained by an imbalance of the model's boundary fluxes and diffusive fluxes of both salt and heat and not due to an incorrect formulation of the inverse method. There is large information content in the system, and the resulting solution shows only little sensitivity to a wide range of row weighting coefficients. From this we conclude that the THIM can be applied to observationally based products to produce well-constrained observationally based estimates of the ocean's diathermohaline circulation and mixing.
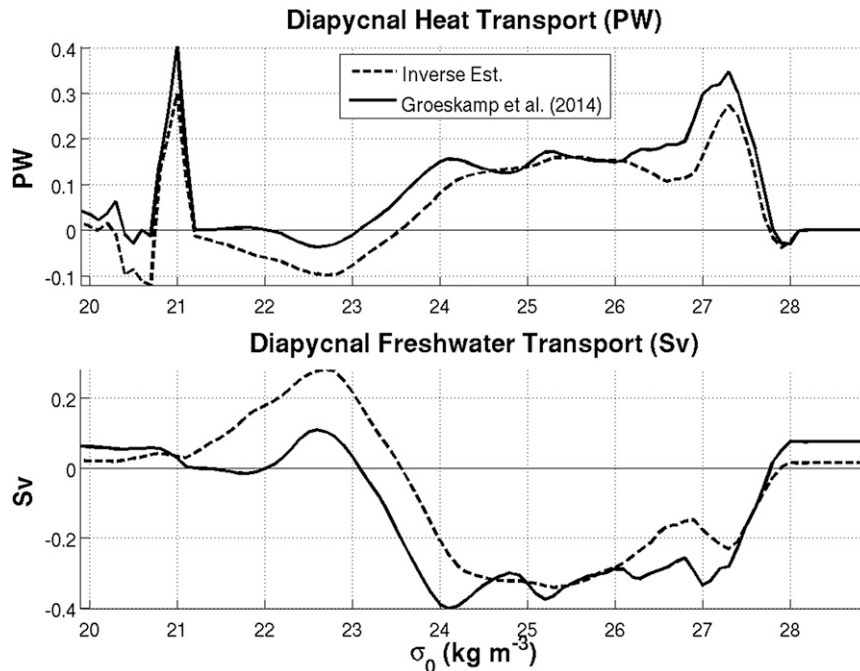
FIG. 8. Diapycnal transport of (top) heat (1 PW = 10^{15} W) and (bottom) freshwater through potential density surfaces (σ_0) for $\Psi_{S_P\theta}^{\text{dia}}$ (solid) and $\Psi_{S_P\theta}^{\text{dia,inv}}$ (dashed).

REFERENCES

Bryan, K., J. L. Sarmiento, and B. Saltzman, 1985: Modeling ocean circulation. *Advances in Geophysics,* Vol. 28, Academic Press, 433–459, doi:10.1016/S0065-2687(08)60232-0.

Cunningham, S., 2000: Circulation and volume flux of the North Atlantic using synoptic hydrographic data in a Bernoulli inverse. *J. Mar. Res.,* **58,** 1–35, doi:10.1357/002224000321511188.

Döös, K., and D. J. Webb, 1994: The Deacon cell and the other meridional cells of the Southern Ocean. *J. Phys. Oceanogr.,* **24,** 429–442, doi:10.1175/1520-0485(1994)024<0429:TDCATO>2.0.CO;2.

——, J. Nilsson, J. Nycander, L. Brodeau, and M. Ballarotta, 2012: The World Ocean thermohaline circulation. *J. Phys. Oceanogr.,* **42,** 1445–1460, doi:10.1175/JPO-D-11-0163.1.

Ferrari, R., and D. Ferreira, 2011: What processes drive the ocean heat transport? *Ocean Modell.,* **38,** 171–186, doi:10.1016/j.ocemod.2011.02.013.

Ganachaud, A., and C. Wunsch, 2000: Improved estimates of global ocean circulation, heat transport and mixing from hydrographic data. *Nature,* **408,** 453–457, doi:10.1038/35044048.

Gent, P. R., J. Willebrand, T. J. McDougall, and J. C. McWilliams, 1995: Parameterizing eddy-induced tracer transports in ocean circulation models. *J. Phys. Oceanogr.,* **25,** 463–474, doi:10.1175/1520-0485(1995)025<0463:PEITTI>2.0.CO;2.

Graham, F. S., and T. J. McDougall, 2013: Quantifying the nonconservative production of Conservative Temperature, potential temperature, and entropy. *J. Phys. Oceanogr.,* **43,** 838–862, doi:10.1175/JPO-D-11-0188.1.

Griffies, S. M., 2004: *Fundamentals of Ocean Climate Models.* Princeton University Press, 518 pp.

Groeskamp, S., J. D. Zika, T. J. McDougall, B. M. Sloyan, and F. Laliberté, 2014: The representation of ocean circulation and variability in thermodynamic coordinates. *J. Phys. Oceanogr.,* **44,** 1735–1750, doi:10.1175/JPO-D-13-0213.1.

Hieronymus, M., J. Nilsson, and J. Nycander, 2014: Water mass transformation in salinity–temperature space. *J. Phys. Oceanogr.,* **44,** 2547–2568, doi:10.1175-JPO-D-13-0257.1.

Hirst, A. C., and T. J. McDougall, 1998: Meridional overturning and dianeutral transport in a z-coordinate ocean model including eddy-induced advection. *J. Phys. Oceanogr.,* **28,** 1205–1223, doi:10.1175/1520-0485(1998)028<1205:MOADTI>2.0.CO;2.

——, D. R. Jackett, and T. J. McDougall, 1996: The meridional overturning cells of a world ocean model in neutral density coordinates. *J. Phys. Oceanogr.,* **26,** 775–791, doi:10.1175/1520-0485(1996)026<0775:TMOCOA>2.0.CO;2.

Huang, R. X., 1993: Real freshwater flux as a natural boundary condition for the salinity balance and thermohaline circulation forced by evaporation and precipitation. *J. Phys.*

Oceanogr., 23, 2428–2446, doi:10.1175/1520-0485(1993)023<2428: RFFAAN>2.0.CO;2.

IOC, SCOR, and IAPSO, 2010: The International Thermodynamic Equation of Seawater—2010: Calculation and use of thermodynamic properties. Intergovernmental Oceanographic Commission, Manuals and Guides 56, 220 pp. [Available online at http://www.teos-10.org/pubs/TEOS-10_Manual.pdf.]

Iudicone, D., G. Madec, and T. J. McDougall, 2008: Water-mass transformations in a neutral density framework and the key role of light penetration. *J. Phys. Oceanogr., 38,* 1357–1376, doi:10.1175/2007JPO3464.1.

Jackett, D. R., and T. J. McDougall, 1997: A neutral density variable for the world's oceans. *J. Phys. Oceanogr., 27,* 237–263, doi:10.1175/1520-0485(1997)027<0237:ANDVFT>2.0.CO;2.

Killworth, P. D., 1986: A Bernoulli inverse method for determining the ocean circulation. *J. Phys. Oceanogr., 16,* 2031–2051, doi:10.1175/1520-0485(1986)016<2031:ABIMFD>2.0.CO;2.

Lumpkin, R., and K. Speer, 2007: Global ocean meridional overturning. *J. Phys. Oceanogr., 37,* 2550–2562, doi:10.1175/JPO3130.1.

McDougall, T. J., 1987: Neutral surfaces. *J. Phys. Oceanogr., 17,* 1950–1964, doi:10.1175/1520-0485(1987)017<1950:NS>2.0.CO;2.

——, 2003: Potential enthalpy: A conservative oceanic variable for evaluating heat content and heat fluxes. *J. Phys. Oceanogr., 33,* 945–963, doi:10.1175/1520-0485(2003)033<0945: PEACOV>2.0.CO;2.

——, D. R. Jackett, D. G. Wright, and R. Feistel, 2003: Accurate and computationally efficient algorithms for potential temperature and density of seawater. *J. Atmos. Oceanic Technol., 20,* 730–741, doi:10.1175/1520-0426(2003)20<730:AACEAF>2.0.CO;2.

——, ——, F. J. Millero, R. Pawlowicz, and P. M. Barker, 2012: A global algorithm for estimating Absolute Salinity. *Ocean Sci., 8,* 1123–1134.

——, S. Groeskamp, and S. M. Griffies, 2014: On geometrical aspects of interior ocean mixing. *J. Phys. Oceanogr., 44,* 2164–2175, doi:10.1175/JPO-D-13-0270.1.

McIntosh, P. C., and S. R. Rintoul, 1997: Do box inverse models work? *J. Phys. Oceanogr., 27,* 291–308, doi:10.1175/1520-0485(1997)027<0291:DBIMW>2.0.CO;2.

Menke, W., 1984: *Geophysical Data Analysis: Discrete Inverse Theory.* Academic Press, 260 pp.

Millero, F. J., R. Feistel, D. G. Wright, and T. J. McDougall, 2008: The composition of standard seawater and the definition of the reference-composition salinity scale. *Deep-Sea Res. I, 55,* 50–72, doi:10.1016/j.dsr.2007.10.001.

Munk, W. H., 1966: Abyssal recipes. *Deep Sea Res. Oceanogr. Abstr., 13,* 707–730, doi:10.1016/0011-7471(66)90602-4.

Nurser, A. J. G., and M.-M. Lee, 2004: Isopycnal averaging at constant height. Part I: The formulation and a case study. *J. Phys. Oceanogr., 34,* 2721–2739, doi:10.1175/JPO2649.1.

Nycander, J., J. Nilsson, K. Döös, and G. Broström, 2007: Thermodynamic analysis of ocean circulation. *J. Phys. Oceanogr., 37,* 2038–2052, doi:10.1175/JPO3113.1.

Pacanowski, R., 1996: MOM2 version 2 documentation, user's guide, and reference manual. NOAA/Geophysical Fluid Dynamics Laboratory Ocean Tech. Rep. 3.2, 350 pp. [Available online at http://gfdl.noaa.gov/cms-filesystem-action/model_development/ocean/manual2.2.pdf.]

Redi, M. H., 1982: Oceanic isopycnal mixing by coordinate rotation. *J. Phys. Oceanogr., 12,* 1154–1158, doi:10.1175/1520-0485(1982)012<1154:OIMBCR>2.0.CO;2.

Schott, F., and H. Stommel, 1978: Beta spirals and absolute velocities in different oceans. *Deep-Sea Res., 25,* 961–1010, doi:10.1016/0146-6291(78)90583-0.

Sijp, W. P., M. Bates, and M. H. England, 2006: Can isopycnal mixing control the stability of the thermohaline circulation in ocean climate models? *J. Climate, 19,* 5637–5651, doi:10.1175/JCLI3890.1.

Sloyan, B. M., and S. R. Rintoul, 2000: Estimates of area-averaged diapycnal fluxes from basin-scale budgets. *J. Phys. Oceanogr., 30,* 2320–2341, doi:10.1175/1520-0485(2000)030<2320: EOAADF>2.0.CO;2.

——, and ——, 2001: The Southern Ocean limb of the global deep overturning circulation. *J. Phys. Oceanogr., 31,* 143–173, doi:10.1175/1520-0485(2001)031<0143:TSOLOT>2.0.CO;2.

Speer, K. G., 1993: Conversion among North Atlantic surface water types. *Tellus, 45A,* 72–79, doi:10.1034/j.1600-0870.1993.00006.x.

——, and E. Tziperman, 1992: Rates of water mass formation in the North Atlantic Ocean. *J. Phys. Oceanogr., 22,* 93–104, doi:10.1175/1520-0485(1992)022<0093:ROWMFI>2.0.CO;2.

Stommel, H., and F. Schott, 1977: The beta spiral and the determination of the absolute velocity field from hydrographic station data. *Deep-Sea Res., 24,* 325–329, doi:10.1016/0146-6291(77)93000-4.

Walin, G., 1982: On the relation between sea-surface heat flow and thermal circulation in the ocean. *Tellus, 34,* 187–195, doi:10.1111/j.2153-3490.1982.tb01806.x.

Wunsch, C., 1978: The North Atlantic general circulation west of 50°W determined by inverse methods. *Rev. Geophys., 16,* 583–620, doi:10.1029/RG016i004p00583.

——, 1996: *The Ocean Circulation Inverse Problem.* Cambridge University Press, 442 pp.

Zhang, H.-M., and N. G. Hogg, 1992: Circulation and water mass balance in the Brazil basin. *J. Mar. Res., 50,* 385–420, doi:10.1357/002224092784797629.

Zika, J. D., T. J. McDougall, and B. M. Sloyan, 2010a: A tracer-contour inverse method for estimating ocean circulation and mixing. *J. Phys. Oceanogr., 40,* 26–47, doi:10.1175/2009JPO4208.1.

——, ——, and ——, 2010b: Weak mixing in the eastern North Atlantic: An application of the tracer-contour inverse method. *J. Phys. Oceanogr., 40,* 1881–1893, doi:10.1175/2010JPO4360.1.

——, M. H. England, and W. P. Sijp, 2012: The ocean circulation in thermohaline coordinates. *J. Phys. Oceanogr., 42,* 708–724, doi:10.1175/JPO-D-11-0139.1.