The DOI for this manuscript is doi: 10.1175/JTECH-D-17-0198.1

The final published version of this manuscript will replace the preliminary version at the above DOI once it is available.

If you would like to cite this EOR in a separate work, please use the following full citation:

Mackay, N., C. Wilson, J. Zika, and N. Holliday, 2018: A Regional Thermohaline Inverse Method for Estimating Circulation and Mixing in the Arctic and subpolar North Atlantic. J. Atmos. Oceanic Technol. doi:10.1175/JTECH-D-17-0198.1, in press.

A Regional Thermohaline Inverse Method for Estimating Circulation and Mixing in the Arctic and subpolar North Atlantic

Neill Mackay*, Chris Wilson

National Oceanography Centre, Joseph Proudman Building, 6 Brownlow Street, Liverpool, L3 5DA, United Kingdom

Jan Zika

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

N. Penny Holliday

National Oceanography Centre, European Way, Southampton, SO14 3ZH, United Kingdom

**Corresponding author address:* National Oceanography Centre, Joseph Proudman Building, 6 Brownlow Street, Liverpool, L3 5DA, United Kingdom

E-mail: nmack@noc.ac.uk

ABSTRACT

A Regional Thermohaline Inverse Method (RTHIM) is presented that estimates velocities through the section bounding an enclosed domain and transformation rates due to interior mixing within the domain, given inputs of surface boundary fluxes of heat and salt and interior distributions of salinity and temperature. The method works by invoking a volumetric balance in thermohaline coordinates between the transformation due to mixing, surface fluxes and advection, while constraining the mixing to be down tracer gradients. The method is validated using a 20-year mean of outputs from the NEMO model in an Arctic and subpolar North Atlantic domain, bound to the south by a section with a mean latitude of 66°N. RTHIM solutions agree well with the NEMO model 'truth' and are robust to a range of parameters; the MOC, heat and freshwater transports calculated from an ensemble of RTHIM solutions are within 12%, 10% and 19%, respectively, of the NEMO values. There is also bulk agreement between RTHIM solution transformation rates due to mixing and those diagnosed from NEMO. Localized differences in diagnosed mixing may be used to guide the development of mixing parameterizations in models such as NEMO, whose downgradient diffusive closures with prescribed diffusivity may be considered oversimplified and too restrictive.

1. Introduction

The complex spatial structure of the global ocean circulation often obscures its key physical phenomena and drivers, including mechanisms involving both internal mixing and external forcing. Lagrangian observations have shown that cartoon-like schematics of the thermohaline circulation may oversimplify the true complexity of the flow pathways (Bower et al. 2009; Lozier 2012), yet it remains desirable to seek a view that is both simple and physically accurate. Thermohaline coordinates have long been used to characterize different water masses. Recent studies have shown that expressing circulation and mixing in this non-geographical coordinate system (e.g. Zika et al. 2012; Hieronymus et al. 2014) filters out adiabatic fluctuations and therefore enables insight into diabatic process which are relevant to the climate system over long timescales.

Globally, it has been demonstrated that a thermohaline-coordinate circulation streamfunction and internal mixing may be diagnosed from observations or model data (Groeskamp et al. 2014a). A companion study further showed that an inverse method, an extension of Walin (1982), may be used to infer the circulation and internal mixing from knowledge of the temperature and salinity combined with surface fluxes - the Thermohaline Inverse Method (THIM) (Groeskamp et al. (2014b); hereafter G14b). THIM elucidates the connectivity of the global transport of heat and freshwater and their drivers, but the tradeoff is that detailed regional information may be lost through the transformation from geographical to thermodynamic coordinates, due to the integration of water masses with the same temperature-salinity properties.

Motivated by the need to understand regional circulation and mixing, this study describes a new Regional Thermohaline Inverse Method (RTHIM) and its validation against output from the Nucleus for European Modelling of the Ocean (NEMO) model (Madec 2008), with a particular application to the Arctic and subpolar North Atlantic (Fig. 1). The equivalent transformation in the regional domain once again achieves a simplified view of the complex circulation and mixing, but the integration of water masses retains information specific to the region of interest, rather than this information being a smaller fraction of a global integral. The outputs from traditional "box inverse" models (Wunsch 1978) have been analyzed in thermohaline coordinates to provide insight into the circulation in specific regions such as the Weddell Gyre (Naveira Garabato et al. 2016) and the Arctic Ocean (Tsubouchi et al. 2012), but RTHIM is unique as an inverse method which directly uses a thermohaline coordinate system applied to a specific region of the ocean, rather than the global ocean, as in THIM. In THIM, water of particular $S-\theta$ properties may exist in more than one region. Due to the averaging process, such regional distinction is lost, but in RTHIM the diagnosics remain regionally specific. There is a tradeoff between sufficient averaging to obtain a simpler view and dilution of regional information by excessive averaging. Both approaches are arguably valid, depending on the particular question or goal. In addition, previous methods that have solved for interior mixing have done so by imposing a simplified spatial structure on the diagnosed diffusivities. RTHIM does not impose such a structure, instead applying only simple constraints on mixing which we describe in section 2b.

Climate models suggest that buoyancy changes in the Arctic and subpolar North Atlantic strongly influence the Atlantic Meridional Overturning Circulation, yet observational evidence is actively being sought to support the model predictions (Lozier et al. 2017). This paper is a first step towards providing relevant robust observational estimates based on RTHIM, which will be used to complement more traditional direct observations provided by new trans-basin arrays such as in the Overturning in the Subpolar North Atlantic Programme (OSNAP; www.o-snap.org).

The paper is organized as follows: in section 2 we describe the principles of the new inverse method and detail the background to its development. Section 3 explains how we have applied

RTHIM to outputs from the NEMO model for validation. In section 4 we present the results of the validation, and section 5 contains discussion and conclusions.

2. The Regional Thermohaline Inverse Method

a. Distributions in thermohaline coordinates

When working in thermohaline coordinates, it is helpful to consider the distribution of water masses in $S-\theta$ space. For example, if we are looking at the transport through a section (e.g. the section indicated by the red line on Fig. 1), the section itself can be divided up using contours of S and θ (Fig. 2 (a)). The section area distribution in $S-\theta$ space from a 20 year time-mean of the NEMO fields (Fig. 2 (b)) is calculated by summing the areas of the section which fall between contours of S and θ, defined by $S \in (S \pm \frac{\Delta S}{2})$, $\theta \in (\theta \pm \frac{\Delta\theta}{2})$, where $\Delta S = 0.1$ PSU and $\Delta\theta = 0.2$ °C. Similarly, the volumetric distribution (Fig. 2 (c)) is calculated by summing the volumes enclosed by contours of S and θ over the whole physical domain. Both distributions are concentrated in comparatively small regions of $S-\theta$ space. For the area distribution, 76% of the total section area of $2.6 \times 10^9\,\mathrm{m}^2$ is covered by 10% of the occupied (S,θ) bins (that is 10% of the bins which have non-zero associated section area). For the volumetric distribution, 95% of the total volume of $1.8 \times 10^{16}\,\mathrm{m}^3$ is contained within 10% of the occupied bins.

Processes acting on and within the ocean affect the distribution of water masses in $S-\theta$ space (Groeskamp et al. 2014a; Pemberton et al. 2015). Surface heat fluxes increase the spread of water masses along the θ axis, while freshwater fluxes increase their spread along the S axis. Interior mixing processes on the other hand act to reduce the spread of water masses in $S-\theta$ space, since mixing acts down tracer gradients. In addition to the most familiar effects of heating/cooling and freshening/salinification, ice melting causes water masses below to cool. This is particularly relevant along the Atlantic Water inflow (Fram Strait and around Svalbard) where surface waters are relatively warm and sea ice melting on top creates large cold freshwater lenses, which get mixed during storms. The competition between boundary and interior processes allows us to diagnose interior mixing in the absence of direct observations, by inferring it from other quantities which are more easily measured (see e.g. Zika et al. 2015).

b. The volume budget

The aim is to estimate, as outputs from the method, section volume advection and interior diffusive processes (described later), given inputs of surface boundary fluxes and interior S and θ. As in G14b, we begin by considering a volume element bound by pairs of isohaline and isothermal surfaces, $V = V(S \pm \frac{\Delta S}{2}, \theta \pm \frac{\Delta\theta}{2}, t)$ (Fig. 3). G14b have absolute salinity S_A and conservative temperature Θ as their coordinates which is appropriate for application to real ocean observations; here we use practical salinity S and potential temperature θ which are conserved in NEMO. The conservation equations for volume, salt and heat mirror those in equations 7-9 of G14b for the global THIM, except that we express the conservation per unit S and θ, and we add terms involving $I_{adv} = I_{adv}(S, \theta)$, which we define as the volume advection (inward normal velocity component times area of volume element at the bounding section) into the regional volume at its open boundary per unit S and θ, leading to the following:

$$\frac{\partial}{\partial t}(V) + \nabla_{S\theta} \cdot \boldsymbol{U}^{dia}_{S\theta} = \frac{1}{\rho_0} f_m + \ I_{adv}, \tag{1}$$

where $\frac{\partial}{\partial t}(V)$ is the volume tendency per unit (S, θ), $\nabla_{S\theta} \cdot \boldsymbol{U}^{dia}_{S\theta}$ is the diathermohaline volume transport divergence per unit (S, θ), $\rho_0 = 1024\,\mathrm{kg\,m^{-3}}$ is a reference density of seawater, f_m is the surface boundary mass flux per unit (S,θ) and I_{adv} is the section volume advection per unit (S,θ);

$$\frac{\partial}{\partial t}(VS)+\nabla_{S\theta}\cdot(S\boldsymbol{U}_{S\theta}^{dia})=\tfrac{1}{\rho_0}\,m_S+I_{adv}S, \tag{2}$$

where $\frac{\partial}{\partial t}(VS)$ is the salt tendency per unit (S,θ), $\nabla_{S\theta}\cdot(S\boldsymbol{U}_{S\theta}^{dia})$ is the diathermohaline salt transport divergence per unit (S,θ), m_S is the salt convergence due to interior diffusion per unit (S,θ), and $I_{adv}S$ is the section salt advection per unit (S,θ); and

$$\frac{\partial}{\partial t}(V\theta)+\nabla_{S\theta}\cdot(\theta\boldsymbol{U}_{S\theta}^{dia})=\tfrac{1}{\rho_0 c_p^0}\,(f_h\quad+\quad m_\theta)\;+\;I_{adv}\theta, \tag{3}$$

where $\frac{\partial}{\partial t}(V\theta)$ is the heat tendency per unit (S,θ), $\nabla_{S\theta}\cdot(\theta\boldsymbol{U}_{S\theta}^{dia})$ is the diathermohaline heat transport divergence per unit (S,θ), c_p^0 is the specific heat capacity of seawater $(4000\,\mathrm{J\,kg^{-1}\,K^{-1}})$, f_h is the heat convergence due to boundary fluxes per unit (S,θ), m_θ is the heat convergence due to interior diffusion per unit (S,θ), and $I_{adv}\theta$ is the section heat advection per unit (S,θ). Note that the superscript '*dia*' denotes a transport through a surface of constant temperature or salinity.

We begin the separation into inputs and outputs by identifying the component terms of the time-mean diathermohaline volume transport $\boldsymbol{U}_{S\theta}^{dia}=(U_{|S}^{dia},U_{|\theta}^{dia})$. Here $U_{|S}^{dia}$ is the transport across an S surface between two θ surfaces per unit θ; and $U_{|\theta}^{dia}$ is the transport across a θ surface between two S surfaces per unit S. We repeat the algebraic manipulation used in G14b on the analogous equations to $(1-3)$, but using a different definition of the thermohaline divergence operator (and similar for the gradient operator), $\nabla_{S\theta}\cdot[\,]=(\frac{\partial}{\partial S},\frac{\partial}{\partial\theta})\cdot[\,]$. The discrete $\frac{\partial}{\partial S}$ operator is defined at the central S value of the grid cell corresponding to $(S\pm\frac{\Delta S}{2},\theta\pm\frac{\Delta\theta}{2})$ and analogous for θ. Note that $\frac{\partial S}{\partial S}=1$ and $\frac{\partial\theta}{\partial S}=0$, and similar when S is swapped with θ. The rightmost terms involving I_{adv} in equations $(1-3)$ above are found to cancel, leading to the following expressions for each of the components in thermohaline coordinates (an overbar represents a time-mean):

$$U_{|S}^{dia}(S,\theta)=U_{|S}^{surf}+U_{|S}^{mix}-U_{|S}^{loc}, \tag{4}$$

$$U^{dia}_{|\theta}(S,\theta) = U^{surf}_{|\theta} + U^{mix}_{|\theta} - U^{loc}_{|\theta}, \tag{5}$$

where

$$U^{surf}_{|S} = -\frac{1}{\rho_0}\overline{Sf_m}, \qquad U^{surf}_{|\theta} = \frac{1}{\rho_0 c^0_p}\overline{f_h - c^0_p \theta f_m} \tag{6}$$

are due to surface processes;

$$U^{mix}_{|S} = \frac{1}{\rho_0}\overline{m_S}, \qquad U^{mix}_{|\theta} = \frac{1}{\rho_0 c^0_p}\overline{m_\theta} \tag{7}$$

are due to interior diffusive processes, defined in more detail later; and

$$U^{loc}_{|S} = \frac{1}{\rho_0}\overline{l_S}, \qquad U^{loc}_{|\theta} = \frac{1}{\rho_0 c^0_p}\overline{l_\theta} \tag{8}$$

are related to the local response, $\overline{l_S} = \rho_0 \overline{V\frac{\partial S}{\partial t}}$ and $\overline{l_\theta} = \rho_0 c^0_p \overline{V\frac{\partial \theta}{\partial t}}$. The local response expresses the rate of change of heat (salt) due to the rate of change of temperature (salinity), for a given volume within a discrete (S,θ) grid cell. Equivalently, it is the amount of heat (salt) that can be added or removed from the system, changing the temperature (salinity) of a water mass by an amount whereby it remains within that grid cell.

We may now use equations (4) and (5) to rewrite the time-mean of equation (1) in terms of desired output estimates (LHS) balanced by specified inputs (RHS):

$$I_{adv} - \nabla_{S\theta}\cdot \boldsymbol{U}^{mix}_{S\theta} = \frac{\partial V}{\partial t} - \frac{1}{\rho_0}\overline{f_m} + \nabla_{S\theta}\cdot \boldsymbol{U}^{surf}_{S\theta} - \nabla_{S\theta}\cdot \boldsymbol{U}^{loc}_{S\theta} \tag{9}$$

For the sake of simplicity, we follow the approach of Groeskamp et al. (2017) and neglect the local response term, since it is small for the (S,θ) grids used in this study. In addition, the mass flux term, f_m, is also negligible for the NEMO solution used here, so that is omitted too.

At this point, we again deviate from G14b by including a more general formulation for the interior mixing, which no longer depends on the assumption of a constant diffusivity with a fixed geographical distribution. We may write the diathermohaline volume transport vector due to interior diffusive processes, $\boldsymbol{U}_{S\theta}^{mix}$, as the divergence of a tensor, $\boldsymbol{F}$, whose properties we may constrain to ensure net downgradient mixing:

$$\boldsymbol{U}_{S\theta}^{mix} = \nabla_{S\theta} \cdot \boldsymbol{F} = \left(\frac{\partial}{\partial S} F_{SS} + \frac{\partial}{\partial \theta} F_{S\theta}, \frac{\partial}{\partial S} F_{\theta S} + \frac{\partial}{\partial \theta} F_{\theta\theta} \right) \tag{10}$$

where $\boldsymbol{F}$ is constrained to be a symmetric diffusive flux tensor in thermohaline coordinates (see the Appendix for an explanation of the symmetric constraint).

Combining this definition of the diffusive flux tensor with equation (9), we may define the inverse problem,

$$I_{adv} - \nabla_{S\theta}^2 \boldsymbol{F} = \frac{\partial V}{\partial t} + \nabla_{S\theta} \cdot \boldsymbol{U}_{S\theta}^{surf} + \varepsilon, \tag{11}$$

where ε is the error in the inverse solution, which we aim to minimize. Equation 11 is our volume budget, representing the fact that the time-mean change in volume of a volume element results from a balance between advection, surface fluxes and interior mixing.

We solve equation (11) for I_{adv} and the elements $F_{C_1C_2}$ of $\boldsymbol{F}$ using constrained optimization (MATLAB's fmincon), with constraints $F_{SS}, F_{\theta\theta} \geq 0$ (net downgradient mixing) , and $F_{S\theta} = F_{\theta S}$ (symmetric diffusive flux tensor). Note that, since equation (11) involves the $\nabla_{S\theta}^2$ operator applied to $\boldsymbol{F}$, it is possible to conceive of an arbitrary, non-divergent gauge, N, to be added to $\boldsymbol{F}$, e.g. $\boldsymbol{F}' = \boldsymbol{F} + \boldsymbol{N}$, where $\nabla_{S\theta}^2 \boldsymbol{N} = 0$, and $\boldsymbol{F}'$ is the modified $\boldsymbol{F}$, yet to still obtain the same volume balance: $\nabla_{S\theta}^2 \boldsymbol{F} = \nabla_{S\theta}^2 \boldsymbol{F}'$. Often a Helmholtz decomposition is used to solve similar gauge problems, given appropriate boundary conditions on $\boldsymbol{F}$. However, there is no appropriate physical metric in thermohaline coordinates for such a decomposition (in Cartesian coordinates (x, y), one metric could be a measure of distance, proportional to $(\Delta x^2 + \Delta y^2)^{\frac{1}{2}}$). Therefore, we emphasize that the interpretation of the RTHIM solution should be focused solely on the role of $\nabla_{S\theta} \cdot \boldsymbol{U}_{S\theta}^{mix}$, rather than on the individual terms in $\boldsymbol{F}$.

3. Application to NEMO

a. NEMO model domain

The model data used for the validation of RTHIM come from a 1° resolution NEMO simulation run at the National Oceanography Centre (NOC) known as ORCA1-N403. It is a z-level Boussinesq global model, with the OPA ocean model (Madec 2008) coupled to the LIM2 sea-ice model (Timmermann et al. 2005). The horizontal grid is tripolar, giving a resolution of ~ 50 km over the Arctic ocean. There are 75 vertical level ranging in thickness from ~ 1 m at the surface to ~ 200 m at the bottom. The model is forced by CORE atmospheric forcing (Large and Yeager 2004) which has 2.5° resolution. We took monthly mean outputs of net surface fluxes of heat and freshwater and of ocean temperature and salinity, from 20 years of the simulation from 1988-2007 and averaged them to produce a monthly climatology. The limited resolution of the simulation means that mesoscale features of the circulation are not resolved; however our focus is to validate the inverse method against an internally consistent system where subgrid mixing is parameterized. The Arctic and subpolar North Atlantic domain volume to which we apply RTHIM is bounded to the south by a line of constant latitude index on the ORCA1-N403 grid. The model has a tripolar grid, so the constant latitude index line is curved in latitude-longitude coordinates (see red lines on Fig. 1). The flow into and out of the domain is through a full-depth section defined by this line. In order for the balance of Eq. 11 to hold, the water masses that flow through the section must be geographically connected within the model so that they can mix. We therefore exclude ranges from the section corresponding to Hudson Bay and part of the Canadian Archipelago (94°W to 68°W, although the latter is connected to the rest of the Arctic in the real ocean it is not connected in the model); and the Gulf of Bothnia (18°E to 22°E). We also mask out the surface fluxes and interior water masses associated with these ranges since these do not contribute to the volume budget for the domain.

b. $S-\theta$ grid

We design an irregular $S-\theta$ grid for use with RTHIM (see Appendix for grid node definitions). We saw in Fig. 2 that the majority of the area of the section occupies a small region in $S-\theta$ space. The along-section distance Δx that can be resolved in the RTHIM solution is related to the θ grid spacing $\Delta\theta$ by $\Delta x \approx (\frac{\partial\theta}{\partial x})^{-1}\Delta\theta$, where $\frac{\partial\theta}{\partial x}$ is the horizontal gradient of θ (the equivalent applies in the S coordinate). An irregular grid allows us to resolve details of the flow through the section, while keeping the size of the inverse problem manageable. We want to focus the resolution in the regions of $S-\theta$ space where gradients of S and θ are small, with additional focus given to parts of the section where the transports are large. In order to do this, we first construct cumulative transport functions in S and θ which measure the magnitude of the transport through the section below a contour, based on the absolute model velocities $|v|$ normal to the section at each model grid cell:

$$|T|_S(S) = \sum_{S_{min}}^{S} |v|\Delta A, \quad |T|_\theta(\theta) = \sum_{\theta_{min}}^{\theta} |v|\Delta A \tag{12}$$

Here the sums are over the area of the section for which the salinity (temperature) is below the salinity (temperature) in the bracket. We then plot the transport functions normalized by their totals, which allows the θ and S ranges to be discretized according to an equipartition of $|T|$ (Fig. 4). Each point in θ (or S) identified by the red dashed lines on Fig. 4(b) (4(c)) as they intersect the $x-$axis then becomes a node in the θ (S) vectors used to construct the $S-\theta$ grid.

c. *Advective flux*

During the constrained optimization, we want RTHIM to compute the advection term I_{adv} from a surface reference level velocity v_{ref}. I_{adv} is calculated by assuming a relative velocity v_{rel} to be added to v_{ref}:

$$v(x,z) = v_{ref}(x) + v_{rel}(x,z) \tag{13}$$

$$I_{adv}(S,\theta) = I_{ref}(S,\theta) + I_{rel}(S,\theta) \tag{14}$$

where (x,z) are coordinates in the along-section and vertical directions. The components of I_{adv} are found by binning relative velocities at the section:

$$I_{rel}(S,\theta) = \frac{1}{\Delta S \Delta\theta} \iint \Pi(S)\Pi(\theta) v_{rel}\, dA, \tag{15}$$

(and similarly for I_{ref}), where $v_{rel}(x,z) = v(x,z) - v_{surf}(x)$ is the NEMO section velocity relative to the fixed surface velocity. The 'Π's mask in only regions of the section between isohalines $S \pm \frac{\Delta S}{2}$ and isotherms $\theta \pm \frac{\Delta\theta}{2}$, and the quantities are integrated over the whole section. We construct a matrix $\mathbf{A}$ such that $I_{ref} = \mathbf{A} v_{ref}$, then $I_{adv}(S,\theta)$ can be calculated using Eq. 14.

In the initial condition, v_{ref} is taken as the model values with random noise added of between -5 cm/s and +5 cm/s with 5-cell boxcar smoothing applied (equivalent to smoothing over 5 degrees longitude). This initial condition is designed to reflect the uncertainty on the surface velocities that might be expected when applying RTHIM to observations and using surface geostrophic velocity from satellite altimetry data for the initial v_{ref}. The 5 cm/s is slightly larger than the 3 cm/s estimated by Gourcuff et al. (2011) as the uncertainty on altimetrically derived surface geostrophic velocities, but also represents an upper limit on the uncertainty that might be expected from in-situ current meter measurements (see e.g. Tsubouchi et al. 2012). We also constrain v_{ref} in the solution to be within 10 cm/s of the NEMO model 'truth', which limits the range of solutions while allowing for a generous amount of uncertainty on our 'best guess' by doubling the expected value.

d. Surface flux

Here we describe how the fluxes of heat and freshwater that forced the NEMO run are used to calculate $\nabla_{S\theta} \cdot \boldsymbol{U}_{S\theta}^{surf}$ for input to RTHIM. The contribution of the surface fluxes to the volume budget, $\nabla_{S\theta} \cdot \boldsymbol{U}_{S\theta}^{surf}$, is a prescribed term in the RTHIM solution, assumed perfectly known. This is appropriate for validating with a model which is an internally consistent system; we discuss the implications for applying RTHIM to observations in section 5. The freshwater flux in $m\,s^{-1}$ (see Fig. 1, left panel) is divided by the thickness of the model's top layer, Δz^{surf}, obtaining f_w in s^{-1}, and $U_{|S}^{surf}$ is then calculated as follows:

$$U_{|S}^{surf}(S^*, \theta^*) = \frac{-S_0}{\Delta S \Delta\theta} \iiint \Pi(S^*)\Pi(\theta^*) f_w \, dV. \tag{16}$$

Here the Πs mask in regions on the surface layer between isohalines S^* and $S^* + \Delta S$ and between isotherms $\theta^* \pm \frac{\Delta\theta}{2}$. The different masking here compared with Eq. 15 for I_{adv} ensures that $\nabla_{S\theta} \cdot \boldsymbol{U}_{S\theta}^{surf}$ will coincide with I_{adv} on the discrete grid, since $U_{|S}^{surf}$ is calculated between S nodes and on θ nodes, while $U_{|\theta}^{surf}$ (below) is calculated on S nodes and between θ nodes, therefore $\nabla_{S\theta} \cdot \boldsymbol{U}_{S\theta}^{surf} = (\frac{\partial}{\partial S}, \frac{\partial}{\partial \theta}) \cdot (U_{|S}^{surf}, U_{|\theta}^{surf})$ will be on S and θ nodes, as is I_{adv}. The heat flux in $W\,m^{-2}$ (Fig. 1, right panel) is divided by $(\rho_0 c_p^0 \Delta z^{surf} = 4.1 \times 10^6 \Delta z^{surf})$ to obtain f_θ in $K\,s^{-1}$, and $U_{|\theta}^{surf}$ calculated as follows:

$$U_{|\theta}^{surf}(S^*, \theta^*) = \frac{1}{\Delta S \Delta\theta} \iiint \Pi(S^*)\Pi(\theta^*) f_\theta \, dV \tag{17}$$

The Πs mask in regions on the surface layer between isohalines $S^* \pm \frac{\Delta S}{2}$ and between isotherms θ^* and $\theta^* + \Delta\theta$.

e. Diffusive flux

We described constructing an initial condition for the reference level velocities in section 3c. We also require an initial condition for the components of the diffusive flux tensor $\boldsymbol{F}$. Using the NEMO fields, we calculate the x, y and z components of the along-isopycnal gradients of S and θ and their vertical gradients, assuming small isopycnal slopes (note that this assumption only applies to the initial condition for $\boldsymbol{F}$, and is not required for the solution). We then calculate the 4 components of $\boldsymbol{F}$ by defining a diffusion tensor $\mathbf{K}$ with a uniform isopycnal and vertical component for the purposes of the initial condition (see Appendix). We also apply some smoothing to the components of $\mathbf{F}$ in the initial condition, since with our variable $S-\theta$ grid they can be noisy where ΔS or $\Delta\theta$ is small. We therefore apply a simple 2D boxcar smoothing to these terms. Along with the surface flux term $\nabla_{S\theta} \cdot \boldsymbol{U}_{S\theta}^{surf}$, the diffusive fluxes are calculated for each month of our NEMO monthly climatology, and the mean of each set of 12 fields is calculated for input to RTHIM.

f. Optimization

Having established the inputs to Eq. 11, RTHIM can now be solved by minimizing the cost function r^2 in units of $\mathrm{Sv}^2 (1\ Sv = 10^6\ \mathrm{m}^3\mathrm{s}^{-1})$ in the following:

$$r^2 = \sum_{S,\theta} (\varepsilon \Delta S \Delta\theta)^2 + w \left[\sum_{S,\theta} I_{adv} \Delta S \Delta\theta - T_{net} \right]^2 \tag{18}$$

where the term in square brackets on the RHS is a constraint on the net transport through the section, calculated as the difference between the integrated section transport from the RTHIM solution advection term I_{adv} and the known 1988-2007 time-mean NEMO net section transport including Bering Strait ($T_{net} = -0.41\mathrm{Sv}$). This net section transport constraint is multiplied by a weighting factor w, which adjusts the relative importance of the constraint compared to that of the volume budget (the sensitivity of our results to the factor w is explored in section 4b). When applying RTHIM to observations, T_{net} would be set to zero or a known value; since we have an enclosed domain, zero net transport would be an appropriate assumption in the absence of other information.

An initial control vector is constructed in two sections: the first from the 4 components of $\mathbf{F}$, which are arrays of values at each S and θ on our grid; and the second from v_{ref}. Each section is normalized by the mean of that section's values to make all the values in the control vector of order 1; this enables the minimization algorithm to more efficiently search the parameter space. Within Matlab/fmincon, we use the Interior Point algorithm to carry out the minimization (Byrd et al. 1998). There is more than one option for terminating the optimisation on arrival at a satisfactory solution. We explored using the default criterion based on cost function gradient becoming smaller than a certain value (it should tend to zero at an extremum), but this was susceptible to gradient noise and oscillations in the solution. Instead we used a criterion based on the cost function value itself. When r^2 drops below $10^{-4}\,\mathrm{Sv}^2$, we accepted the solution (corresponding to a misfit, r, of 0.01 Sv, chosen to be significantly less than observational uncertainty of hydrographic sections or trans-basin mooring arrays).

4. Solution

a. Section transport

In this section we will examine the part of the RTHIM solution that diagnoses the transport through the section bounding the domain. We quantify the method's skill in reproducing the model 'truth' from NEMO using a number of metrics and test the sensitivity of the solution to a range of reasonable parameters. The purpose is to demonstrate that RTHIM will find a good solution given, for example, an uncertainty on v_{ref} that reflects that which would be present when applying the method to observations (i.e. the uncertainty in the surface absolute geostrophic velocities). RTHIM finds a surface v_{ref} by solving Eq. 11, given an initial condition obtained by taking the surface velocities from NEMO and adding some random noise. Given such an initial guess, RTHIM has sufficient skill to converge on the model truth with a high degree of precision (top panel of Fig. 5). The rms error between the RTHIM v_{ref} and the model truth for this solution was $v_{rms} = 0.06\,\mathrm{cm\,s^{-1}}$. The longitude range shown corresponds to the part of the section indicated by the red line on Fig. 1 between Baffin Island to the west and Norway to the east. The longitude range west of Baffin Island including the Hudson bay is not part of our solution as explained in section 3a; Bering Strait is included in the solution but is not plotted. We then obtain the full section velocities by adding the known model shear to v_{ref} (in applying RTHIM to observations the shear will be obtained from thermal wind balance). Even with a high degree of agreement between the RTHIM and NEMO surface velocities, it is still possible to see differences in the full section velocities in the deep part of the section east of the Icelandic Plateau (note that since we use the actual NEMO model shear to construct the section velocities from v_{ref}, any differences at depth must be due to differences at the surface). This highlights the importance of obtaining a good solution for v_{ref}, particularly where the ocean is deepest. We define an additional skill measure comparing the section transport from the RTHIM solution as follows:

$$T_{skill} = \frac{1}{N}\sqrt{\sum_{x=1}^{N}\left[\sum_{z} T_{sol}(x,z) - \sum_{z} T_{mod}(x,z)\right]^2} \tag{19}$$

where $T_{sol}(x,z) = v_{sol}(x,z)dA(x,z)$ and $T_{mod}(x,z) = v_{mod}(x,z)dA(x,z)$ are the volume transports in Sv through the geographical grid cell (x,z) at the section from RTHIM and NEMO, respectively. The sums over z in Eq. 19 are over the water column, and N is the number of x-grid cells along the section. T_{skill} therefore measures the rms error (RTHIM vs NEMO) for the depth-integrated transports at each grid cell in x. For the RTHIM solution shown in Fig. 5, T_{skill}=0.03 Sv, which is less than 4% of the rms of the NEMO depth-integrated transports along the section.

The RTHIM solution v_{ref} is robust to a range of parameters. If we take an ensemble of initial conditions (blue lines on Fig. 6, top panel) while varying a number of other model parameters which will be described in the next subsection, the solutions (red lines on the same figure) cluster around the model truth (black line). We can further demonstrate the model skill in solving for v_{ref} by examining the solution as it converges (Fig. 6, bottom panel). With a convergence tolerance of $r^2 = 1$, the solution (solid blue line) has only deviated from the initial condition (dotted blue line) in a few places; however we note that in this initial period of optimization the cost function r^2 has reduced from $\sim 10^6$ to < 1. This is achieved mainly through adjustments of the elements of $\boldsymbol{F}$, with some small adjustments to v_{ref} required to reduce the net transport constraint in Eq. 18 from $\sim 10^3$ to $\sim 10^{-5}$. As the convergence tolerance is decreased, the v_{ref} solutions converge towards the model truth, with a tolerance of 10^{-4} yielding a solution (solid green line) that almost matches the model (dotted black line).

b. Solution sensitivities

We wish to establish some useful metrics for the RTHIM solution appropriate to a wider oceanographic context and assess the sensitivity of the solution to a range of parameters in terms of those metrics. First we define familiar streamfunctions of the section transport diagnosed from the RTHIM solution, binned in terms of density, temperature and salinity:

$$\psi_\sigma(\sigma^*) = \int_{\sigma \leq \sigma^*} v \, dA \tag{20}$$

where $v\,dA$ is the transport through the section, in this case integrated below contours of constant potential density σ_2 referenced to 2000 m . The similar streamfunctions $\Psi_\theta(\theta^*)$ and $\Psi_S(S^*)$

are defined in the same way, but integrated below contours of temperature and salinity. We then calculate the Meridional Overturning Circulation (MOC), Meridional Heat Transport (MHT) and Meridional Freshwater Transport (MFWT) through the section as follows:

$$MOC = \sum_{\sigma_{min}}^{\sigma_{max}} T(\sigma) \qquad \text{for } T(\sigma) > 0 \tag{21}$$

$$MHT = \rho_0 c_p^0 \int_{\theta_{min}}^{\theta_{max}} \psi_\theta \, d\theta \tag{22}$$

$$MFWT = \frac{-1}{S_0} \int_{S_{min}}^{S_{max}} \psi_S \, dS \tag{23}$$

where in this case $T(\sigma)$ represents the transport integrated between density contours 0.001 $\mathrm{kgm^{-3}}$ apart, $\rho_0 c_p^0 = 4.1 \times 10^6\ \mathrm{Jm^{-3}K^{-1}}$, and $S_0 = 35$ PSU is a reference salinity. Note that we calculate the MOC from $T'(\sigma)$ as in Li et al. (2017) (their Eq. 7) rather than from the streamfunction.

There are a number of parameters which can affect the RTHIM solution. We have already mentioned adding random noise to the v_{ref} initial condition. We also produce an initial condition for the components of the diffusive flux $\boldsymbol{F}$, the other unknown in our inverse problem, based on a uniform diffusivity $\mathbf{K}$ made up of along-isopycnal and vertical components, K and D. We test K values of 10 $\mathrm{m^2s^{-1}}$, 20 $\mathrm{m^2s^{-1}}$, 50 $\mathrm{m^2s^{-1}}$, 100 $\mathrm{m^2s^{-1}}$, 200 $\mathrm{m^2s^{-1}}$, 500 $\mathrm{m^2s^{-1}}$, 1000 $\mathrm{m^2s^{-1}}$, and 2000 $\mathrm{m^2s^{-1}}$; and D values of $10^{-5}\ \mathrm{m^2s^{-1}}$ and $10^{-4}\ \mathrm{m^2s^{-1}}$. There is also the 2D boxcar smoothing applied to the components of $\boldsymbol{F}$ in the initial condition, which has an associated parameter for the number of cells over which to apply the smoothing (we smooth uniformly in both the S and θ directions). We test integer values of 3, 5, 7 and 9 of this parameter, and an additional case where we simply replace each element of $F_{C_1C_2}(S,\theta)$ with the mean of that component over all $S-\theta$ space. In section 3b we described the design of an irregular $S-\theta$ grid over which we bin all the terms of our volume budget, to focus the available resolution where it is needed by dividing up the transport through the section equally in S or θ coordinates. We can choose the number of segments to divide the ranges into, and test grids of 10×10, 13×13, 15×15, and 17×17; at higher resolution than this the number of possible solutions becomes very large and RTHIM has difficulty converging. Finally the constraint on the net section transport in Eq. 18 has a weighting factor w which we vary, testing values of 0, 1, 10, 20, 50, 100, 200, 500, and 1000. We summarize the sensitivity to all these parameters of the streamfunctions and metrics defined above, in Fig. 7 and Table 1.

The principle sensitivities of the RTHIM solutions are to the $S-\theta$ grid and the diffusivity K applied to the $\boldsymbol{F}$ initial condition. This is evident in the purple and dark blue lines on Fig. 7 and in the first 5 rows below the headings of Table 1. In particular a value of $K = 10\ \mathrm{m^2s^{-1}}$ significantly degrades the solution (see Fig. 8 and upper values in the ranges of row 4, columns 3 and 4 on Table 1). The sensitivity to the other parameters is small; however we find that if no smoothing is applied to $\boldsymbol{F}$ in the initial condition, RTHIM will not converge, and if we remove the net transport constraint (i.e. set $w = 0$), the solution suffers dramatically (see Fig. 8). If we consider that this exercise has informed the appropriate choices of parameters as a minimum $S-\theta$ grid size of (13×13) and a K of $100-500\ \mathrm{m^2s^{-1}}$, from the last row of Table 1 we can say that the MHT and MFWT are determined by RTHIM with uncertainties of ± 0.03 PW, and ± 0.03 Sv, respectively, and their mean values agree perfectly with the NEMO model truth. In the case of the MOC, the distribution is slightly skewed, with a range of 10.76-12.91 Sv surrounding a mean of 11.64 Sv; the mean is still in good agreement with the model truth of 11.53 Sv. These uncertainties on the MOC, MHT and MFWT metrics translate to percentage errors of less than 12%, 10% and 19%, respectively, when compared with the model truth.

A strength of this method is the ability to view the circulation in thermohaline coordinates. Fig. 9 (left panel) shows I_{adv} as it is used within RTHIM: binned in $S-\theta$ coordinates. The small residual between the term provided by the RTHIM solution and the term constructed by binning the NEMO section transports on the same grid (right panel) shows the two are very similar. The transport through the section is dominated by a northward flow of warm, salty waters (red colors) and a compensating southward flow of cooler, fresher waters (blue colors). This is what would be expected for this region, since warm, salty North Atlantic water flowing northward is balanced by cooler, fresher water created by boundary processes in the Arctic and Subpolar Gyre.

c. Volume budget

In section 2b we described the volume budget in thermohaline coordinates which RTHIM uses to obtain a solution. We can examine the terms in the volume budget to gain an insight into the contributions of each process to the increase or decrease in volume of water masses in each region of $S-\theta$ space (Fig. 10). Mixing processes (Fig 10 (a)) tend to increase the volume of water towards the center of the $S-\theta$ distribution while decreasing the volume of water at its fringes. This is in accordance with what we expect, since mixing acts down tracer gradients to homogenize extremes in S and θ. By contrast, surface fluxes (b) tend to increase the volume of water at the fringes of the $S-\theta$ distribution while decreasing the volume towards the center. This again is an expected consequence of surface forcing: in particular we can see cooling of the warm surface waters due to the Arctic winds and salinification of the cooler water due to brine rejection in ice formation, both leading to the formation of dense waters. We have already discussed the advection term (c) in terms of the inflow and outflow of different water masses; what we can also see on Fig. 10 is how the residual created by adding the mixing and surface flux terms is largely accounted for by the advection term. The volume trend (d) is comparatively small. After adding together the terms in the volume budget, we are left with a residual, ε. This reduces by several orders of magnitude between the initial condition and the RTHIM solution (Fig. 11).

An alternative way to view the volume budget is by integrating in one direction or the other in thermohaline coordinates. This helps us to see more quantitatively the size of the contributions of each process to the volume budget, and their relative roles at each point in the S or θ range. Integrated first through S and then cumulatively in θ (Fig. 12, left panel), we see that at low temperatures the balance is dominated by a competition between mixing and surface fluxes; whereas at higher temperatures the advection is the larger term, balancing out the other two. The y-values of the lines on Fig. 12 at a given x represent the transformation in volume below that isotherm (or isohaline). This means that a range in x where the gradient $\frac{dy}{dx}$ is positive corresponds to an increase in the volume of water in that temperature or salinity range by a given process, and a range with negative gradient corresponds to a decrease in volume. Therefore on the left panel of Fig. 12 we see surface fluxes increasing the volume of the coldest water and decreasing the volume of the warmer water; meanwhile mixing decreases the volume of the coldest ($-3°$ to $-1.3°$C) and warmest waters ($> 2.5°$C) and increases the volume of waters at intermediate temperatures; and finally advection decreases the volume of water up to 2.5°C (outflow) and increases the volume above that temperature (inflow). The trend term is relatively small. In salinity coordinates (right panel of Fig. 12) the balance is between mixing and surface fluxes between 29.5 and 32.5 PSU; and between mixing and advection at higher salinities, with surface fluxes playing a smaller role. Mixing increases the volume between 29.5-31.5 PSU and decreases the volume between 31.5-32.5 PSU; meanwhile surface fluxes have the opposing effect. Mixing then increases volume between 32.5-34.5 PSU and decreases volume of the saltiest water; in this range balanced by advection.

In section 4b we validated one of the unknowns in our RTHIM solution, the advection term, by comparing section velocities with a known model ‘truth’ from NEMO and exploring the solution sensitivities. We now examine the other unknown: the mixing term. As noted in section 2b, the term constrained in the RTHIM solution is $\nabla_{S\theta} \cdot \boldsymbol{U}_{S\theta}^{mix}$, the mixing term from Fig. 10. Using NEMO diagnostics of the tendencies in temperature and salinity due to mixing, we can construct the components of the transformation vector due to mixing as:

$$U_{|S}^{mix-NEMO}(S^*,\theta^*) = \frac{-S_0}{\Delta S \Delta\theta} \iiint \Pi(S^*)\Pi(\theta^*) \frac{dS}{dt}^{mix} dV. \tag{24}$$

$$U_{|\theta}^{mix-NEMO}(S^*,\theta^*) = \frac{1}{\Delta S \Delta\theta} \iiint \Pi(S^*)\Pi(\theta^*) \frac{d\theta}{dt}^{mix} dV \tag{25}$$

where $\frac{d\theta}{dt}^{mix}$ and $\frac{dS}{dt}^{mix}$ are the NEMO temperature and salinity trends due to mixing in Ks^{-1} and s^{-1}, respectively. We can then calculate $\nabla_{S\theta} \cdot \boldsymbol{U}_{S\theta}^{mix-NEMO}$, the NEMO equivalent of our RTHIM mixing term, to compare with the RTHIM solutions.

We can compare the mixing terms from the RTHIM solution and NEMO by plotting each on the same $S-\theta$ grid (Fig. 13). There are some broadly similar patterns such as the decrease in volume of waters at low temperatures and an increase in volume towards the center of the $S-\theta$ distribution, and there are differences such as the larger values in NEMO at lower salinities. It is not surprising that RTHIM does not capture the details of the NEMO diffusive transformation. The mixing term is highly differentiated (it is calculated from $\nabla_{S\theta}^2 \boldsymbol{F}$) and the dominant signals are likely the advection at the boundary and surface fluxes, with mixing being a small residual in many cases. So a small percentage error in the flow at the boundary may manifest as a large sized error in the mixing terms. We can also integrate the mixing terms from Fig. 13 in $S-\theta$ space, in one dimension or the other (Fig. 14). Although the structure of the NEMO and RTHIM terms differs, the convergence of the blue and black lines at the top end of the ranges on each panel demonstrates that the total diffusive transformation integrated over all $S-\theta$ space is almost the same (0.42 Sv and 0.14 Sv in RTHIM and NEMO, respectively). This shows that the bulk diffusive transformation is in the same balance as in the NEMO 'truth', in equilibrium with advection and surface fluxes when integrated over the control volume.

5. Discussion and conclusions

a. Summary

We have developed 'RTHIM', a new inverse method in thermohaline coordinates to use temperature, salinity and surface flux observations to provide estimates of water mass transformation and circulation in an enclosed domain. The method builds on the work of Groeskamp et al. (2014a,b) and others, and allows analysis previously done globally to be applied to a specific region of the ocean, thereby retaining additional useful geographical information about the circulation. Using inputs of time-mean S, θ, and surface fluxes of heat and salt, RTHIM solves for a surface reference level velocity v_{ref} at the section bounding the domain and interior mixing within it, while constraining the mixing to be down tracer gradients and v_{ref} to be within $\pm 10\,\mathrm{cm\,s^{-1}}$ of an initial guess. Full section velocities are constructed from v_{ref} assuming a known vertical velocity shear. RTHIM was applied to model data from NEMO, and the solution compared to the model 'truth'. The rms error between RTHIM and NEMO for v_{ref} was less than $0.63\,\mathrm{cm\,s^{-1}}$ from an ensemble of solutions with varying parameters in RTHIM. Estimates of the MOC from the ensemble were 10.76-12.91 Sv with a mean of 11.64 Sv, compared to 11.53 Sv in NEMO. The MHT and MFWT from RTHIM were 0.29 ± 0.03 PW and -0.16 ± 0.03 Sv, in exact agreement with NEMO. The transformation due to mixing from the RTHIM solutions was also compared with that obtained diagnostically from NEMO, and while there are some similar patterns the agreement is not as strong. However this is expected because the mixing in NEMO is governed by a parameterization including a specified horizontal eddy diffusivity, whereas in RTHIM we only constrain the diffusive flux tensor $\boldsymbol{F}$.

b. Previous studies of the region

We have applied RTHIM to a domain including the Arctic and part of the subpolar North Atlantic, since our future aim is to use it to estimate the circulation and mixing from observations in the OSNAP study region. There have been some previous efforts to study the circulation in these regions using a thermohaline coordinate system. In a purely diagnostic study, Pemberton et al. (2015) calculated the contributions of mixing, surface fluxes and advection to water mass transformation in the Arctic using NEMO model outputs. They used much higher resolution in $S-\theta$ space than we were able to achieve with RTHIM, and they also separated the effects of isopycnal and diapycnal mixing; however since their method was diagnostic it could not be applied to observations where the complete fields are not available. In comparing our Fig. 10 to their Fig. 16, there is striking similarity between the advection terms. The surface and interior mixing terms are broadly similar, with formation of water at the fringes of the $S-\theta$ distribution by the surface fluxes with destruction towards the center, and vice-versa for the interior mixing (note the opposite sign convention for the formation/destruction of water masses). However there are some differences in the details. We see relatively little transformation of the coldest water between salinities of 29-32 PSU in the RTHIM solution where Pemberton see significant transformation. On the other hand, we see significant transformation up to $\sim 2°$C in the salinity range 32-34 PSU where Pemberton find almost all the transformation occurs below 0°C in the same salinity range. The differences are likely due to a combination of the higher resolution in the Pemberton study, the different domains (Pemberton's Arctic does not include the GIN seas), and the differences between RTHIM and NEMO's treatment of the interior mixing that we have already discussed. An inverse estimate based on observations was carried out by Tsubouchi et al. (2012) for one month of data in the summer of 2005. Their method relied on current data from moorings and CTD sections, and they calculated boundary fluxes and water mass transformation using a box-inverse approach. RTHIM could be applied to observational reanalysis to obtain such estimates over a much longer time period.

c. RTHIM sensitivity

In section 4b we found that RTHIM solutions were robust provided some constraints are placed on the parameters involved in the inverse model setup. RTHIM demonstrated good skill for an $S-\theta$ grid of at least 13×13 cells; however the remainder of the sensitivity study was carried out with a grid size of 17×17, the highest resolution for which RTHIM was able to converge. When applying RTHIM to observations it would make sense to use the highest possible resolution so as to retain the maximum information both in $S-\theta$ space and in physical space for the determination of v_{ref}. The trade-off would be one of computation time, and this would become a factor if large numbers of RTHIM runs were required. For the diapycnal diffusivity D applied to the initial condition of the diffusive flux tensor, we only explored two values: 10^{-4} and 10^{-5} m^{2}s^{-1}. This is reasonable because the average diapycnal diffusivity over the global ocean is well constrained, although its value can vary locally by several orders of magnitude (see e.g. Munk and Wunsch 1998). The volume-weighted time- and spatial-mean of the vertical diffusivity in NEMO for our domain was 5.6×10^{-5} m^{2}s^{-1}. We explored a wider range of the isopycnal diffusivity K since this is a more uncertain parameter involved in the parameterization of unresolved eddies in the NEMO model. Our optimal initial condition values of $K=100-500$ m^{2}s^{-1} are also consistent with NEMO, where the volume-weighted mean lateral diffusivity for our domain was 473 m^{2}s^{-1}. The inability of RTHIM to converge on a solution when $\boldsymbol{F}$ is not smoothed in the initial condition reflects the fact that $\boldsymbol{F}$ is twice differentiated in $S-\theta$ space, so that any noise on the initial fields may be amplified causing very large initial values of $\nabla_{S\theta}\cdot\boldsymbol{U}_{S\theta}^{mix}$. A minimal amount of smoothing of $\boldsymbol{F}$ proved to be enough to overcome this problem. The sensitivity to the other parameters explored was small.

d. Future application to observations

This validation exercise has proven the effectiveness of RTHIM when applied to an internally consistent system, where the ocean has been forced by the prescribed surface fluxes, and where there are no gaps in the data input to the inverse model. When applying the model to observations, there will be some additional challenges to overcome.

We can obtain the temperature and salinity fields needed to construct the diffusive fluxes initial condition and bin the advective fluxes in $S-\theta$ space from observational reanalysis such as EN4 (Good et al. 2013). However, the sparseness of observations over much of the Arctic means that these regions are heavily weighted towards model data. It remains to be seen how much of an obstacle this will pose to obtaining good RTHIM solutions, however since the interior observations are only required for constructing the initial condition, and the validation has proven some resilience to this, it seems reasonable to be optimistic.

In our validation we have also assumed the surface fluxes to be a known quantity. This will not be the case for the application, where again there is quite significant uncertainty in places due to a lack of available observations. This will contribute a degree of uncertainty to the mixing solution, which is affected through the volume budget. Our intended approach is to run RTHIM using a range of surface flux products (for example Lindsay et al. 2014, suggest ERA-INTERIM, CFSR and MERRA are the most consistent with observations in the Arctic) to constrain our estimates. Recent efforts to address the lack of observations such as the N-ICE2015 expedition (Granskog et al. 2016) and the upcoming MOSAiC expedition (`http://www.mosaicobservatory.org/index.html`) may also provide valuable data for the application of RTHIM to the region.

During the validation we have calculated an initial condition for the reference level velocity by adding noise to the model truth, and have taken the vertical velocity shear as the model shear. When applying to observations, we will obtain our v_{ref} initial condition from AVISO surface geostrophic velocities. We have accounted for uncertainties in the AVISO velocities through the noise added to the initial condition, but these errors can be larger at land boundaries. Gourcuff et al. (2011) found that uncertainties in altimetric surface geostrophic velocities near the oceanic western boundary with Greenland led to uncertainties in the East Greenland Irminger Current from an inverse method of $\sim 30\%$, and that in general altimetric velocities were weaker than those obtained from Ship Acoustic Doppler Current Profiler observations along the OVIDE section between Greenland and Portugal. These factors will need to be considered when making use of AVISO surface geostrophic velocities with RTHIM. The assumption of known velocity shear could be met by a variety of observational constraints. For example, current meter moorings could give direct measurements of velocity. Alternatively, below the Ekman layer and away from strong currents, the flow is geostrophic, so thermal wind shear may be derived from observations of S and along the section. For the OSNAP section, both types of observational constraint are present. The region is also well sampled by the ARGO program, so reanalyses are well constrained by observations (Good et al. 2013). Furthermore, for the OSNAP region, the Ekman component of the advection is weak, contributing only 1 Sv to the overturning (Mercier et al. 2015).

Finally, we have carried out the current analysis for a 20-year mean of NEMO data. Ultimately our intention is to use RTHIM to examine temporal variability in the circulation, which would mean shortening the averaging period, perhaps to inter-annual or shorter variability. This could increase the size of the trend term, which in the validation proved small enough to neglect (although we have not done so); since it will be harder to calculate from uncertain observed interior distributions of S and θ, its importance will place a lower limit on the timescale of variability we can study.

e. Comparison with earlier methods

RTHIM has built on previously developed inverse methods for diagnosing the oceanic circulation from hydographic observations, and has several advantages over its predecessors. It is a natural extension of the traditional box inverse method for determining the flow through a section by solving for a reference level velocity, but instead of dividing the section into layers of depth or density, we use contours of S and θ, thereby defining a coordinate system suitable for studying the conservation of heat and salt. The Bernoulli method (Killworth 1986) and beta-spiral method (Stommel and Schott 1977) both make use of gradients of tracers such as potential vorticity on isopycnals, but assume no mixing in their solution. The Tracer Contour Inverse Method (TCIM) of Zika et al. (2010) and Thermohaline Inverse Method of Groeskamp et al. (2014a) have mixing as a leading-order process, however as for the box inverse, the mixing has a fixed spatial structure. In the case of TCIM and the box inverse they assume one coefficient of isopycnal and diapycnal mixing in each layer. The ability of RTHIM to constrain mixing to be downgradient without imposing such a simplified structure makes it a major advance on earlier methods.

f. Conclusion

In summary, we have presented and carried out a successful validation of a new inverse method, RTHIM. While its application to observations will present some challenges, the method has the potential to provide independent estimates of circulation and interior mixing complementary to the OSNAP array observations, and could ultimately be applied to any enclosed region of the ocean in future work. A verification test of RTHIM and its assumptions may be derived from a hindcast of independently-observed section transport, such as from the OSNAP array. If the hindcast proves to be skillful, that further supports the likely validity of RTHIM solutions for the OSNAP section at other times.

Acknowledgments. NM, CW, and NPH acknowledge funding from the UK Natural Research Council under the UK OSNAP Large Grant (NE/K010875.1). NPH is additionally supported by the UK Natural Environment Research Council's Atlantic Climate System Integrated Study program (NE/N018044/1). We would also like to thank George Nurser (NOC) and Adam Blaker (NOC) for their help with the NEMO outputs. The model output used for the inverse model validation is stored at the British Oceanographic Data Centre under doi:10.5285/6371e0e7-8197-2959-e053-6c86abc0630d.

APPENDIX

Symmetric diffusive flux tensor constraint

Here we derive the constraint that the diffusive flux tensor is symmetric as used in RTHIM. Consider temperature and salinity iso-surfaces with unit normal vectors $\nabla\theta/|\nabla\theta|$ and $\nabla S/|\nabla S|$ respectively. The flux of heat and salt per unit θ and S across these surfaces is given by the flux tensor $\mathbf{F}$ with components

$$F_{\theta\theta}(S^*,\theta^*) = \frac{1}{\Delta S}\iint \Pi(S^*)\left(\mathbf{K}\nabla\theta\cdot\nabla\theta/|\nabla\theta|\right) dA_{\theta^*}, \tag{A1}$$

$$F_{\theta S}(S^*,\theta^*) = \frac{1}{\Delta \theta}\iint \Pi(\theta^*)\left(\mathbf{K}\nabla\theta\cdot\nabla S/|\nabla S|\right) dA_{S^*}, \tag{A2}$$

$$F_{S\theta}(S^*,\theta^*) = \frac{1}{\Delta S}\iint \Pi(S^*)\left(\mathbf{K}\nabla S\cdot\nabla\theta/|\nabla\theta|\right) dA_{\theta^*}, \tag{A3}$$

$$F_{SS}(S^*,\theta^*) = \frac{1}{\Delta \theta}\iint \Pi(\theta^*)\left(\mathbf{K}\nabla S\cdot\nabla S/|\nabla S|\right) dA_{S^*}, \tag{A4}$$

where $\Pi(S^*)$ is a boxcar function over $S^* \pm \Delta S/2$, $\mathbf{K}$ is a 3D and time dependent diffusion tensor and $\iint dA_S^*$ is an area integral over the surface where $S = S^*$. Each component of $\mathbf{F}$, $F_{C_1C_2}$ represents the diffusive flux of tracer C_1 across and in the direction perpendicular to the iso-surface of tracer C_2.

Considering specifically the diffusive flux of temperature across the salinity surface and salt across the temperature surface in the limit as $\Delta\theta$ and ΔS become small we have

$$F_{\theta\theta}(S^*,\theta^*) = \frac{1}{\Delta S \Delta\theta}\iiint \Pi(\theta^*)\Pi(S^*)\mathbf{K}\nabla\theta\cdot\nabla\theta dV \tag{A5}$$

$$F_{\theta S}(S^*,\theta^*) = \frac{1}{\Delta S \Delta\theta}\iiint \Pi(\theta^*)\Pi(S^*)\mathbf{K}\nabla\theta\cdot\nabla S dV \tag{A6}$$

$$F_{S\theta}(S^*,\theta^*) = \frac{1}{\Delta S \Delta\theta}\iiint \Pi(\theta^*)\Pi(S^*)\mathbf{K}\nabla S\cdot\nabla\theta dV \tag{A7}$$

$$F_{SS}(S^*,\theta^*) = \frac{1}{\Delta S \Delta\theta}\iiint \Pi(\theta^*)\Pi(S^*)\mathbf{K}\nabla S\cdot\nabla S dV \tag{A8}$$

Typically coarse resolution ocean models use diffusion tensors with off diagonal terms to account for sub-grid adiabatic advective effects which arise in Eulerian coordinates (e.g. Stoke's drift and transient eddy advection). These are not relevant to transport across iso-surfaces which move with such adiabatic advection and the diffusion tensor can be assumed to have only diagonal elements. We assume the same diffusion tensor applies to both heat and salt (i.e. that turbulent dominates over molecular diffusion). The diffusivity tensor having only diagonal elements and being equivalent for both heat and salt implies that $\mathbf{K}\nabla\theta\cdot\nabla S = \mathbf{K}\nabla S\cdot\nabla\theta$ and hence that $F_{S\theta} = F_{\theta S}$.

$S-\theta$ grid nodes

The grid nodes constructed from the transport functions detailed on Fig. 4, including additional nodes at either end to ensure binning of the whole range of (S,θ) found within the control volume, are as follows:

θ: [−4, −3.05, −1.65, −1, −0.75, −0.1, 0.55, 0.95, 1.3, 2, 3.05, 6.1, 7.2, 7.55, 8.8, 13.15, 14]

S: [6, 6.93, 32.15, 32.86, 33.76, 34.12, 34.46, 34.67,
34.75, 34.78, 34.83, 34.88, 34.95, 35.01, 35.05, 35.55, 36].

Sensitivity study breakdown

References

Bower, A. S., M. S. Lozier, S. F. Gary, and C. W. Böning, 2009: Interior pathways of the North Atlantic meridional overturning circulation. *Nature*, **459 (7244)**, 243–247, doi:10.1038/nature07979, URL http://www.nature.com/doifinder/10.1038/nature07979.

Byrd, R. H., M. E. Hribar, and J. Nocedal, 1998: An Interior Point Algorithm for Large-Scale Nonlinear Programming. *SIAM Journal on Optimization*, **9 (4)**, 877–900, doi:10.1137/S1052623497325107, URL https://doi.org/10.1137/S1052623497325107.

Good, S. A., M. J. Martin, and N. A. Rayner, 2013: EN4: Quality controlled ocean temperature and salinity profiles and monthly objective analyses with uncertainty estimates. *Journal of Geophysical Research: Oceans*, **118 (12)**, 6704–6716, doi:10.1002/2013JC009067, URL http://doi.wiley.com/10.1002/2013JC009067.

Gourcuff, C., P. Lherminier, H. Mercier, and P. Y. Le Traon, 2011: Altimetry combined with hydrography for ocean transport estimation. *Journal of Atmospheric and Oceanic Technology*, **28 (10)**, 1324–1337, doi:10.1175/2011JTECHO818.1.

Granskog, M., P. Assmy, S. Gerland, G. Spreen, H. Steen, and L. Smedsrud, 2016: Arctic Research on Thin Ice: Consequences of Arctic Sea Ice Loss. *Eos*, **97**, doi:10.1029/2016EO044097, URL https://eos.org/project-updates/arctic-research-on-thin-ice-consequences-of-arctic-sea-ice-loss.

Groeskamp, S., B. M. Sloyan, J. D. Zika, and T. J. McDougall, 2017: Mixing Inferred from an Ocean Climatology and Surface Fluxes. *Journal of Physical Oceanography*, **47 (3)**, 667–687, doi:10.1175/JPO-D-16-0125.1, URL http://journals.ametsoc.org/doi/10.1175/JPO-D-16-0125.1.

Groeskamp, S., J. D. Zika, T. J. McDougall, B. M. Sloyan, and F. Laliberté, 2014a: The Representation of Ocean Circulation and Variability in Thermodynamic Coordinates. *Journal of Physical Oceanography*, **44 (7)**, 1735–1750, doi:10.1175/JPO-D-13-0213.1, URL http://journals.ametsoc.org/doi/abs/10.1175/JPO-D-13-0213.1.

Groeskamp, S., J. D. Zika, B. M. Sloyan, T. J. McDougall, and P. C. McIntosh, 2014b: A Thermohaline Inverse Method for Estimating Diathermohaline Circulation and Mixing. *Journal of Physical Oceanography*, **44 (10)**, 2681–2697, doi:10.1175/JPO-D-14-0039.1, URL http://journals.ametsoc.org/doi/abs/10.1175/JPO-D-14-0039.1.

Hieronymus, M., J. Nilsson, and J. Nycander, 2014: Water Mass Transformation in SalinityTemperature Space. *Journal of Physical Oceanography*, **44 (9)**, 2547–2568, doi:10.1175/JPO-D-13-0257.1, URL http://journals.ametsoc.org/doi/abs/10.1175/JPO-D-13-0257.1.

Killworth, P. D., 1986: A Bernoulli Inverse Method for Determining the Ocean Circulation. *Journal of Physical Oceanography*, **16 (12)**, 2031–2051, doi:10.1175/1520-0485(1986)016⟨2031:ABIMFD⟩2.0.CO;2, URL http://journals.ametsoc.org/doi/abs/10.1175/1520-0485{\%}281986{\%}29016{\%}3C2031{\%}3AABIMFD{\%}3E2.0.CO{\%}3B2.

Large, W. G., and S. G. Yeager, 2004: Diurnal to decadal global forcing for ocean and sea–ice models: {The} data sets and flux climatologies. *NCAR Tech. Note*, **TN–460+ST (May)**, 105pp, doi:10.5065/D6KK98Q6.

Li, F., M. S. Lozier, and W. E. Johns, 2017: Calculating the Meridional Volume, Heat, and Freshwater Transports from an Observing System in the Subpolar North Atlantic: Observing System Simulation Experiment. *Journal of Atmospheric and Oceanic Technology*, **34 (7)**, 1483–1500, doi:10.1175/JTECH-D-16-0247.1, URL http://journals.ametsoc.org/doi/10.1175/JTECH-D-16-0247.1.

Lindsay, R., M. Wensnahan, a. Schweiger, and J. Zhang, 2014: Evaluation of seven different atmospheric reanalysis products in the arctic. *Journal of Climate*, **27 (7)**, 2588–2606, doi:10.1175/JCLI-D-13-00014.

Lozier, M. S., 2012: Overturning in the North Atlantic. *Annual Review of Marine Science*, **4 (1)**, 291–315, doi:10.1146/annurev-marine-120710-100740, URL http://www.annualreviews.org/doi/10.1146/annurev-marine-120710-100740, arXiv:1011.1669v3.

Lozier, M. S., and Coauthors, 2017: Overturning in the Subpolar North Atlantic Program: A New International Ocean Observing System. *Bulletin of the American Meteorological Society*, **98 (4)**, 737–752, doi:10.1175/BAMS-D-16-0057.1, URL http://journals.ametsoc.org/doi/10.1175/BAMS-D-16-0057.1.

Madec, G., 2008: *NEMO ocean engine*. Note du Pôle de modélisation, Institut Pierre-Simon Laplace (IPSL), France, No 27, ISSN No 1288-1619.

Mercier, H., and Coauthors, 2015: Variability of the meridional overturning circulation at the GreenlandPortugal OVIDE section from 1993 to 2010. *Progress in Oceanography*, **132**, 250–261, doi:10.1016/j.pocean.2013.11.001, URL http://linkinghub.elsevier.com/retrieve/pii/S0079661113002206.

Munk, W., and C. Wunsch, 1998: Abyssal recipes II: energetics of tidal and wind mixing. *Deep Sea Research Part I: Oceanographic Research Papers*, **45 (12)**, 1977–2010, doi:10.1016/S0967-0637(98)00070-3, URL http://linkinghub.elsevier.com/retrieve/pii/S0967063798000703.

Naveira Garabato, A. C., J. D. Zika, L. Jullion, P. J. Brown, P. R. Holland, M. P. Meredith, and S. Bacon, 2016: The thermodynamic balance of the Weddell Gyre. *Geophysical Re-*

search Letters, **43 (1)**, 317–325, doi:10.1002/2015GL066658, URL http://doi.wiley.com/10.1002/2015GL066658.

Pemberton, P., J. Nilsson, M. Hieronymus, and H. E. M. Meier, 2015: Arctic Ocean Water Mass Transformation in ST Coordinates. *Journal of Physical Oceanography*, **45 (4)**, 1025–1050, doi:10.1175/JPO-D-14-0197.1, URL http://journals.ametsoc.org/doi/abs/10.1175/JPO-D-14-0197.1http://journals.ametsoc.org/doi/10.1175/JPO-D-14-0197.1.

Stommel, H., and F. Schott, 1977: The beta spiral and the determination of the absolute velocity field from hydrographic station data. *Deep Sea Research*, **24 (3)**, 325–329, doi:10.1016/0146-6291(77)93000-4, URL http://linkinghub.elsevier.com/retrieve/pii/0146629177930004.

Timmermann, R., H. Goosse, G. Madec, T. Fichefet, C. Ethe, and V. Dulière, 2005: On the representation of high latitude processes in the ORCA-LIM global coupled sea iceocean model. *Ocean Modelling*, **8 (1-2)**, 175–201, doi:10.1016/j.ocemod.2003.12.009, URL http://linkinghub.elsevier.com/retrieve/pii/S146350030300074X.

Tsubouchi, T., and Coauthors, 2012: The Arctic Ocean in summer: A quasi-synoptic inverse estimate of boundary fluxes and water mass transformation. *Journal of Geophysical Research: Oceans*, **117 (1)**, 1–28, doi:10.1029/2011JC007174.

Walin, G., 1982: On the relation between sea-surface heat flow and thermal circulation in the ocean. *Tellus*, **34 (2)**, 187–195, doi:10.1111/j.2153-3490.1982.tb01806.x, URL http://tellusa.net/index.php/tellusa/article/view/10801.

Wunsch, C., 1978: The North Atlantic general circulation west of 50W determined by inverse methods. *Reviews of Geophysics*, **16 (4)**, 583–620, doi:10.1029/RG016i004p00583, URL http://doi.wiley.com/10.1029/RG016i004p00583.

Zika, J. D., M. H. England, and W. P. Sijp, 2012: The Ocean Circulation in Thermohaline Coordinates. *Journal of Physical Oceanography*, **42 (5)**, 708–724, doi:10.1175/JPO-D-11-0139.1, URL http://journals.ametsoc.org/doi/abs/10.1175/JPO-D-11-0139.1.

Zika, J. D., T. J. McDougall, and B. M. Sloyan, 2010: A Tracer-Contour Inverse Method for Estimating Ocean Circulation and Mixing. *Journal of Physical Oceanography*, **40 (1)**, 26–47, doi:10.1175/2009JPO4208.1, URL http://journals.ametsoc.org/doi/abs/10.1175/2009JPO4208.1.

Zika, J. D., N. Skliris, A. J. George Nurser, S. A. Josey, L. Mudryk, F. Laliberté, and R. Marsh, 2015: Maintenance and broadening of the ocean's salinity distribution by the water cycle. *Journal of Climate*, **28 (24)**, 9550–9560, doi:10.1175/JCLI-D-15-0273.1.

LIST OF TABLES

TABLE 1. A summary of the sensitivities of RTHIM metrics to various parameters. Each row details an ensemble of runs where a particular parameter (column 1) is varied within a range (column 2). All other parameters are kept constant while exploring a particular parameter. The remaining columns give the metrics of the rms difference between NEMO and RTHIM reference level velocities and depth-integrated transports v_{rms} and T_{skill}; and the Meridional Overturning Circulation (MOC), Meridional Heat Transport (MHT) and Meridional Freshwater Transport (MFWT) as described in section 4. Where 'range' appears at the column head, the lower and upper values of the metric from the ensemble are given. Where a value appears in italics at the column head, these are the NEMO values. The parameters are the number of nodes in S and θ on the RTHIM $S-\theta$ grid, the vertical and lateral diffusivities D and K applied to the diffusive flux initial condition, the weighting factor w applied to the section net transport constraint, the random noise applied to the v_{ref} initial condition, and the number of cells of boxcar smoothing applied to the diffusive flux initial condition. The MOC and MFWT are in Sv; the MHT is in PetaWatts (PW). The last row of the table summarizes the results excluding rows 1 and 4. The break down of the results used to construct this table is given in Table 2 in the Appendix

Parameter	Range	v_{rms}	T_{skill}	MOC *(11.53)*	MOC	MHT *(0.29)*	MHT	MFWT *(-0.16)*	MFWT
		(range)	(range)	(mean)	(range)	(mean)	(range)	(mean)	(range)
Grid size	10-17	0.14, 0.92	0.09, 0.74	12.61	11.00, 15.05	0.32	0.28, 0.37	-0.13	-0.18, -0.06
Grid size	13-17	0.14, 0.35	0.09, 0.20	11.80	11.00, 12.25	0.30	0.28, 0.32	-0.15	-0.18, -0.13
D ($m^2 s^{-1}$)	$10^{-4}-10^{-5}$	0.07, 0.15	0.04, 0.10	11.58	11.44, 11.72	0.30	0.29, 0.30	-0.17	-0.17, -0.17
K ($m^2 s^{-1}$)	10-2000	0.07, 1.99	0.04, 1.90	14.49	11.44, 28.30	0.26	0.21, 0.31	-0.18	-0.32, -0.14
K ($m^2 s^{-1}$)	100-500	0.07, 0.63	0.04, 0.34	11.85	11.44, 12.36	0.28	0.26, 0.31	-0.16	-0.17, -0.14
w	1-1000	0.09, 0.32	0.06, 0.23	11.50	10.76, 12.91	0.29	0.28, 0.29	-0.16	-0.19, -0.14
v_{ref} I.C.	± 5 cms^{-1} noise	0.07, 0.23	0.04, 0.17	11.74	11.34, 12.25	0.29	0.28, 0.31	-0.17	-0.17, -0.15
$\boldsymbol{F}$ smoothing	3-9 and mean	0.06, 0.09	0.04, 0.06	11.42	11.36, 11.55	0.29	0.29, 0.29	-0.17	-0.17, -0.16
Overall		0.06, 1.99	0.04, 1.90	12.78	10.76, 28.30	0.28	0.21, 0.37	-0.17	-0.32, -0.06
Excl. outliers		0.06, 0.63	0.04, 0.34	11.64	10.76, 12.91	0.29	0.26, 0.32	-0.16	-0.19, -0.13

TABLE 2. Breakdown of the results of the sensitivity study summarised in Table 1. For each group of runs (column 2) all parameters (columns 3-7) are kept constant except the one being tested for the sensitivity of the metrics (columns 8-12) to that parameter. The parameters are the number of nodes on the RTHIM S and θ grids, the vertical and lateral diffusion coefficients applied to the diffusive flux initial condition D and K in $\mathrm{m^2\,s^{-1}}$, the weighting factor w on the section solution net transport constraint, the random noise applied to the v_{ref} initial condition (no column), and the smoothing applied to the diffusive flux initial condition. The metrics of v_{rms} in $\mathrm{cm\,s^{-1}}$, T_{skill} in Sv, MOC in Sv, MHT in PW and MFWT in Sv are as described in section 4. Where metric values are reported as N/A RTHIM failed to converge on a solution. $5\,\mathrm{cm\,s^{-1}}$ random noise is added to the initial v_{ref} for all runs, but for each grouping except for the v_{ref} I.C. noise grouping the same noise profile is used for each run within a group.

Run	Parameter varied	Grid size	D	K	w	F smoothing	v_{rms}	T_{skill}	MOC	MHT	MFWT
1	Grid size	10×10	10^{-4}	200	10	3-cell boxcar	0.92	0.74	15.05	0.37	-0.06
2	Grid size	13×13	10^{-4}	200	10	3-cell boxcar	0.35	0.19	12.25	0.32	-0.13
3	Grid size	15×15	10^{-4}	200	10	3-cell boxcar	0.23	0.20	11.00	0.28	-0.15
4	Grid size	17×17	10^{-4}	200	10	3-cell boxcar	0.14	0.09	12.14	0.29	-0.18
5	**F** I.C. diffusivities	17×17	10^{-4}	10	10	3-cell boxcar	1.84	1.81	22.08	0.25	-0.12
6	**F** I.C. diffusivities	17×17	10^{-4}	20	10	3-cell boxcar	1.03	0.69	12.64	0.25	-0.24
7	**F** I.C. diffusivities	17×17	10^{-4}	50	10	3-cell boxcar	N/A	N/A	N/A	N/A	N/A
8	**F** I.C. diffusivities	17×17	10^{-4}	100	10	3-cell boxcar	0.34	0.32	12.36	0.27	-0.16
9	**F** I.C. diffusivities	17×17	10^{-4}	200	10	3-cell boxcar	0.07	0.04	11.72	0.29	-0.17
10	**F** I.C. diffusivities	17×17	10^{-4}	500	10	3-cell boxcar	0.27	0.16	11.64	0.31	-0.17
11	**F** I.C. diffusivities	17×17	10^{-4}	1000	10	3-cell boxcar	0.80	0.51	13.74	0.22	-0.16
12	**F** I.C. diffusivities	17×17	10^{-4}	2000	10	3-cell boxcar	1.02	0.63	14.36	0.21	-0.15
13	**F** I.C. diffusivities	17×17	10^{-5}	10	10	3-cell boxcar	1.99	1.90	28.30	0.30	-0.32
14	**F** I.C. diffusivities	17×17	10^{-5}	20	10	3-cell boxcar	0.85	0.70	12.84	0.27	-0.20
15	**F** I.C. diffusivities	17×17	10^{-5}	50	10	3-cell boxcar	N/A	N/A	N/A	N/A	N/A
16	**F** I.C. diffusivities	17×17	10^{-5}	100	10	3-cell boxcar	0.63	0.34	12.07	0.27	-0.14
17	**F** I.C. diffusivities	17×17	10^{-5}	200	10	3-cell boxcar	0.15	0.10	11.44	0.30	-0.17
18	**F** I.C. diffusivities	17×17	10^{-5}	500	10	3-cell boxcar	0.37	0.23	11.85	0.26	-0.17
19	**F** I.C. diffusivities	17×17	10^{-5}	1000	10	3-cell boxcar	0.70	0.41	13.24	0.24	-0.16
20	**F** I.C. diffusivities	17×17	10^{-5}	2000	10	3-cell boxcar	1.06	0.66	14.52	0.21	-0.15
21	v_{ref} I.C. noise	17×17	10^{-4}	200	10	3-cell boxcar	0.11	0.09	12.25	0.29	-0.17
22	v_{ref} I.C. noise	17×17	10^{-4}	200	10	3-cell boxcar	0.21	0.17	11.34	0.30	-0.16
23	v_{ref} I.C. noise	17×17	10^{-4}	200	10	3-cell boxcar	0.07	0.04	11.70	0.29	-0.17
24	v_{ref} I.C. noise	17×17	10^{-4}	200	10	3-cell boxcar	0.08	0.05	11.59	0.29	-0.17
25	v_{ref} I.C. noise	17×17	10^{-4}	200	10	3-cell boxcar	0.08	0.06	11.66	0.28	-0.17
26	v_{ref} I.C. noise	17×17	10^{-4}	200	10	3-cell boxcar	0.23	0.13	11.93	0.31	-0.15
27	**F** I.C. smoothing	17×17	10^{-4}	200	10	3-cell boxcar	0.06	0.04	11.37	0.29	-0.17
28	**F** I.C. smoothing	17×17	10^{-4}	200	10	5-cell boxcar	0.09	0.06	11.38	0.29	-0.17
29	**F** I.C. smoothing	17×17	10^{-4}	200	10	7-cell boxcar	0.09	0.06	11.46	0.29	-0.17
30	**F** I.C. smoothing	17×17	10^{-4}	200	10	9-cell boxcar	0.08	0.05	11.36	0.29	-0.16
31	**F** I.C. smoothing	17×17	10^{-4}	200	10	Mean of all cells	0.09	0.06	11.55	0.29	-0.17
32	w	17×17	10^{-4}	200	1	3-cell boxcar	0.14	0.10	10.98	0.28	-0.15
33	w	17×17	10^{-4}	200	10	3-cell boxcar	0.09	0.06	11.81	0.28	-0.17
34	w	17×17	10^{-4}	200	20	3-cell boxcar	0.15	0.09	11.25	0.29	-0.15
35	w	17×17	10^{-4}	200	50	3-cell boxcar	0.32	0.23	11.20	0.29	-0.14
36	w	17×17	10^{-4}	200	100	3-cell boxcar	0.16	0.10	11.18	0.29	-0.15
37	w	17×17	10^{-4}	200	200	3-cell boxcar	0.24	0.17	10.76	0.28	-0.15
38	w	17×17	10^{-4}	200	500	3-cell boxcar	0.14	0.10	11.87	0.28	-0.18
39	w	17×17	10^{-4}	200	1000	3-cell boxcar	0.25	0.20	12.91	0.28	-0.19

LIST OF FIGURES

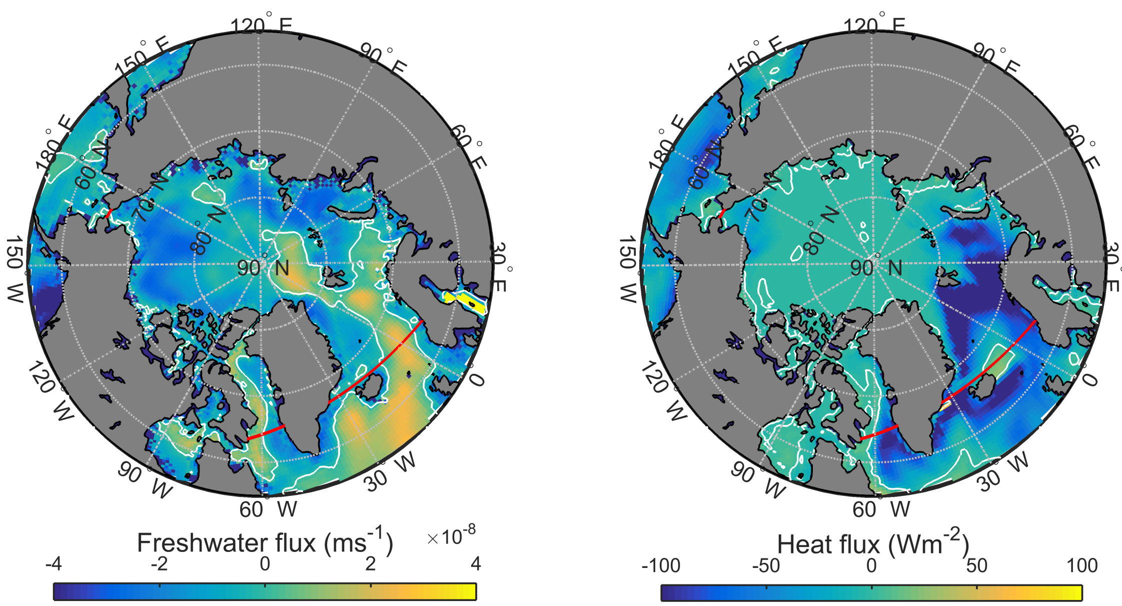


FIG. 1. Plan view of the Arctic and subpolar North Atlantic domain to which RTHIM is applied. The background colors show the 1988-2007 time-mean of the surface fluxes of freshwater (left) and heat (right) forcing NEMO. Fluxes are positive going into the ocean. The red line on each panel is the section bounding the domain, defined at a constant pseudo-latitude. The model's tripolar grid means that the section latitude varies with longitude, with the mean for the section equal to 65.6°N. The white contour is the zero line for the fluxes.

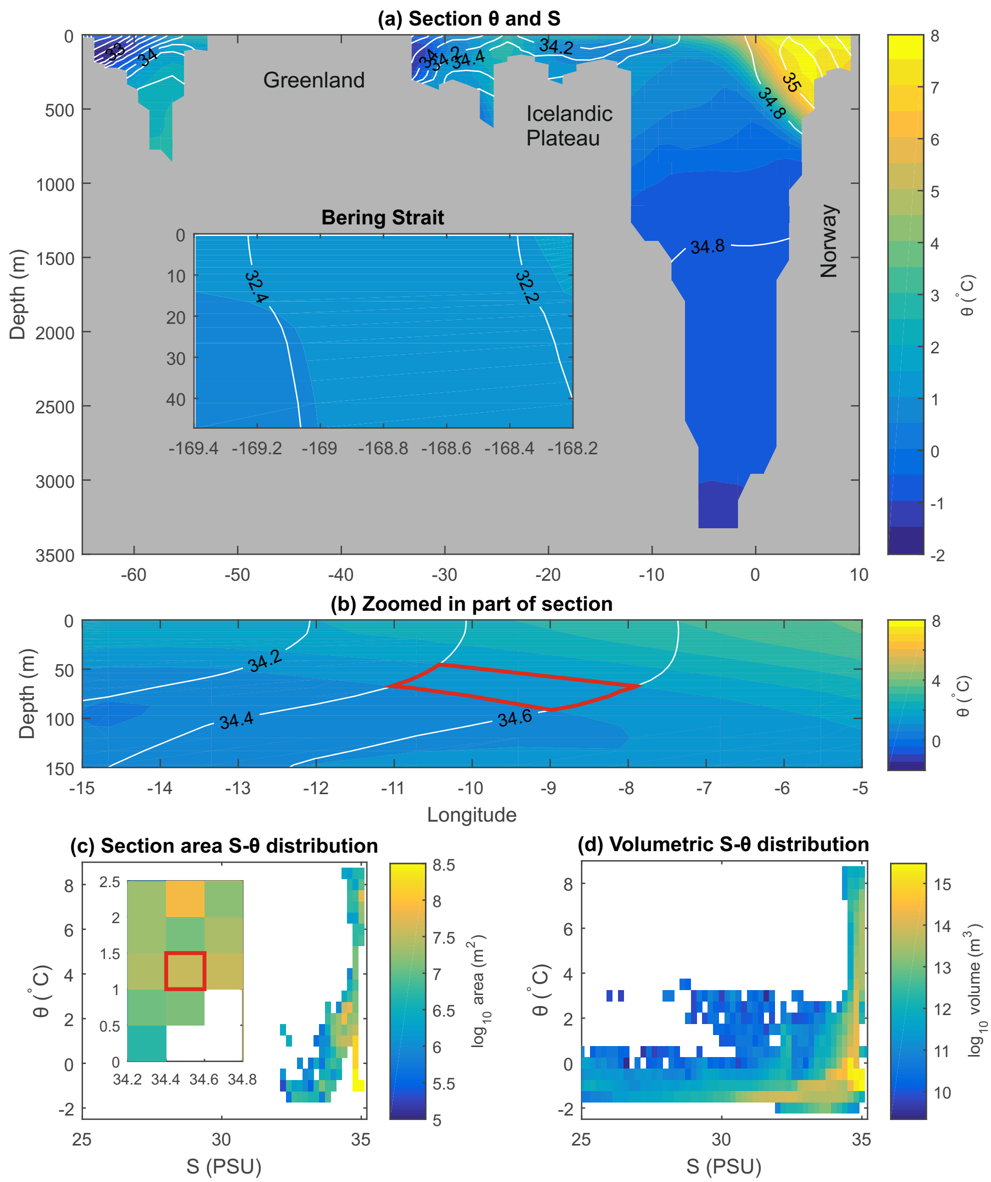


FIG. 2. (a) The 1988-2007 time-mean temperature and salinity from the section defined on Fig. 1. The background colors show temperature; the white contours show salinity. The two basins of the main plot are the Labrador Sea (left) and the Greenland-Iceland-Norway (GIN) Seas (right); the inset shows Bering Strait. (b) A zoomed in part of the section in the Western Norwegian Basin with a red polygon indicating a region enclosed by pairs of S and θ contours. (c) The distribution in $S-\theta$ space of the section area. The grid spacing is $\Delta S = 0.2$ PSU and $\Delta\theta = 0.5$ °C; the same as the contour intervals on (a) and (b). The inset shows a zoomed in part of the distribution with a red rectangle corresponding to the red polygon in (b). (d) The volumetric distribution in $S-\theta$ space for the domain bounded by the section (note that the color scales on (c) and (d) are logarithmic). A more detailed description of the domain is given in section 3a.

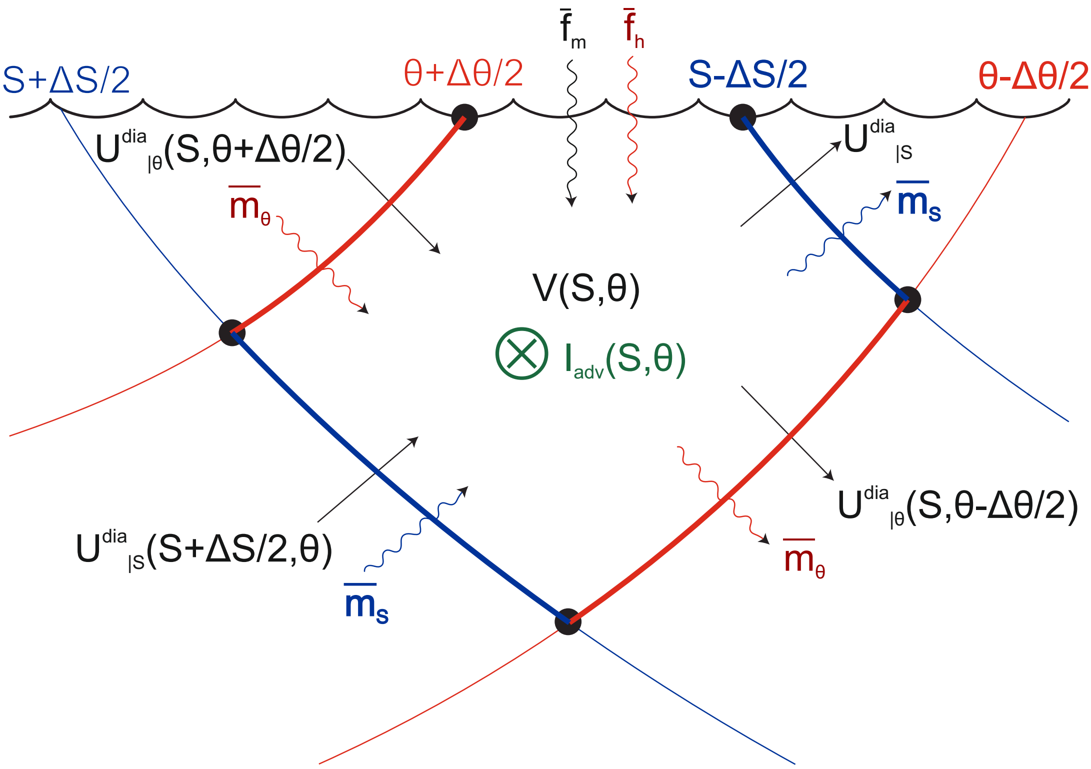


FIG. 3. A volume element V illustrating the RTHIM volume budget. The element can be defined between pairs of isotherms $(\theta \pm \frac{\Delta\theta}{2})$ and isohalines $(S \pm \frac{\Delta S}{2})$, such as is indicated by the polygon on the inset of Fig. 2(a). $I_{adv}(S,\theta)$ is the advective flux through the section into the volume element per unit S,θ; $\overline{f}_m$ and $\overline{f}_h$ are boundary fluxes of mass and heat; and $\overline{m}_S$ and $\overline{m}_\theta$ are diffusive fluxes of salt and heat. The change in volume of the element depends on the movement of its bounding surfaces, $\mathbf{U}^{dia}$, and is governed by Eq. 9. Figure adapted from Groeskamp et al. (2014b).

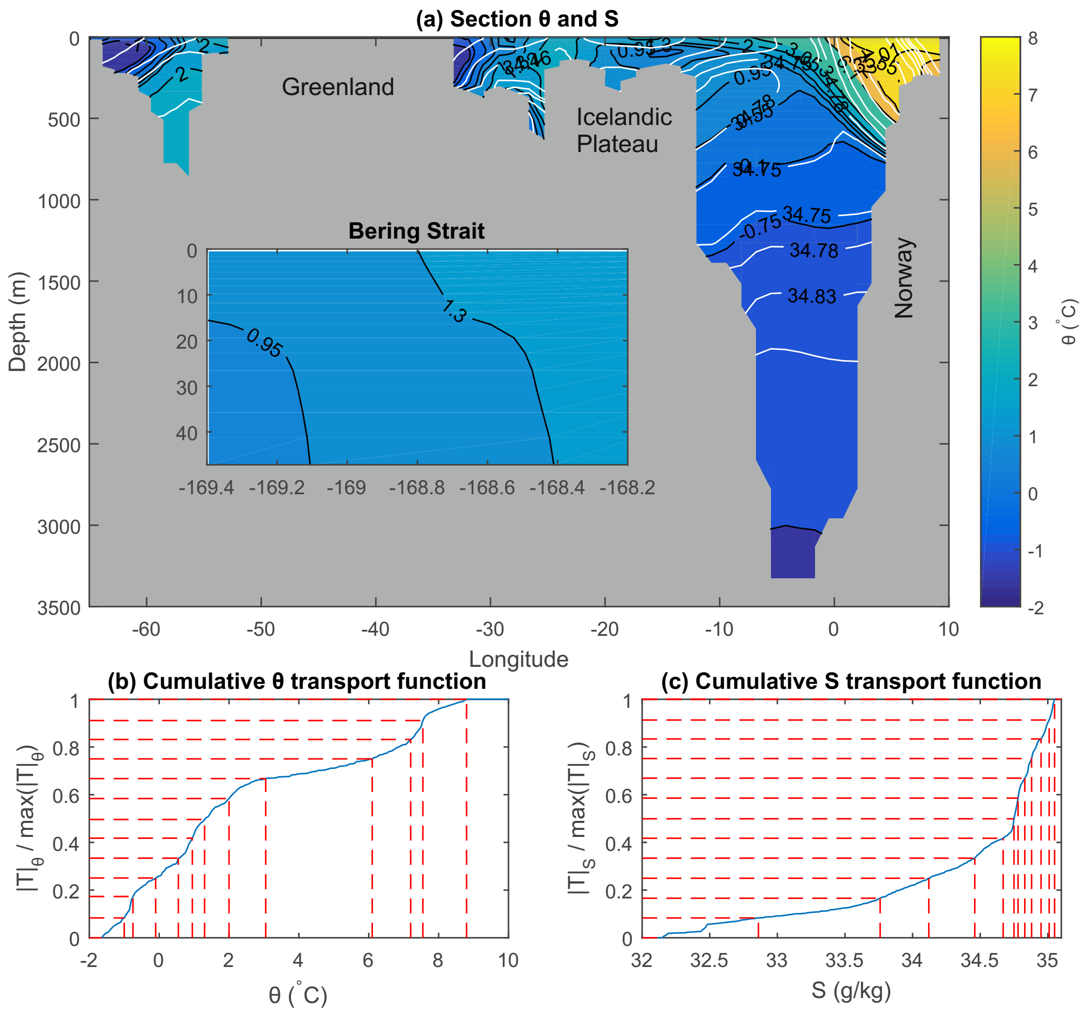


FIG. 4. (a) The section θ (colors and black contours) and S (white contours) from the NEMO simulation time-mean. This is the same as Fig. 2 (a) except that the contour intervals have been determined for use in the RTHIM $S-\theta$ grid using the transport functions in (b) and (c) as detailed in section 3b. The blue lines on (b) and (c) represent the normalized transport functions, and the red dashed lines indicate the selection of the θ and S grid nodes. In this case the transport functions have been divided up into 12 approximately equal segments; however we can choose to construct a grid using more or fewer nodes. The grids constructed from these plots, including additional nodes at either end to ensure binning of the whole range of (S,θ) found within the control volume, are detailed in the appendix.

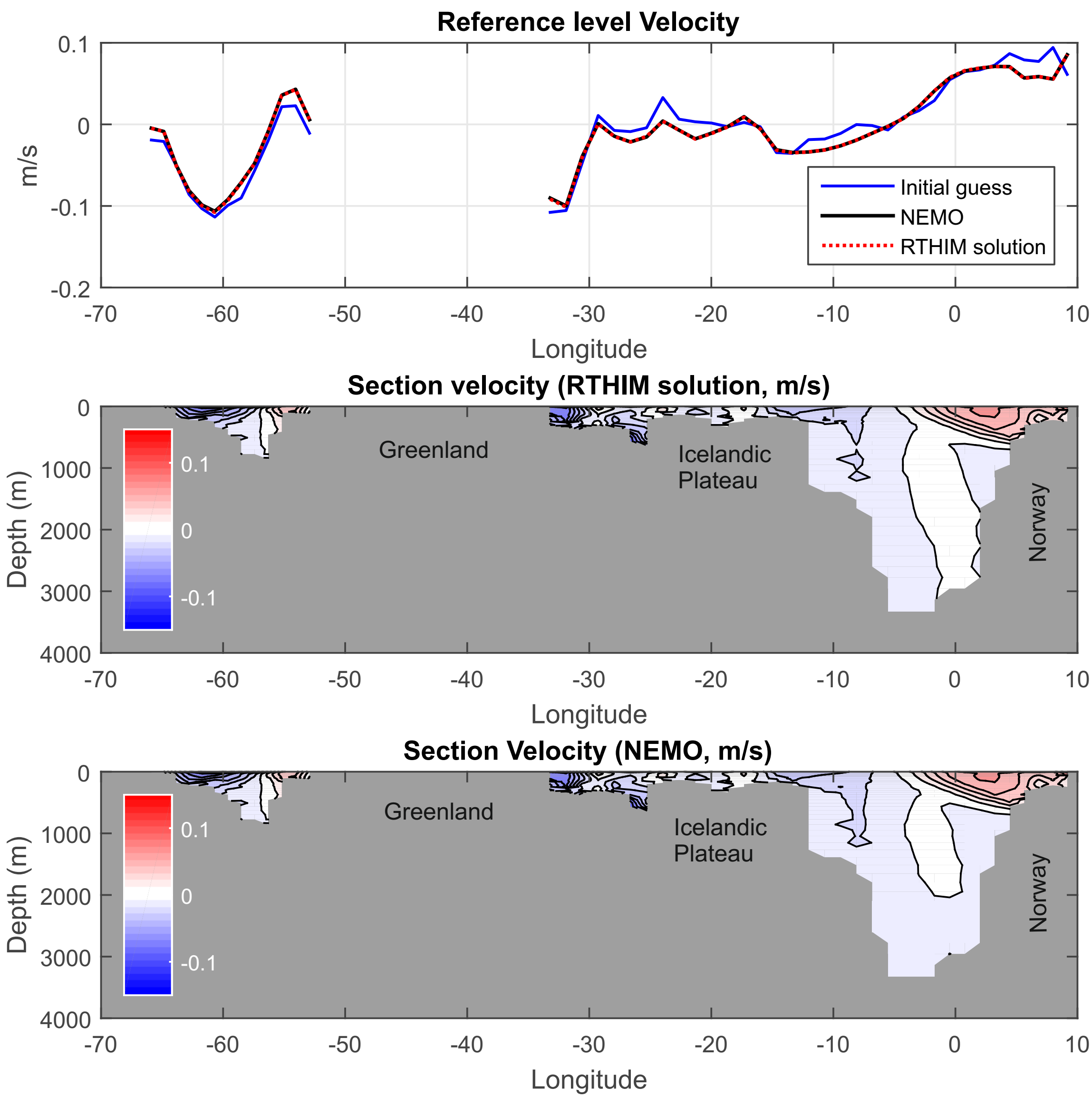


FIG. 5. Velocities corresponding to the section identified by the red lines in Fig. 1. The top panel shows surface v_{ref} from RTHIM solution (red) against the NEMO model 'truth' (black) with the RTHIM initial condition in blue. The middle panel shows full section velocities from the RTHIM solution obtained by adding the known model shear to the solution v_{ref}. The bottom panel shows full section velocities from NEMO.

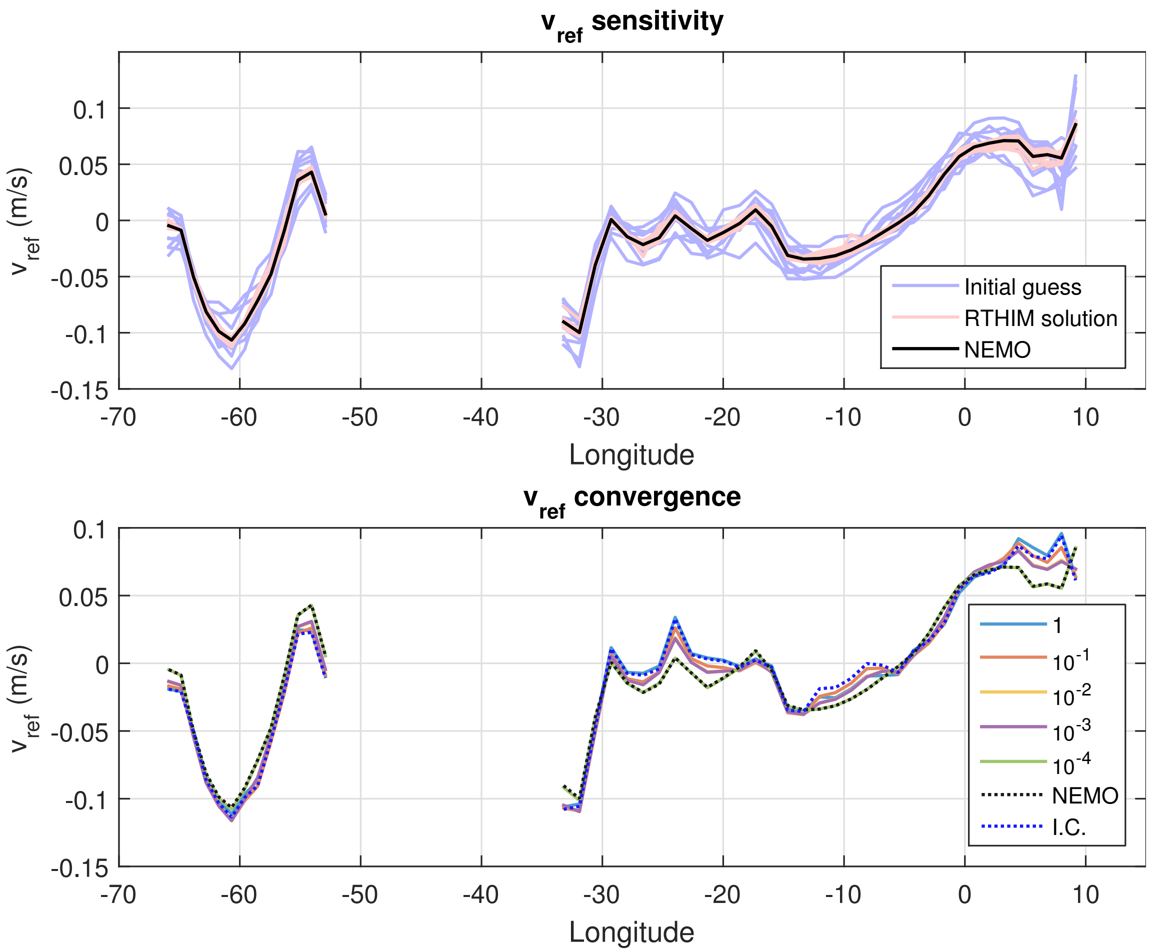


FIG. 6. The top panel shows v_{ref} from an ensemble of RTHIM solutions (red lines) with varying initial conditions (blue lines) and varying model parameters (described in section 4b), compared with the model truth (black line). The bottom panel shows the convergence of v_{ref} from its initial condition (blue dotted line) towards the model truth (black dotted line). Each solid line is an RTHIM solution with a different convergence tolerance applied to the optimization algorithm.

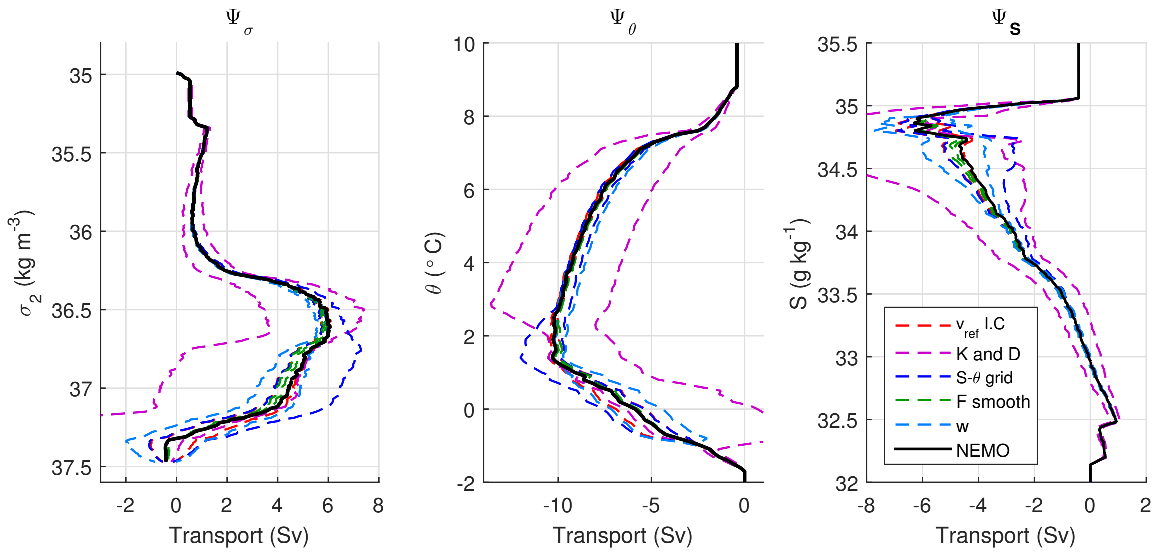


FIG. 7. Streamfunctions in σ (left), θ (middle) and S (right) coordinates calculated from RTHIM solution section velocities from ensembles varying a range of parameters. The same-colored pairs of dashed lines give the upper and lower limits at each point on the y-axis of the streamfunctions for an ensemble. The parameters varied were the v_{ref} initial condition (red lines), the isopycnal and diapycnal mixing coefficients applied to the $\boldsymbol{F}$ initial condition (purple lines), the $S-\theta$ grid used for binning the terms in the volume budget (dark blue lines), the smoothing factor applied to the $\boldsymbol{F}$ initial condition (green lines), and the weighting factor on the net transport constraint (light blue lines). The streamfunctions calculated from the NEMO section velocities are shown in black, and a black dotted zero line is also plotted for reference. Note that the x-axis ranges have been set to assist the reader in discerning details of the contributions of different parameters to the uncertainty, for which it is necessary not to include the full range of the K and D lines. Their full extent can be seen in the solid red lines on Fig. 8. For Ψ_σ the lower bounds have contributions from RTHIM runs runs 6, 12 and 13 detailed in Table 2 in the Appendix; the upper bounds have contributions from runs 2 and 13. For Ψ_θ the lower bounds have contributions from runs 2 and 13; the upper bounds have contributions from runs 6, 11 and 13. For Ψ_S the lower bound comes from run 13 and the upper bound from run 5.

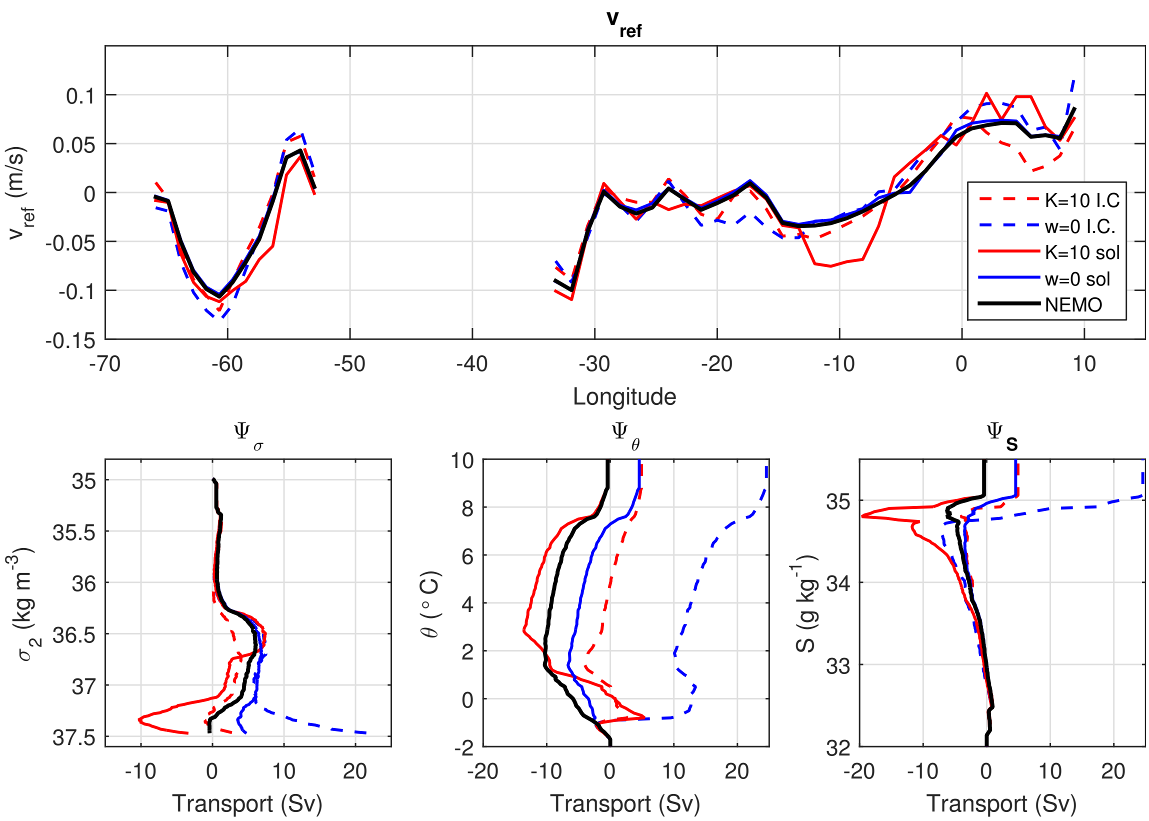
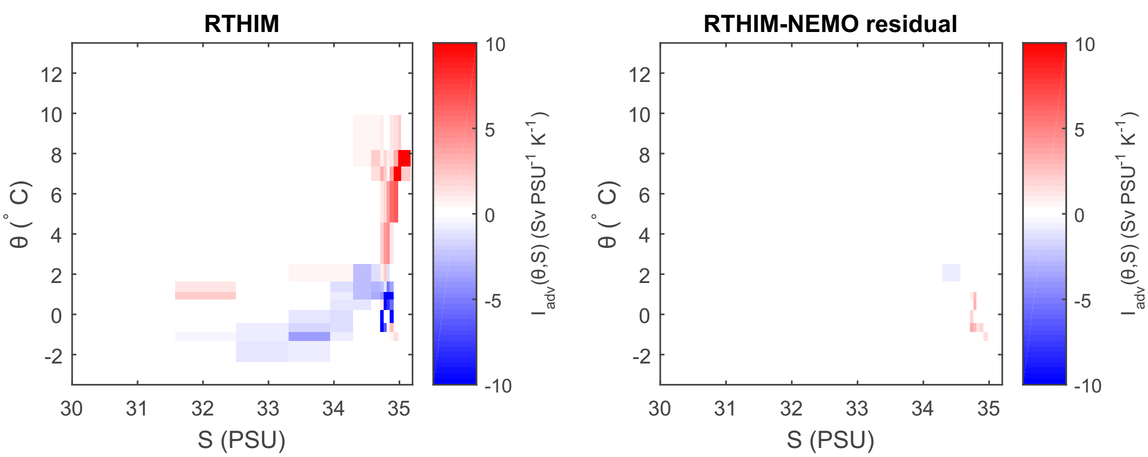


FIG. 8. Two RTHIM runs where the selection of model parameters has significantly degraded the solution. In red is the result of setting $K = 10$ m^{2}s^{-1} in the calculation of the $\boldsymbol{F}$ initial condition. In blue is the result of removing the net transport constraint. The top panel shows v_{ref} for the initial condition (dashed lines) and solution (solid lines) for each run. The bottom 3 panels show the streamfunctions in density (left), temperature (middle) and salinity (right) coordinates.



928 FIG. 9. The advection term I_{adv} plotted in $S-\theta$ coordinates from an RTHIM solution (left panel) and the

929 residual with the same term based on the NEMO section velocities (right panel).

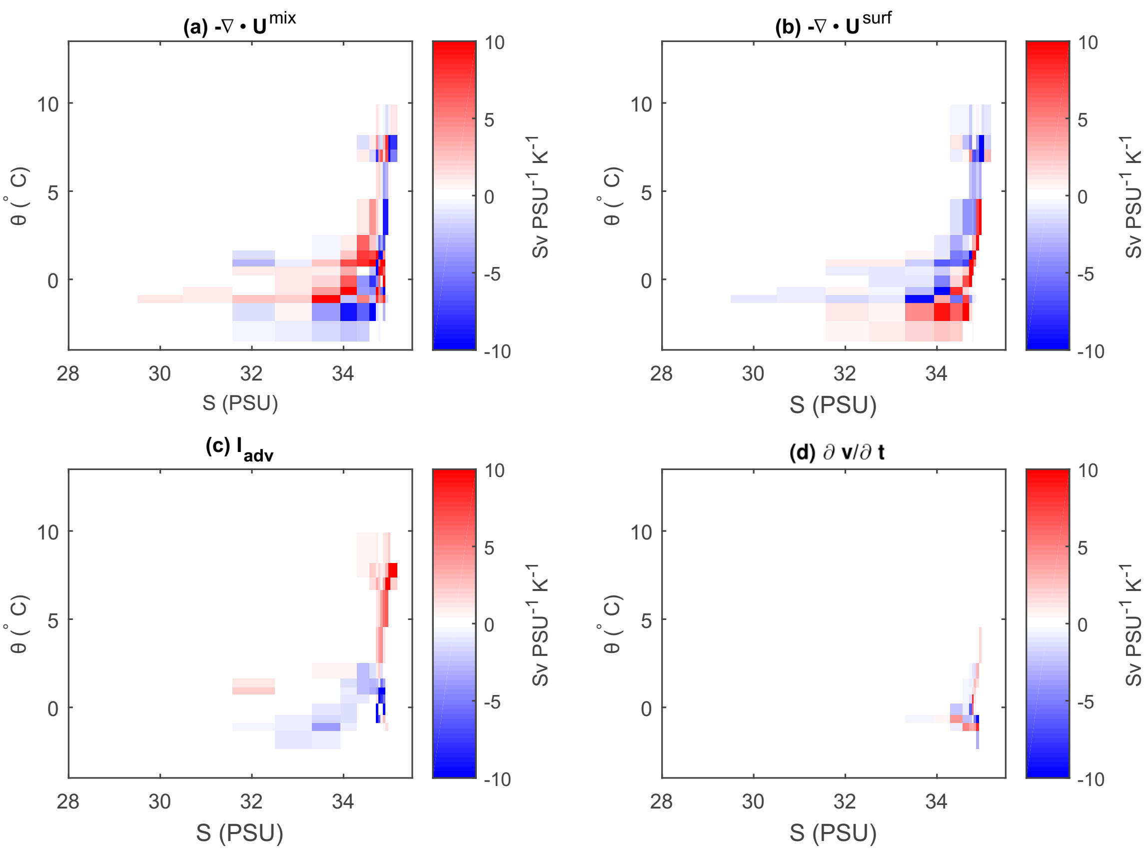
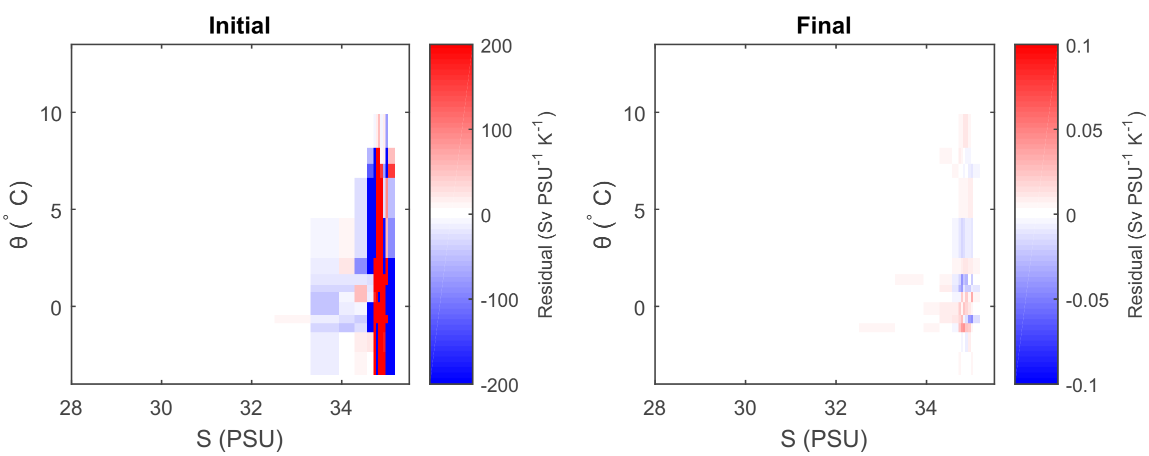


FIG. 10. Each term from the volume budget of Eq. 11, plotted in $S-\theta$ coordinates, from an RTHIM solution. The terms are (a) the mixing term, (b) the surface flux term (c) the advection term and (d) the volume trend. Red colors indicate a net positive contribution to the volume of water in that (S,θ) bin by a given process; blue colors indicate a net negative contribution.



934 FIG. 11. The residual, ε, in the volume budget in $S-\theta$ coordinates from the RTHIM initial condition (left)
935 and after optimization (right). Note the different color scales.

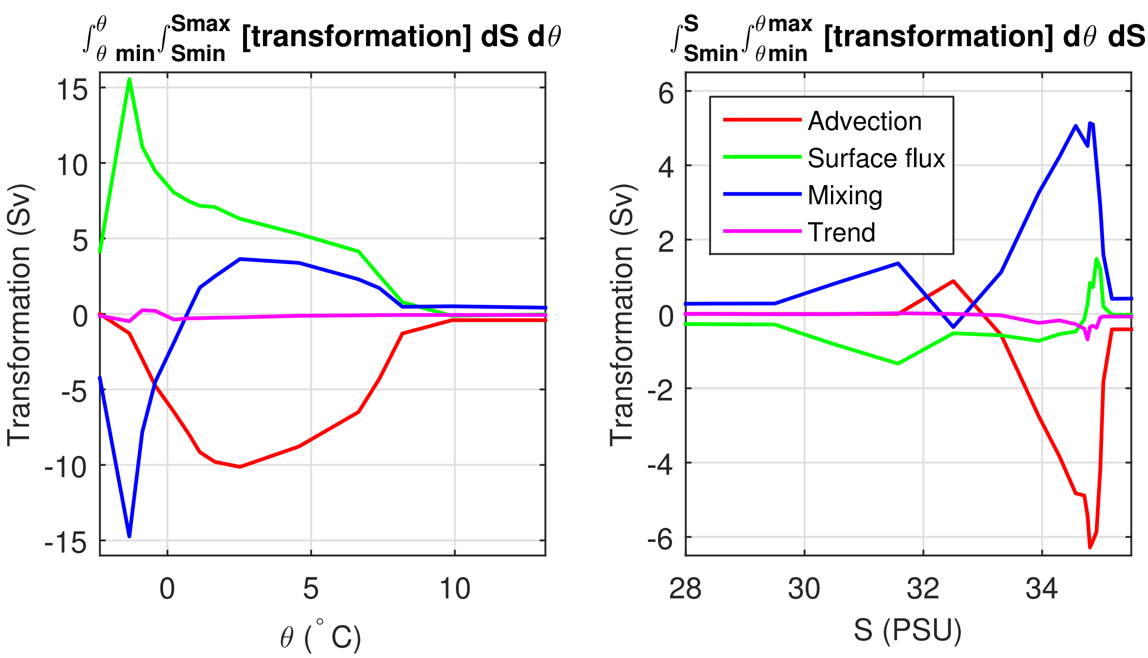


FIG. 12. Terms in the volume budget integrated in $S-\theta$ space. On the left panel, each term has been integrated first through all S and then cumulatively in θ, and plotted against θ. On the right panel, each term has been integrated first through all θ and then cumulatively in S, and plotted against S.

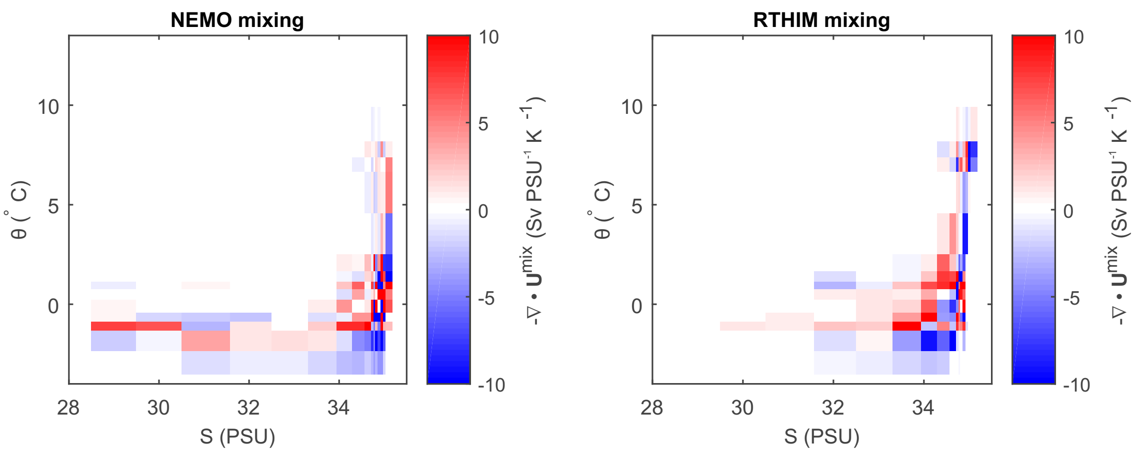


FIG. 13. Mixing terms diagnosed directly from NEMO tendencies (left panel) and from an RTHIM solution (right panel). Red colors indicate a net positive contribution to the volume of water in that (S, θ) bin; blue colors indicate a net negative contribution.

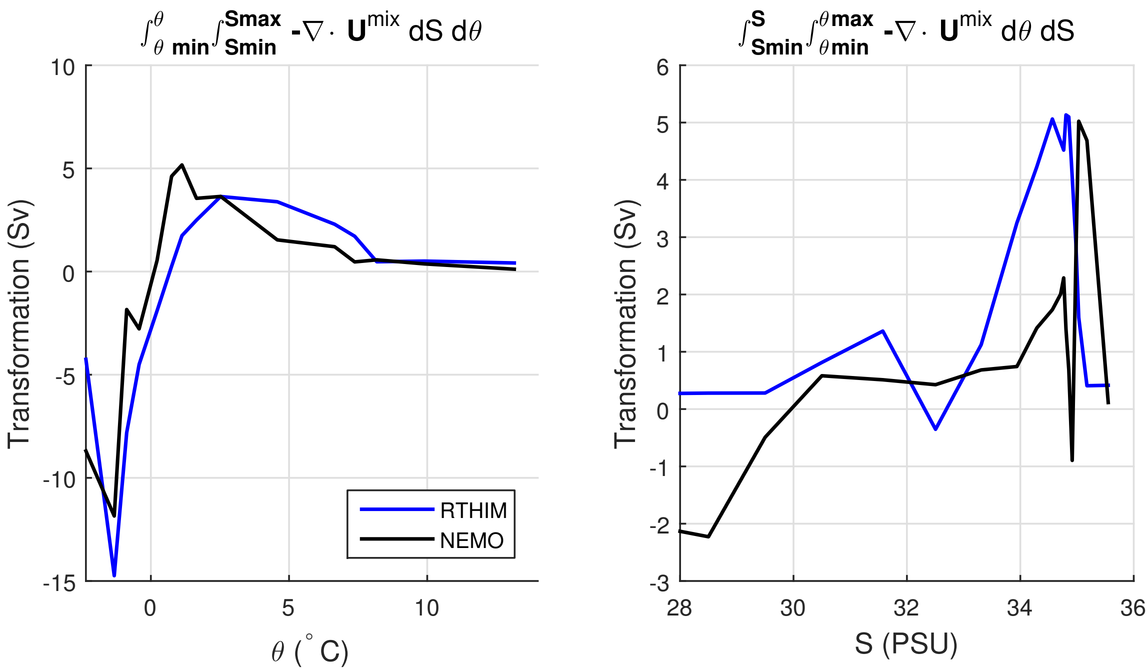


FIG. 14. Mixing terms from an RTHIM solution (blue line) and NEMO (black line) integrated in $S-\theta$ space. On the left panel, each term has been integrated first through all S and then cumulatively in θ, and plotted against θ. On the right panel, each term has been integrated first through all θ and then cumulatively in S, and plotted against S.