The observation-based application of a Regional Thermohaline Inverse Method to diagnose the formation and transformation of water masses north of the OSNAP array from 2013-2015

Neill Mackay*

National Oceanography Centre

Chris Wilson and N. Penny Holliday

National Oceanography Centre

Jan D. Zika

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

**Corresponding author address:* College of Life and Environmental Sciences, University of Exeter, Prince of Wales Road, Exeter, UK.

E-mail: n.mackay@exeter.ac.uk

Early Online Release: This preliminary version has been accepted for publication in *Journal of the Physical Oceanography*, may be fully cited, and has been assigned DOI 10.1175/JPO-D-19-0188.1. The final typeset copyedited article will replace the EOR at the above DOI when it is published.

© 2020 American Meteorological Society

ABSTRACT

The strength of the Meridional Overturning Circulation (MOC) in the North Atlantic is dependent upon the formation of dense waters that occurs at high northern latitudes. Wintertime deep convection in the Labrador and Irminger Seas forms the intermediate water mass known as Labrador Sea Water (LSW). Changes in the rate of formation and subsequent export of LSW are thought to play a role in MOC variability, but formation rates are uncertain and the link between formation and export is complex. We present the first observation-based application of a recently developed Regional Thermohaline Inverse Method (RTHIM) to a region encompassing the Arctic and part of the North Atlantic subpolar gyre for the years 2013, 2014 and 2015. RTHIM is a novel method which can diagnose the formation and export rates of water masses such as the LSW identified by their temperature and salinity, apportioning the formation rates into contributions from surface fluxes and interior mixing. We find LSW formation rates of up to 12 Sv during 2014-15, a period of strong wintertime convection, and around half that value during 2013 when convection was weak. We also show that the newly convected water is not exported directly, but instead is mixed isopycnally with warm, salty waters that have been advected into the region, before the products are then exported. RTHIM solutions for 2015 volume, heat and freshwater transports are compared with observations from a mooring array deployed for the Overturning in the Subpolar North Atlantic Program (OSNAP) and show good agreement, lending validity to our results.

Accepted for publication in *Journal of Physical Oceanography*. DOI 10.1175/JPO-D-19-0188.1.

1. Introduction

a. The subpolar North Atlantic

The meridional overturning circulation (MOC) of the ocean is characterised in the North Atlantic by a northward flow of warm, salty waters in the upper layers and a compensating southward flow of cooler, fresher waters at depth (see Fig. 1). The transformation into denser waters occurs at high latitudes, where heat is lost to the atmosphere and freshwater added, and it has been proposed that the strength of the MOC is linked to the rate of production of these dense waters (Lozier 2012). Observations suggest that rates of dense water formation are fairly constant in the Nordic Seas, but more variable in the Labrador and Irminger Seas (Smeed et al. 2014), and climate models have indicated a link between changes in Labrador and Irminger Sea convection and the MOC (e.g. Zhang 2010; Danabasoglu et al. 2012); however direct observational evidence for this link is lacking (Lozier et al. 2017). Paleo-oceanographers have also linked deep convection in the Labrador Sea with the strength of the AMOC over the last 1500 years using proxies (Thornalley et al. 2018). The formation of dense waters in the Labrador and Irminger basins is therefore a subject of interest.

The Overturning in the Subpolar North Atlantic Program (OSNAP) has been continuously monitoring the flow through the section shown on Fig. 1 since August 2014. OSNAP aims to quantify the strength and variability of the MOC and heat and freshwater transports through the section, using a combination of moored instruments and glider surveys described in Lozier et al. (2017). OSNAP results reported by Lozier et al. (2019) (hereafter L2019) suggested that despite the deployment coinciding with a period of strong deep convection in the Labrador Sea, the MOC variability was dominated by changes in the Irminger and Iceland basins. The present study complements the OSNAP array observations with new information about the processes that transform the

Accepted for publication in *Journal of Physical Oceanography*. DOI 10.1175/JPO-D-19-0188.1.

waters north of the section, while also providing independent estimates of the section transports. A key test of our results is that they are validated by the OSNAP observations.

b. Labrador Sea Water

Labrador Sea Water (LSW) is an intermediate water mass formed when convection creates mixed layers of cold, fresh water as deep as 1500m in winter (Yashayaev 2007, hereafter Y2007a). It is formed predominantly in the Labrador Sea but also occasionally in the Irminger Sea (Pickart et al. 2003; Fröb et al. 2016), and can be found at mid depths throughout the North Atlantic north of 40°N and further south along the western boundary. It intersects with high salinity lower Mediterranean water and mixes isopycnally to produce the upper North Atlantic Deep Water (NADW), thereby contributing to the upper cell of the MOC (Talley and McCartney 1982).

In a review of 45 years of observations, Y2007a described the LSW and the processes behind its formation and transformation over its life cycle. Over a period of years, repeated convection events in winter form a class of LSW identifiable by its temperature (T) and salinity (S) properties (see e.g. Y2007a Figs. 6 and 7), which then evolves with time. The T and S of a given LSW class depends on the conditions that led to its formation, and the atmospheric forcing both throughout the year of formation and in previous years play a role in preconditioning the ocean for convection. For example, heating and freshwater flux in summer increases surface stratification which works against convection the following winter; on the other hand deep convection one winter homogenises the water column meaning that it occurs more easily during the next. The role of preconditioning in convection was confirmed by Yashayaev and Loder (2017) (hereafter YL2017). During convection and after it finishes, the newly formed LSW mixes isopycnally with warm saline intermediate waters arriving into the Labrador Sea in the boundary current that flows westward around the southern tip of Greenland (Fig. 1). Gradually the LSW drains away to other

Accepted for publication in *Journal of Physical Oceanography*. DOI 10.1175/JPO-D-19-0188.1.

parts of the North Atlantic, being replaced by lighter waters as they flow into the Labrador Sea. Much of the water also recirculates within the Labrador and Irminger Seas, steadily mixing with warmer and saltier waters, thereby maintaining a steady transfer of heat and salt into the LSW. As it spreads eastwards across the Irminger sea, modified LSW can also come into contact with North East Atlantic Deep Water (NEADW) spreading west. The NEADW is modified Iceland-Scotland Overflow Water (ISOW) that mixes with warm/salty older NEADW, with DSOW, and with the eastward spreading LSW before it is advected into the Labrador Sea (Yashayaev and Dickson 2008).

The formation rate of LSW is variable interanually and estimates of its mean also vary: Pickart and Spall (2007) compiled estimates of 1-8.5 Sv from studies between 1972-2003; Y2007a reported an average annual formation rate of 2 Sv for 1970-1995, with a higher rate of 4.5 Sv for 1987-1994; Haine et al. (2008) summarised estimates of 1.3-12.7 Sv from a range of studies; and LeBel et al. (2008) suggested a long-term formation rate of 11.9 Sv based on CFC-11 inventories. There is also conflicting evidence of the link between deep convection (and subsequent LSW formation) and the rate of LSW export from the subpolar gyre. Schott et al. (2004) found similar export rates in two different time periods with significantly different amounts of convection; on the other hand Yashayaev and Loder (2016) (hereafter YL2016) found LSW export rates between 2002-2015 were larger in strong convection years. The main export pathway for LSW has been assumed to be in the Deep Western Boundary Current, but recent studies using real (Bower et al. 2009, 2011) and simulated (Gary et al. 2011; Zou and Lozier 2016) Lagrangian floats have found that much, if not most LSW is exported via interior pathways, and that there is significant recirculation within the subpolar gyre. In fact the exchange of LSW between the Labrador and Irminger Seas was noted in Y2007a as a "broad well-established communication" between the two basins.

Accepted for publication in *Journal of Physical Oceanography*. DOI 10.1175/JPO-D-19-0188.1.

c. The ocean in thermohaline coordinates

Water masses such as the LSW may be defined in various ways, e.g. as water within a T-S class or density class. Transformation of water masses occurs when water moves between the relevant defined classes, via a flux of volume. The transformation may lead to water mass formation (destruction), when there is a net increase (decrease) of volume in a particular class. Fig. 2 shows a volumetric distribution in TS space for the Arctic/subpolar North Atlantic ocean volume used in this study, with the definitions of various important water masses used in our analysis indicated by black boxes. The warm, saline water mass at the top right of the plot is the North Atlantic Water (NAW). Just cooler than the NAW are the Labrador Sea Water (LSW) and Overflow Water (OW), the latter of which includes the Iceland-Scotland Overflow Water and Denmark Strait Overflow Water; these water masses collectively occupy a small region of TS space but considerable volume. Just saltier than these is Arctic/Atlantic Water (AAW, see e.g. Polyakov et al. 2004; Shu et al. 2019). The bottom of the plot has the cold, fresh Arctic Surface Water (ASW) and the cold, salty Arctic Deep Water (ADW); the latter being the densest and most voluminous water mass.

There are a number of processes which transform water masses, moving them around in TS space. These processes and their effects are summarised on Fig. 3. The tendency of surface fluxes to increase the spread of water masses in TS space is counteracted by the tendency of mixing to bring them together. In this work, we utilise this competition to determine the relative roles of the different processes involved in water mass formation and transformation.

The Regional Thermohaline Inverse Method (RTHIM) was developed to investigate water mass transformation north of the OSNAP section, and to diagnose the relative roles of the transforming processes for each water mass. Using inputs of surface fluxes of heat and freshwater, conservative temperature (Θ) and absolute salinity (S_A) along the section, and initial estimates of interior

Accepted for publication in *Journal of Physical Oceanography*. DOI 10.1175/JPO-D-19-0188.1.

mixing and section velocities, RTHIM solves for the section flow and mixing within the control volume bounded by the section. The method was validated by Mackay et al. (2018) (hereafter M2018) using model data and a control volume bounded by a simple section on the model grid. In this study we define a volume-bounding section that follows the OSNAP section and includes Bering Strait and apply RTHIM to gridded observations from 2013-2015. These years overlap with the OSNAP observations, allowing independent validation of the section velocities component of our solution for 2015. The period also coincides with the development of a new LSW class between 2012-2016 reported by YL2016, which was one of the deepest and thickest LSW layers observed since the 1990s. There are contrasts in atmospheric forcing over the three years, with 2013 seeing relatively weak convection and 2014-15 relatively strong convection in the Labrador Sea. RTHIM will enable us to diagnose the formation (or destruction) and export rates of LSW in each year, and partition that formation (destruction) into contributions from surface fluxes and interior mixing. In the results presented, rates of formation or destruction of a water mass will be referred to generically as ‘formation rates’, where negative values are associated with destruction. Previous estimates of LSW formation rates have tended to be from inverse box models, inferred from air-sea fluxes or tracer inventories, or using the thickness of an LSW layer identified using hydographic sections as a proxy for its volume (note that YL2016 were more sophisticated, combining ship and float data to analyse the evolution of LSW week by week). By combining RTHIM with observational products that use a wealth of recent observations from remote sensing and autonomous profiling floats, we can improve upon these estimates. RTHIM combines a number of the best aspects of previous methods. It conserves volume like a box model, but also implicitly conserves heat and salt. It uses the available surface fluxes while ensuring consistency with the conservation of volume, heat and salt. And finally it apportions the water mass transformation into contributions from the surface flux and interior mixing while imposing physically realistic con-

Accepted for publication in *Journal of Physical Oceanography*. DOI 10.1175/JPO-D-19-0188.1.

straints on that mixing (see section 2 and Appendix D, or M2018 for full details). We will capture the episodic nature of the LSW formation and export by analysing three individual years of data. In addition, Y2007a suggest that their LSW production rates are likely to be underestimates as they do not account for the loss by mixing and entrainment during and shortly after convection; our solutions account for these losses.

The paper is organised as follows: in section 2 we summarise the Regional Thermohaline Inverse Method and detail how we have applied it to observation-based data. In section 3 we compare the section transports from the RTHIM solutions with observations from the OSNAP array. In section 4 we examine the formation rates for important water masses and the relative roles of the processes effecting this formation. Section 5 discusses our results and section 6 presents a summary and conclusions.

2. Regional Thermohaline Inverse Method

a. Method theory

The full details of RTHIM and its validation are laid out in M2018. Here we summarise the method and describe how we have applied it to observations in this study. We begin by defining an ocean control volume consisting of the Arctic and part of the NASPG, bounded to the south by the circumpolar section shown on Fig. 4(a). This section is chosen because it coincides with the OSNAP array, giving us the opportunity of comparing our inverse model solutions with the OSNAP observations. We can subdivide the section into areas enclosed by isotherms and isohalines (Fig. 4(b)), which project into the volume and may outcrop. We then consider a volume element V between a pair of isotherms and a pair of isohalines, the volume of which will change when these isosurfaces move. The rate of volume change $\frac{\partial V}{\partial t}$ of the element is governed by the flow I_{adv}

Accepted for publication in *Journal of Physical Oceanography*. DOI 10.1175/JPO-D-19-0188.1.

through the section in between the isosurfaces, surface fluxes of heat and freshwater in between the same isosurfaces where they outcrop, and interior mixing across the element isosurfaces within our control volume. This volumetric balance can be expressed in thermohaline coordinates as:

$$I_{adv} - \nabla^2_{S\theta}\boldsymbol{F} = \frac{\partial V}{\partial t} + \nabla_{S\theta} \cdot \boldsymbol{U}^{surf}_{S\theta} + \varepsilon, \qquad (1)$$

where $\nabla^2_{S\theta}\boldsymbol{F}$ is the mixing term, $\nabla_{S\theta} \cdot \boldsymbol{U}^{surf}_{S\theta}$ is the surface flux term and ε is the error in the inverse solution which we minimise. The operator $\nabla_{S\theta}\cdot$ is the divergence operator in thermohaline coordinates. The full derivation of this equation is given in M2018, and some more details of the terms can be found in Appendix D; we note here only that the diffusive flux tensor $\boldsymbol{F}$ is constrained such that it is symmetric and that mixing acts down tracer gradients. The fact that no spatial structure is imposed on the interior mixing is a unique advantage of RTHIM over other inverse methods such as the Tracer Contour Inverse Method (Zika et al. 2010) which has uniform isopycnal K and diapycnal D diffusivities on each neutral density surface, or the Thermohaline Inverse Method (Groeskamp et al. 2014b) which has uniform K and imposes a horizontally uniform vertical structure on D.

In RTHIM, the advection term I_{adv} is obtained similarly to a box inverse method (e.g. Wunsch 1978), by solving a reference level velocity v_{ref} (we use the surface) and imposing a velocity shear to reconstruct the velocities for the whole section (see section 2d). The volume trend and surface flux terms are also imposed, calculated from the time-evolving Θ, S_A, and surface flux fields, also described in the next section. For full details of the calculation of the terms in Eq 1, the reader is referred to M2018.

Accepted for publication in *Journal of Physical Oceanography*. DOI 10.1175/JPO-D-19-0188.1.

b. Datasets

We have applied RTHIM to two distinct datasets as a means of exploring the uncertainty in our inverse solutions that is due to uncertainties in the input data. The additional uncertainty due to choices of inverse model parameters is also explored (see section 3b). We apply RTHIM to each dataset for the years 2013, 2014 and 2015, giving six sets of results.

The first dataset is ECCO v4r3, a dynamically consistent time-evolving ocean state estimate with known heat and buoyancy forcing (Forget et al. 2015; Fukumori et al. 2017), which includes all of the variables required by RTHIM: surface heat and freshwater fluxes, time-evolving T and S for the control volume, and sea surface height (ssh) for the calculation of an initial geostrophic surface v_{ref}. ECCO v4r3 has not been constrained by any observations collected for OSNAP. The second dataset uses a combination of products. T and S fields are obtained from EN4 version 4.2.1, an objective analysis of quality controlled T and S profiles (Good et al. 2013); surface fluxes of heat and freshwater are from the ERA-Interim reanalysis (Dee et al. 2011); and ssh are from AVISO satellite altimetry[1]. In what follows, 'EN4 solution' or 'EN4 dataset' in the context of our RTHIM results refers to this combination of products. Of course, since ECCO uses the available observations to calculate its state estimate, the two datasets are not independent; however they are different enough to help in the exploration of uncertainty.

The ECCO dataset acts as a good test bed for the application of RTHIM to real-world observations, because its time-evolving fields must obey the dynamical equations of the general circulation model on which it is based, including the conservation of volume, heat and salt consistent with

[1] for this study, SSALTO/DUACS (www.aviso.altimetry.fr) all-satellite, merged, DUACS2014 absolute dynamic gopography (MADT-H) is used. The dynamic topography is provided on a 1/4 deg. latitude-longitude grid, at daily interval [citation: Pujol, M.-I., and SL-TAC Team (2017). Copernicus marine environment monitoring service quality information document (Tech. rep. CMEMS-SL-QUID-008–032-051). EU Copernicus. France: Mercator Ocean, Ramonville Saint-Agne. Retrieved from http://cmems-resources.cls.fr/documents/QUID/CMEMS-SL-QUID-008-032-051.pdf]

Accepted for publication in *Journal of Physical Oceanography*. DOI 10.1175/JPO-D-19-0188.1.

its surface forcing. This strength is also a weakness however when it comes to making inferences about the real ocean from the application of RTHIM to the ECCO fields, since the ability of the model to represent the real ocean is constrained by its resolution and by any limitations in its subgrid-scale parameterisations. It is therefore expected that the model will depart from available observations to some degree in order that the fields remain dynamically consistent (see Forget et al. (2015) for an explanation of how they minimise the model-data misfit and Carton et al. (2019) for a comparison of ECCOv4r3 with temperature and salinity observations). By contrast, we expect the combined dataset of EN4/ERA-Interim/AVISO to most closely match the available observations; however where observations are absent EN4 is relaxed to climatology, and consequently Good et al. (2013) urge caution when using the dataset to diagnose trends. Given the sparsity of observations under the polar ice cap in particular, this must be taken into consideration when interpreting our RTHIM solutions. Away from Argo float coverage, EN4 relies on hydrographic section data which are necessarily sparse in time, therefore the seasonal evolution of the T and S fields is unlikely to be resolved in these regions. In addition, while in the depth range of the Argo floats there are plentiful observations contributing to EN4, their number is considerably reduced below 2000m. Finally there are some known issues with the EN4 fields high in the Arctic which result in some unphysical spatial distributions of T and S where it seems that sparse point observations of the North Pole Environmental Observatory (NPEO) may have been too heavily weighted (A.T Blaker, personal communication). We explore the sensitivity of our solutions to uncertainties in the EN4 dataset in sections 3 and 4.

c. TS grid design

We have identified a number of specific water masses on which we focus our analysis, defined according to their TS properties. Taking the volumetric TS distribution of water masses in the

Accepted for publication in *Journal of Physical Oceanography*. DOI 10.1175/JPO-D-19-0188.1.

control volume from each dataset, we draw boxes around each water mass in TS space as shown on Fig. 2 for EN4 and on Fig. A1 for ECCO (see Table A1 for water mass definitions). We are restricted by our inverse model grid to the use of rectangles in TS space; this has some implications at the boundaries between the Labrador Sea Water and Overflow Waters which will be discussed later. The total volume of the water in all our defined water masses represents 95% of the whole control volume for EN4, and 98% for ECCO. Having established our water mass boundaries in TS space, we then construct TS grids which allow us to determine formation rates for those water masses by integrating terms in the RTHIM solutions over the relevant parts of TS space. Details of these grids are in Appendix B.

We have chosen a simple approach to defining water masses based on their volumetric TS distributions for ease of comparison across different datasets when evaluating the formation rates (section 4). In the case of the LSW, the focus of this paper, there is some variation in the volumetric peak between datasets and over different years (Fig. A2). Our LSW definition includes some water that is colder than the temperature range suggested by the volumetric peaks for 2013 and 2014, but allows for a consistent definition across all three years and both datasets. We have carried out some investigations using alternative definitions of the LSW and found that regions of TS space with low volume do not contribute significantly to the formation rates, and consequently do not impact our results. We also note that our LSW definition is a general one for the whole of our control volume, and therefore the volumetric peaks in subregions of the volume (such as the Labrador Sea, for example) are likely to be different. The use of a thermohaline coordinate system necessitates this approach, but has the advantage that we make no assumptions about where in our control volume a water mass is located.

Accepted for publication in *Journal of Physical Oceanography*. DOI 10.1175/JPO-D-19-0188.1.

d. Further model adaptations

To build the section shown on Fig. 4(a), we construct a series of subsections consisting of approximately evenly spaced points joining the lat/lon coordinates of the OSNAP mooring arrays, plus additional subsections where needed to make the section circumpolar. The dataset T and S fields are then bi-linearly interpolated onto these points in the lateral plane on each depth level, giving a 2D section of T and S in the along-section/depth plane, coinciding with the OSNAP array. We then calculate relative velocities through the section using the thermal wind relation:

$$v_{geos}(x,z) = \frac{-g}{\rho f}\int_{z}^{0}\frac{\partial \rho}{\partial x}dz, \tag{2}$$

where g is the acceleration due to gravity $= 9.81\,ms^{-2}$, f is the Coriolis parameter, ρ is the density of water at that point along the section calculated from the T and S fields, and z and x are the depth and distance along the section in metres. The relative velocities v_{geos} are then added to surface reference level velocities v_{ref} from the RTHIM solution to construct the full section velocities. The gridded EN4 and ECCO datasets have resolutions of 1 degree and 0.5 degrees, respectively. In order that the section definitions from the two datasets match as well as possible, we bi-linearly interpolate the EN4 fields onto a 1/4 degree grid before they are input to RTHIM. We used a 1/4 degree interpolation in order that the EN4 grid points can line up with the ECCO grid points which fall every X.25 and X.75 of a degree, while at the same time avoiding the loss of any information from EN4. We then apply a land mask constructed using 1/12 degree ETOPO5 bathymetry data [2] to both datasets.

An initial condition for the surface reference level velocity v_{ref} is calculated from geostrophy

[2]Data Announcement 88-MGG-02, Digital relief of the Surface of the Earth. NOAA, National Geophysical Data Center, Boulder, Colorado, 1988

Accepted for publication in *Journal of Physical Oceanography*. DOI 10.1175/JPO-D-19-0188.1.

using the annual-mean ssh η:

$$v_{ref} = \frac{g}{f}\frac{\partial \eta}{\partial x}. \tag{3}$$

To reduce near-gridscale noise in the initial condition (originating mainly from the observation error of the sea surface height, the observation and representation error of the geoid, and errors in the optimal interpolation in the gridded product), we explore smoothing v_{ref} using a boxcar filter of different widths. In the case of the AVISO η fields, we find that some smoothing is required to produce a realistic v_{ref}, which we apply using a moving average. We explore a range of parameters for the smoothing in the RTHIM ensembles (see section 3b). The ECCO ssh anomalies are quite smooth, and so do not require this step.

In order to more closely match the transport calculation method in L2019 for our comparison to the OSNAP observations, we introduce an additional constraint to RTHIM. It is known from long term observations that the net transport through Davis Strait has a long-term mean value of -1.6 Sv (Curry et al. 2014), so we add this constraint, with a weighting factor, to the net transport through the OSNAP-West part of the circumpolar section. This is consistent with the zero total net transport constraint for the whole section that was used in M2018 and L2019, and we explore the sensitivity of the RTHIM solution to the weights on both constraints in section 3. We do not impose an additional constraint on the transport through Bering Strait because this would require a third weighting factor and would therefore increase the size of the parameter space to explore for sensitivity of the RTHIM solutions.

Accepted for publication in *Journal of Physical Oceanography*. DOI 10.1175/JPO-D-19-0188.1.

3. Section transports

a. Qualitative comparison with observations

In this section we compare the section transports obtained from RTHIM solutions with those derived from OSNAP observations. The OSNAP velocity fields were produced using a combination of mooring data, Argo float profiles, glider data, CTD sections and the World Ocean Atlas climatology (for full details the reader is referred to L2019). First we have made a qualitative comparison on Fig. 5 between 2015 annual mean section velocities from RTHIM solutions and those derived from OSNAP observations. Key features such as the Labrador Current, East and West Greenland Currents, Irminger Current, a branch of the North Atlantic Current, and the southward flowing East Reykjanes Ridge Current, are common between the RTHIM solutions and the observations. There is a southward flowing current just to the west of the West Greenland Current in the EN4 solution but not the ECCO solution. This southward flow can be seen in the observed velocity field for the summer of 2016 shown on Fig. 5 of Holliday et al. (2018), but not in the 2014 field from the same paper or in the bottom panel of our Fig. 5. Neither inverse model solution captures all the features seen in the observations, and in both the currents are weaker and broader than those observed. However of the two solutions, EN4 with its altimetry-based surface v_{ref} has sharper gradients, more like the observations.

Overturning streamfunctions calculated from ensembles of RTHIM solutions for 2015 using the EN4 and ECCO datasets are compared with those calculated using the OSNAP observations for the same period on Fig. 6. The ensembles contain RTHIM solutions for a range of model parameters (see section 3b and Table 1). The OSNAP section velocity, temperature and salinity fields were constructed as described in L2019. The monthly mean fields have been further averaged to obtain a 2015 mean, and the section densities calculated from the mean T and S using TEOS-10

Accepted for publication in *Journal of Physical Oceanography*. DOI 10.1175/JPO-D-19-0188.1.

(McDougall and P. Barker 2011). We use the time-mean T and S here so that the OSNAP streamfunctions are comparable to those from RTHIM, for which densities calculated from 2015-mean section T and S from each dataset have been combined with the 2015 solution section velocity. Overturning streamfunctions are calculated according to:

$$\psi_\sigma(\sigma^*) = \iint_{\sigma<\sigma^*} v\,dA, \tag{4}$$

where $v\,dA$ is the transport through the section (or part section), and is integrated under contours of constant density σ^*. For the full OSNAP section the EN4 'best fit' solution (see section 3b) gives the better fit of the two to the observations, and in the case of the full OSNAP section the observations fall within the envelope of the ensemble for most of the density range. The maximum of the EN4 streamfunction just above 27.7 $kg\,m^{-3}$ is around the right density, although too large in magnitude. The ECCO streamfunction maximum is both too large and too deep, and the observed streamfunction is outside the envelope of the ECCO ensemble here.

Both RTHIM solutions correctly show the majority of the overturning occuring across OSNAP East, as reported by L2019, but neither matches the observed structure across OSNAP West, with the observations going outside the ensemble envelopes for much of the density range. The largest discrepancy is the strong peak in the ECCO streamfunction for OSNAP West around 27.75 $kg\,m^{-3}$, which can be explained by looking at the slope of the isopycnals across OSNAP West on the bottom panel of Fig. 5. The 27.7 $kg\,m^{-3}$ and 27.75 $kg\,m^{-3}$ contours are close together on the western side of the basin, where the Labrador Current flows southwards, and farther apart on the eastern side where the West Greenland Current flows northwards. The resulting large net northward transport in this density range gives the peak in the ECCO OSNAP West streamfunction on Fig. 6. By contrast, the same isopycnals are symmetrical across the basin in EN4 (top panel of Fig. 5) so the section transports due to the boundary current on either side of the basin in this density range

Accepted for publication in *Journal of Physical Oceanography*. DOI 10.1175/JPO-D-19-0188.1.

largely cancel out. The density structure in EN4 is much more similar to the observations here. The observed streamfunction across OSNAP West between 27.4 $kg\,m^{-3}$ and 27.9 $kg\,m^{-3}$ is characterised by a northward flow in the lighter waters, a southward flow in the intermediate waters, and northward flow in the densest waters, resulting in a weak net overturning as measured by the maximum of the streamfunction. It may be that the overestimate in the peak from ECCO indicates an overproduction of dense water in the Labrador Sea: Li et al. (2019) suggested that such an overproduction causes a bias in the MOC in some models. We must therefore bear in mind this discrepancy when interpreting our inverse solutions based on the ECCO dataset.

The MOC derived from the full ECCO velocity fields (as opposed to the geostrophic velocity estimate we have used with RTHIM) can be seen in appendix figure C1. It is similar to the RTHIM solution using the ECCO dataset over much of the density range, but with a less pronounced spike at the density of the maximum of the streamfunction, and the maximum also appears at a lighter density. The same figure shows the adjustment that has occurred between an initial condition calculated from geostrophic surface velocities (Eq. 3) and thermal wind shear (Eq. 2) and the solution for the two datasets; this is most significant in the EN4 case where there is significant net transport through the section in the initial condition.

b. Quantitative comparison with observations

We now make a quantitative comparison between section transports derived from the OSNAP array observations and our RTHIM solutions. To do so we define nine metrics: the MOC, meridional heat transport (MHT) and meridional freshwater transport (MFT) for the whole OSNAP section, for OSNAP West, and for OSNAP East. The MOC is the maximum of the streamfunction in Eq 4;

Accepted for publication in *Journal of Physical Oceanography*. DOI 10.1175/JPO-D-19-0188.1.

the MHT and MFT are as follows:

$$MHT = \rho_0 c_p^0 \int_{\Theta_{min}}^{\Theta_{max}} \psi_\Theta \, d\Theta, \quad MFT = \frac{-1}{S_{A0}} \int_{S_{Amin}}^{S_{Amax}} \psi_{S_A} \, dS_A, \tag{5}$$

where $\rho_0 c_p^0 = 4.1 \times 10^6$ Jm^{-3}K^{-1}, $S_{A0} = 35\ gkg^{-1}$ is a reference salinity, and ψ_Θ and ψ_{S_A} are calculated analagously to ψ_σ in Eq. 4, but integrated under contours of constant Θ and S_A, respectively.

In order to determine the uncertainties on our metrics for the RTHIM solutions, we have carried out ensembles of RTHIM runs on both datasets where we explore the sensitivity of the solutions to a range of model parameters, summarised in Table 1. To test the sensitivity for each parameter, we vary one at a time while fixing the other parameters, as was done in M2018. This method is a compromise between the desire to examine fully the uncertainty on our solutions and the available computation time, since there are too many permutations for an exploration of the full parameter space to be feasible. For each ensemble, the result presented is an individual ensemble member chosen to most closely resemble the ensemble mean in terms of our metrics. We do this rather than presenting the ensemble mean of the metrics because individual solutions obey the balance based on volume, heat and salt conservation from equation 1, and as such are more physically realistic than the ensemble mean. The ensemble member most closely resembling the ensemble mean is established using a function C:

$$C = \sum_{i=1}^{9} \left(\frac{TP_i - \overline{TP_i}}{\delta_{TP_i}} \right)^2. \tag{6}$$

The sum is over our nine transport metrics TP_i; TP and $\overline{TP}$ are the individual solution and ensemble mean transport metrics; and δ_{TP} is the ensemble standard deviation. The ensemble member taken as having the ‘best fit’ to the ensemble mean is that which gives the smallest value of C in Eq. 6. The best fit solutions from the 2015 EN4 and ECCO RTHIM ensembles are shown by the coloured bars on Fig. 7, along with their observational equivalents. The RTHIM solutions’ error bounds

Accepted for publication in *Journal of Physical Oceanography*. DOI 10.1175/JPO-D-19-0188.1.

show the ensemble range for each metric; note that the coloured bars do not fall in the centre of the error bounds because they correspond to an individual solution rather than the ensemble mean. The observational error bounds are calculated according to $\Delta MOC = \delta_{MOC}/\sqrt{365/(2\times 16\, days)}$ (and similarly for ΔMHT and ΔMFT), where δ_{MOC} is the standard deviation of the MOCs calculated from the 12 individual months of observations, and 16 days is the integral timescale calculated from the autocorrelation function of the daily MOC time series (see section 1e of L2019 supplementary material). This is reported in L2019 as being a close estimate to the observational error they obtain from a more sophisticated Monte Carlo technique.

The agreement between the transport metrics from the RTHIM ensembles and the observations is generally good, with the MOC across OSNAP West from the ECCO runs the notable exception (Fig. 7). The MHT and MFT from both ensembles mostly agree with the observations within their uncertainties, with the exception being the OSNAP West MFT where both slightly underestimate the (southward) freshwater transport. The reasonable agreement between the RTHIM solution overall section transports and the observations gives us some confidence in the inferred formation rates presented in the next section. We can also diagnose the same metrics for the other two years of RTHIM ensembles as we have done for 2015; these are summarised in Table 2.

4. Formation rates

We now examine the volume budgets from our RTHIM solutions; this is the part of the solution from which we calculate the transformations of individual water masses, and the contributions to these transformations from different physical processes. The volume budget for an RTHIM solution from the EN4 2015 dataset is shown on Fig. 8. The solution is the ensemble member that best fits the ensemble mean, identified as described in section 3b. In the surface flux term, we see the

Accepted for publication in *Journal of Physical Oceanography*. DOI 10.1175/JPO-D-19-0188.1.

formation of the Arctic Water by cooling around $0°C$ and the destruction of the warm, salty North Atlantic Water. There is also a formation signal in the warmer waters between $30-34\,g\,kg^{-1}$ which falls in a region of TS space with little to no volume on the time average (Fig. 2). Examining the monthly surface TS distribution (see animation in supplemental material) reveals that the isotherms and isohalines bounding this region of TS space outcrop in the summer months, which is responsible for this signal in the surface flux term. These waters are then mixed and transformed into other waters at the same average rate as they are formed, as seen in the mixing term. The mixing term is in competition with the surface flux term, having the opposite sign over much of TS space, with the strongest formation around $0°C$ and around $3°C$, $34.5\,g\,kg^{-1}$. In the advection term we can see the northward flow of warm salty waters and southward flow of cooler, fresher waters. The volume trend is generally smaller than the other terms, with some net formation in the main water masses, and destruction of the saltiest Arctic Water.

Fig. 9 shows the volume budget for the EN4 2015 solution, zoomed in to show the LSW and OW water masses. There is a formation signal in the surface flux just cooler and fresher than the box defining the LSW, and at the same density. There is also a northward flow in the advection term just warmer and saltier than the LSW, which presumably comes from the Irminger Current as it enters the Labrador Sea through OSNAP-W (Fig. 1). Meanwhile, the mixing term shows destruction of these adjacent water masses while showing a net formation within the LSW box itself. This suggests that the LSW has been transformed through mixing, predominantly along isopycnals, of the cold, fresh water mass formed by deep convection and the warm, salty water mass advected in. The overflow waters also are formed predominantly by mixing and then advected southwards, with the bulk of the formation in the colder OW box. The water masses mixing to form the OW are likely in the two adjacent blue areas (blue representing net destruction): the AAW slightly saltier and around the same temperature, and the unlabelled water mass fresher and around $2°C$.

Accepted for publication in *Journal of Physical Oceanography*. DOI 10.1175/JPO-D-19-0188.1.

We now compare the formation rates from RTHIM solutions applied to our six datasets: EN4 2013/14/15 and ECCO 2013/14/15 (Fig. 10). We have integrated the formation rates from the RTHIM solutions over the water mass definitions from Fig. 2, grouped as follows: North Atlantic Water (NAW), Labrador Sea Water (LSW), Overflow Water (OW), and Arctic Water (AW = ASW + ADW + AAW). For each of the six ensembles the coloured bars show the solution that most closely resembles the ensemble mean as established using Eq. 6.

NAW is advected into the region and then transformed by a combination of mixing and surface flux, with mixing the larger term. The advection is consistent between the EN4 and ECCO solutions and constant over the three years. The picture is more complex for the mixing and surface flux terms, as mixing in EN4 reduces between 2013 (weak convection) and 2014-15 (strong convection), whereas ECCO shows mixing increasing and then decreasing while the surface flux increases steadily. The difference in this pattern is explained by the differences in the trend term, which is diagnosed directly from the T and S fields and must also be the sum of the other three terms. In the EN4 solution for 2013 the trend term is negative, meaning that the net effect of the inflow and transformation by mixing and surface flux is to reduce the volume of NAW within the region. In ECCO for 2013 the trend is positive, i.e. the volume of NAW is increasing. In 2014 the opposite is true: the ECCO solution shows decreasing NAW volume while in EN4 there is a small increase. In order to close the volume budget in each case, RTHIM has attributed differing amounts of transformation to mixing in each solution for each year.

In the LSW there is larger variation between the solutions from the different datasets. Part of this may be attributable to the large volumes of water in the small region of TS space where the LSW resides, which means the solutions are sensitive to differences in the volumetric distributions of ECCO and EN4 (see Figs. 2 and A1). However for simplicity we have kept the definition of the LSW the same across all the ensembles. In 2013, LSW is formed by a combination of mixing

Accepted for publication in *Journal of Physical Oceanography*. DOI 10.1175/JPO-D-19-0188.1.

and surface fluxes in both ensembles, with mixing dominating, although the ECCO solution has a larger role for the surface flux. These waters are then exported (negative advection term), and the ECCO solution shows a decrease in volume (negative trend term). As the years progress both sets of solutions show a steady increase in the rate of production of LSW by mixing, but the export rates follow an upward trend in the EN4 solutions versus a downward trend in ECCO. The surface flux and trend terms are very small for the EN4 solutions, but play a significant role in the balance for ECCO; in the latter the increased production of LSW in 2014-15 results in storage (positive trend term), rather than export. Both the EN4 and ECCO solutions show a net destruction of the LSW by the surface flux in 2015, our strongest deep convection year. This perhaps surprising result can be understood by referring back to Fig. 9: the surface fluxes are cooling our LSW and some fresher waters adjacent to it while forming a cooler and fresher water mass, which meanwhile gets mixed back to form the LSW.

In the OW, both sets of solutions show similar rates of export in all three years, although the rates are higher in ECCO. In 5 out of the 6 ensembles mixing forms the OW, the exception being the ECCO solution in 2015 where the surface flux has a major role. A closer inspection of the volume budget for the ECCO 2015 RTHIM solution (not shown) reveals that the large formation signal is at the warm boundary of the cooler OW box, and likely represents formation due to cooling by convection in the Labrador Sea. This signal appears in the EN4 solution as discussed above, but in fresher waters that are outside our OW definition.

In the AW we see an even balance between formation by surface flux and distruction by mixing, with almost no signal in the advection since these water masses do not generally flow through the section. The EN4 and ECCO solutions disagree on the magnitudes of the rates of formation /destruction of AW, and neither show an obvious trend over the 3 years. The effect of random errors on the surface fluxes has been accounted for in our EN4 ensembles, but if there is a bias in the

Accepted for publication in *Journal of Physical Oceanography*. DOI 10.1175/JPO-D-19-0188.1.

surface flux for either dataset then this would explain the disagreement. For example, if the magnitude of the (negative) heat flux from ERA-Interim were increased this would increase the rate of production of AW in the surface flux term, and RTHIM would increase the rate of destruction by mixing to maintain the volumetric balance, reducing the disagreement with the ECCO solution.

5. Discussion

a. Datset comparison

We have presented RTHIM solutions based on two distinct datasets: one composed of outputs from the ECCO state estimate and one using a combination of EN4 and ERA-Interim reanalyses and AVISO satellite altimetry. Both datasets produced solutions for the advection through the OSNAP section which agree reasonably well with observations from the OSNAP array, but the EN4 dataset was the closer of the two. On examining the formation rates, the two datasets generally agree on the sign, but not the magnitude, of contributions to the volume budget from advection, mixing, surface fluxes and the volume trend. Where there is disagreement between the solutions derived from each dataset, it is difficult to know which is more realistic, since EN4 and ECCO include different types of observational influence. Where observations are always missing, EN4's persistence-based forecast form of objective analysis adjusts the solution towards climatology Good et al. (2013). However, ECCO's 4DVar state estimate is effectively a free-running forward model solution with forcing/parameters chosen so that its state most closely fits the observations. This means that information from near-surface observations can, in principle, continue to propagate into the deep ocean in the case of ECCO, provided there are dynamical connections in the forward model (Forget et al. 2015). We also look at how the formation rates differ between 2013-2015, three years with varying rates of convection in the Labrador Sea. Both datasets show

Accepted for publication in *Journal of Physical Oceanography*. DOI 10.1175/JPO-D-19-0188.1.

an increase in the production rate of LSW by mixing as convection increased between 2013 and 2015, but EN4 and ECCO show opposite trends in export rates, while the ECCO solutions also show a larger role for surface fluxes and the volume trend in the balance.

In the high Arctic, the interior T and S of our control volume are poorly constrained by observations, leading to issues with EN4 described in section 2b. This affects two aspects of RTHIM: the initial condition for the mixing term which is based on TS gradients, and the fixed trend term which is calculated from changes in the volumetric distribution of T and S. We have explored the effects of variations in the mixing term initial condition through our ensembles, and their ranges are plotted in the error bars on Figs. 7 and 10. For the EN4 ensembles these ranges include runs where we have added random errors to the T and S fields. We also briefly explore the uncertainty on the trend term in both datasets by including an ensemble member where it is set to zero (Table 1). This can be seen on Fig. 10 in the fact that the trend term uncertainties all have zero as either an upper or lower bound.

There are differences in the magnitudes of the formation rates due to surface flux for the two datasets, and with only two to compare it is difficult to be confident in a preference for one over the other. In most cases the differences are not large enough to change the nature of the volumetric balance, but there are a few exceptions in the Labrador Sea Water and Overflow Waters. It is the large volumes associated with these water masses (combined they occupy 12.5% and 12.2% of the total control volume for the EN4 and ECCO datasets, respectively) and their close proximity in TS space which make them difficult to distinguish in this coordinate system.

b. Labrador Sea Water

The period of 2013-2015 addressed in this study coincides with the development of a class of LSW defined in YL2017 as $LSW_{2012-2016}$. On their Fig. 2 this water mass has potential tem-

Accepted for publication in *Journal of Physical Oceanography*. DOI 10.1175/JPO-D-19-0188.1.

perature around $3.2-3.7°$C and practical salinity around $34.80-34.90$ PSU. In our coordinates of (Θ, S_A) the temperature is equivalent to our degree of precision, and the salinity converts to $34.97-35.07\ g\,kg^{-1}$. Our LSW definition based on the ECCO and EN4 2015 volumetric TS distributions of $\Theta = 3.3-4°C$ and $S_A = 35-35.1\ g\,kg^{-1}$ is slightly warmer and slightly saltier than this, but fits with the traditional LSW density range of $\sigma_0 = 27.68-27.80\ kg\,m^3$ quoted by Li et al. (2019) as shown on our Fig. 9. Perhaps the upper end of our LSW salinity range is a little too high and includes some of the OW for the EN4 dataset; the position of the boundary was a compromise reached to try to keep a consistent definition for both EN4 and ECCO. We may also have excluded some LSW with our lower temperature bound of 3.3°C: most of the LSW in the YL2017 figure is in the temperature range 3.2-3.4°C. It is possible that the warmer waters we have identified from the volumetric TS distributions of our Figs. 2 and 6 are legitimate LSW, either formed in the Irminger Sea or formed in the Labrador Sea and recirculated. Equally it is possible that our boundary of 4°C between the LSW and the NAW is a little too high. It is also worth noting that the sections from YL2017 on which these comparisons are based are snapshots from a survey done in May 2016, whereas our definitions are constructed using a time-mean of the 2015 volumetric TS distribution.

We can interpret our results in the context of forcing over the seasonal cycle. The formation rates diagnosed due to each process are annual means, but the formation itself is likely to have taken place over shorter time periods and at different times of the year. For example, the formation of the water mass just cooler and fresher than the LSW on Fig. 9 will have occurred during convection in the winter months; meanwhile the mixing forming the LSW itself probably began during convection but continued for some time after. The isopycnal mixing in the Labrador Sea of newly convected cold, fresh water with warm, salty water brought in by advection is consistent with the description by Y2007a of the evolution of LSW which we introducted in section 1b. We also see

Accepted for publication in *Journal of Physical Oceanography*. DOI 10.1175/JPO-D-19-0188.1.

little export of the newly convected water in the advection term of Fig. 9; instead the waters are mixed into the LSW box before being exported. These findings fit with those of Pickart and Spall (2007) who suggest that LSW is generated at the boundaries of the Labrador Sea via adiabatic eddies, and of Georgiou et al. (2019) who show that dense water formed in the interior of the Labrador Sea is laterally advected into the boundary current by eddies before it can be exported.

The steady warming, salinification and recirculation of LSW over the years following its initial formation described by Y2007a may also explain the fact that much of the water identified in our volumetric TS distributions as LSW is somewhat warmer and saltier than that seen in the sections of YL2016 and YL2017. It is likely that much of what we see in the volumetric census of 2015 is recirculated recently formed LSW that has had more time to mix with other waters. The suggestion of Yashayaev et al. (2007) that LSW spreads to the Irminger Sea 1-2 years after its formation is consistent with this idea.

Our RTHIM solutions using the EN4 dataset give LSW formation rates from a combination of mixing and surface fluxes of 6.2 Sv, 8.3 Sv, and 11.3 Sv for 2013, 2014, and 2015, respectively, with mixing dominating. Using the ECCO dataset the formation rates are 5.6 Sv, 11.9 Sv and 6.0 Sv, with the surface fluxes making a significant contribution to the formation; in particular the effect of the surface flux is to contribute to the formation of LSW in 2014 but to counteract it in 2015 by cooling waters that are already in the LSW class. These formation rates are in the same ballpark as the historical estimates discussed in section 1b, and are also of the same order as the export rates of 3.2 Sv and 8.9 Sv in weak convection and strong convection years, respectively, reported by YL2016. Our results may in fact be closer to the real formation rates as we have explicitly accounted for the contributions of surface fluxes and mixing to formation. As discussed above, the upper bound on our temperature range for defining the LSW may be a little high. If it is reduced to 3.7°C, RTHIM solutions using the EN4 dataset give slightly lower formation rates of

Accepted for publication in *Journal of Physical Oceanography*. DOI 10.1175/JPO-D-19-0188.1.

4.2 Sv, 3.6 Sv and 9.0 Sv for 2013, 2014, and 2015, respectively; the latter value in close agreement with YL2016 for strong convection years.

The question of the link between formation and export rates (i.e. respectively $-\nabla_{S\theta} \cdot \boldsymbol{U}_{S\theta}^{surf} - \nabla_{S\theta}^2 \boldsymbol{F}$ and $-I_{adv}$ from Eq. 1) remains unresolved due to the differences between our solutions from EN4 and ECCO. However, the ECCO solution offers an illustration of the disconnect because while the formation rates were higher in the strong convection years, the export rates as seen in the advection term were lower. The difference is taken up in the trend term: the volume of LSW increased in 2014 following convection (positive trend term on Fig. 10), with a further increase in 2015 partially counteracted by convection creating more source waters for the formation of new LSW by mixing (negative surface flux term, large positive mixing term). YL2016 reported that during the Argo era the average winter LSW volume was about 70% larger in strong convection years than in weak ones; however the reduction in volume from winter to autumn was 180% larger, giving a factor of 2.8 difference in potential LSW export rates in strong convection years. It will be interesting to see what is revealed by the next set of OSNAP observations in the context of the recently observed deep convection. The previously described observation that dense water formation rate is less variable in the Nordic Seas than in the Labrador and Irminger Seas is supported by our results: the AW and OW formation rates have ranges of ~ 2.5 Sv and ~ 3.5 Sv respectively, while the LSW formation rate has a range of ~ 4.5 Sv.

It is also difficult to draw firm conclusions from this study about the role of the LSW in the MOC variability. On the one hand, the MOC across the whole OSNAP section was larger in 2014-2015 when we had higher rates of LSW formation; however the relationship between the whole section MOC and that diagnosed across its western and eastern parts is unclear (see Table 2). If we consider only the EN4 solutions on the grounds of the large discrepancy between the observed MOC across OSNAP West and that derived from the ECCO solutions, the full-section MOC increases

Accepted for publication in *Journal of Physical Oceanography*. DOI 10.1175/JPO-D-19-0188.1.

steadily while the contributions from both parts of the sections fluctuate. It may be neccessary to analyze more years of data in order to establish a possible link between LSW formation and the MOC.

6. Conclusion

We have applied a Regional Thermohaline Inverse Method (RTHIM) to diagnose the water mass transformation in an enclosed region of the Arctic and Subpolar North Atlantic Ocean. Six sets of results were obtained by applying RTHIM to three separate years of data from two different datasets; the year 2013 where the convection in the Labrador Sea was relatively weak and 2014-2015 where it was strong. For each solution we obtain transports through the section bounding the volume, and formation rates due to water mass transformation within the volume. Comparisons between inverse solution section transports and those derived using independent observations from the OSNAP mooring array were good, giving confidence in the formation rates. The latter were summarised in terms of the contributions of different processes to the formation of important water masses, with a particular focus on the Labrador Sea Water (LSW), and the results from the three years and two datasets compared.

Annual mean formation rates for LSW ranged from a low of $\sim$ 6 Sv in 2013 when convection was weak to highs of $\sim$ 12 Sv in either 2014 or 2015 (depending on which dataset was used) when convection was stong, with interior mixing playing a leading role in the formation. The effect of winter convection was to create a water mass slightly colder and fresher than the resident LSW class, but this water was not exported directly. Instead the newly convected water was mixed isopycnally with warm, salty waters carried in by advection. The product, the intermediate temperature and salinity LSW, was then exported.

This was the first application of the recently validated Regional Thermohaline Inverse Method

Accepted for publication in *Journal of Physical Oceanography*. DOI 10.1175/JPO-D-19-0188.1.

to observation-based data. Its success in diagnosing a MOC for the time period coinciding with OSNAP that is consistent with observations indicates its potential for further analysis of the circulation in the region. By applying RTHIM to other years of the observational data products used in this study, or to similar products, yet further context could be provided for the OSNAP observations. It may be possible to look at interannual variability in the meridional overturning circulation as measured across the OSNAP section and explain this in terms of changes in the water mass transformation in the region.

Acknowledgments. NM, CW and NPH acknowledge funding from the U.K. Natural Environment Research Council under the U.K. OSNAP Large Grant (NE/K010875.1). NPH and CW are additionally supported by the U.K. Natural Environment Research Council's North Atlantic Climate System Integrated Study program (NE/N018044/1) and Climate Linked Atlantic Sector Science (CLASS, NE/R015953/1). NPH is also supported by NERC UK-OSNAP-Decade (NE/T00858X/1 and NE/T00858X/2). JDZ acknowledges funding from Australian Research Council grant DP1603130. NM is additionally supported by NERC grant NE/P021298/1, with thanks to Professor Andrew Watson. The OSNAP data are available for download from o-snap.org.

APPENDIX A

Volumetric TS distributions

Fig. A1 shows the full volumetric TS distribution for the ECCO 2015 dataset, and Fig. A2 shows the volumetric distributions in the region of TS space occupied by the LSW for the EN4 and ECCO datasets for 2013, 2014 and 2015. Table A1 shows the water mass definitions which are plotted on Figs. 2 and A1.

APPENDIX B

Accepted for publication in *Journal of Physical Oceanography*. DOI 10.1175/JPO-D-19-0188.1.

TS grid details

The TS grids used in this study are each defined using two row vectors: one in the T dimension and one in the S dimension. These row vectors are detailed in Table B1. The grid points in each vector define the midpoints of TS bins for which the terms in Eq. 1 are calculated. The grids are designed such that TS bin boundaries, which are at the midpoints between row vector points, correspond to defined water mass boundaries. The vectors cover the full range of T and S within the control volume for their respective datasets, plus an additional 'halo' of grid points which is required due to the inverse model discretisation. The water mass boundaries plotted on Figs. 2 and A1 used the 15 x 15 EN4 2015 and ECCO 2015 grids, respectively.

APPENDIX C

RTHIM transport adjustment

Fig. C1 shows overturning streamfunctions for RTHIM solutions from ECCO and EN4 compared to their initial conditions.

APPENDIX D

Details of inverse method

Here we describe some additional details of Eq. 1 and how it is solved to obtain our RTHIM solutions (for a full explanation including the derivation of Eq. 1 the reader is referred to M2018). The first term, I_{adv}, describes the transport through the section surrounding our control volume (the section shown as a red line on Fig. 4a). In each (S,θ) bin of our RTHIM TS grid, $I_{adv}(S,\theta)$ is the transport per unit S and θ perpendicular to the section (positive northwards, into the control volume) between the pairs of isohalines and isotherms defining that bin. The transport is the

Accepted for publication in *Journal of Physical Oceanography*. DOI 10.1175/JPO-D-19-0188.1.

section velocity $v = v_{ref} + v_{geos}$ (see section 2d) integrated over the area between the isohalines and isotherms. I_{adv}, as the other terms in Eq. 1, has units of $m^3\,s^{-1}\,g^{-1}\,kg\,K^{-1}$.

The mixing term $\nabla^2_{S\theta}\boldsymbol{F}$ is the formation rate per unit S and θ within the control volume due to mixing. It is calculated by applying the operator $\nabla^2_{S\theta}$ to the diffusive flux tensor $\boldsymbol{F}$, where $\nabla_{S\theta}\cdot[\,] = (\frac{\partial}{\partial S}, \frac{\partial}{\partial \theta})\cdot[\,]$ is the divergence operator in thermohaline coordinates. The diffusive flux tensor has four components (so that when applying the divergence operator twice we obtain a scalar): $F_{\theta\theta}$, $F_{S\theta}$, $F_{\theta S}$, and F_{SS}, where $F_{C_1C_2}$ represents the diffusive flux of tracer C_1 across and in the direction perpendicular to the iso-surface of tracer C_2. The RTHIM solution is constrained such that $F_{SS}, F_{\theta\theta} \geq 0$, and $F_{S\theta} = F_{\theta S}$.

The volume trend term $\frac{\partial V}{\partial t}(S,\theta)$ is the rate of change of the volume of water in each (S,θ) bin, i.e. that contained within pairs of isohalines and isotherms defining each bin. This is divided by the width of the bins in TS space, $\Delta S \Delta \theta$, so that it has the same units of $m^3\,s^{-1}\,g^{-1}\,kg\,K^{-1}$. It is calculated by taking a volumetric census of the water masses at the start and end of an averaging period (e.g. the year 2015) using the T and S data from our datasets (EN4 or ECCO).

The surface flux term $\nabla_{S\theta}\cdot\boldsymbol{U}^{surf}_{S\theta}(S,\theta)$ is the divergence in thermohaline coordinates of the vector $U^{surf}_{S\theta}(S,\theta)$, which has components of $(U^{surf}_{S}(S,\theta), U^{surf}_{\theta}(S,\theta))$. The components are calculated by integrating the surface fluxes of heat and freshwater from our datasets between the isohalines and isotherms defining each (S,θ) bin where they outcrop.

In the RTHIM inverse calculation, the volume trend and surface flux terms are prescribed, and the advection and mixing terms are solved for. In the case of the mixing term, we prescribe the relative velocities v_{geos} and solve for the surface reference velocity v_{ref}, using Eq. 3 as the initial condition. In the case of the mixing term, we calculate an initial condition for the diffusive flux tensor $\boldsymbol{F}$ from gradients in our dataset temperature and salinity fields, and solve for the components of $\boldsymbol{F}$ given the constraints outlined above. During the optimisation used to obtain our RTHIM

Accepted for publication in *Journal of Physical Oceanography*. DOI 10.1175/JPO-D-19-0188.1.

solution, the sum over all TS space of $(\varepsilon\Delta\theta\Delta S)^2$ is minimised, and this sum has final values of $< 10^{-4}\,Sv^2(1Sv = 10^6\,m^3\,s^{-1})$.

References

Bower, A., S. Lozier, and S. Gary, 2011: Export of Labrador Sea Water from the subpolar North Atlantic: A Lagrangian perspective. *Deep Sea Research Part II: Topical Studies in Oceanography*, **58 (17-18)**, 1798–1818, doi:10.1016/j.dsr2.2010.10.060, URL https://linkinghub.elsevier.com/retrieve/pii/S0967064511000233.

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/articles/nature07979.

Carton, J. A., S. G. Penny, and E. Kalnay, 2019: Temperature and Salinity Variability in the SODA3, ECCO4r3, and ORAS5 Ocean Reanalyses, 1993–2015. *Journal of Climate*, **32 (8)**, 2277–2293, doi:10.1175/JCLI-D-18-0605.1, URL http://journals.ametsoc.org/doi/10.1175/JCLI-D-18-0605.1.

Curry, B., C. M. Lee, B. Petrie, R. E. Moritz, and R. Kwok, 2014: Multiyear Volume, Liquid Freshwater, and Sea Ice Transports through Davis Strait, 2004–10. *Journal of Physical Oceanography*, **44 (4)**, 1244–1266, doi:10.1175/JPO-D-13-0177.1, URL http://journals.ametsoc.org/doi/10.1175/JPO-D-13-0177.1.

Danabasoglu, G., S. G. Yeager, Y. O. Kwon, J. J. Tribbia, A. S. Phillips, and J. W. Hurrell, 2012: Variability of the atlantic meridional overturning circulation in CCSM4. *Journal of Climate*, **25 (15)**, 5153–5172, doi:10.1175/JCLI-D-11-00463.1.

Accepted for publication in *Journal of Physical Oceanography*. DOI 10.1175/JPO-D-19-0188.1.

Daniault, N., and Coauthors, 2016: The northern North Atlantic Ocean mean circulation in the early 21st century. *Progress in Oceanography*, **146**, 142–158, doi:10.1016/j.pocean.2016.06.007, URL https://linkinghub.elsevier.com/retrieve/pii/S0079661116300659.

Dee, D. P., and Coauthors, 2011: The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, **137 (656)**, 553–597, doi:10.1002/qj.828, URL http://doi.wiley.com/10.1002/qj.828.

Forget, G., J.-M. Campin, P. Heimbach, C. N. Hill, R. M. Ponte, and C. Wunsch, 2015: ECCO version 4: an integrated framework for non-linear inverse modeling and global ocean state estimation. *Geoscientific Model Development*, **8 (10)**, 3071–3104, doi:10.5194/gmd-8-3071-2015, URL https://www.geosci-model-dev.net/8/3071/2015/.

Fröb, F., A. Olsen, K. Våge, G. W. K. Moore, I. Yashayaev, E. Jeansson, and B. Rajasakaren, 2016: Irminger Sea deep convection injects oxygen and anthropogenic carbon to the ocean interior. *Nature Communications*, **7 (1)**, 13 244, doi:10.1038/ncomms13244, URL http://www.nature.com/articles/ncomms13244.

Fukumori, I., O. Wang, I. Fenty, G. Forget, P. Heimbach, and R. M. Ponte, 2017: ECCO Version 4 Release 3. **2 (2015)**, 10, doi:1721.1/110380, URL https://dspace.mit.edu/handle/1721.1/110380.

Gary, S. F., M. Susan Lozier, C. W. Böning, and A. Biastoch, 2011: Deciphering the pathways for the deep limb of the Meridional Overturning Circulation. *Deep Sea Research Part II: Topical Studies in Oceanography*, **58 (17-18)**, 1781–1797, doi:10.1016/j.dsr2.2010.10.059, URL https://linkinghub.elsevier.com/retrieve/pii/S0967064511000221.

Georgiou, S., C. G. van der Boog, N. Brüggemann, S. L. Ypma, J. D. Pietrzak, and C. A. Katsman, 2019: On the interplay between downwelling, deep convection and mesoscale eddies in the

Accepted for publication in *Journal of Physical Oceanography*. DOI 10.1175/JPO-D-19-0188.1.

Labrador Sea. *Ocean Modelling*, **135 (February)**, 56–70, doi:10.1016/j.ocemod.2019.02.004, URL https://doi.org/10.1016/j.ocemod.2019.02.004.

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.

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.

Haine, T., and Coauthors, 2008: *North Atlantic Deep Water transformation in the Labrador Sea, recirculation through the subpolar gyre, and discharge to the subtropics*. doi:10.1007/978-1-4020-6774-7_28.

Holliday, N. P., S. Bacon, S. A. Cunningham, S. F. Gary, J. Karstensen, B. A. King, F. Li, and E. L. Mcdonagh, 2018: Subpolar North Atlantic Overturning and Gyre-Scale Circulation in the Summers of 2014 and 2016. *Journal of Geophysical Research: Oceans*, **123 (7)**, 4538–4559, doi:10.1029/2018JC013841.

Accepted for publication in *Journal of Physical Oceanography*. DOI 10.1175/JPO-D-19-0188.1.

LeBel, D. A., and Coauthors, 2008: The formation rate of North Atlantic Deep Water and Eighteen Degree Water calculated from CFC-11 inventories observed during WOCE. *Deep Sea Research Part I: Oceanographic Research Papers*, **55 (8)**, 891–910, doi:10.1016/j.dsr.2008.03.009, URL https://linkinghub.elsevier.com/retrieve/pii/S0967063708000472.

Li, F., and Coauthors, 2019: Local and downstream relationships between labrador sea water volume and North Atlantic meridional overturning circulation variability. *Journal of Climate*, **32 (13)**, 3883–3898, doi:10.1175/JCLI-D-18-0735.1.

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.

Lozier, M. S., and Coauthors, 2019: A sea change in our view of overturning in the subpolar North Atlantic. *Science*, **363 (6426)**, 516–521, doi:10.1126/science.aau6592, URL http://www.sciencemag.org/lookup/doi/10.1126/science.aau6592.

Mackay, N., C. Wilson, J. Zika, and N. P. Holliday, 2018: A Regional Thermohaline Inverse Method for Estimating Circulation and Mixing in the Arctic and Subpolar North Atlantic. *Journal of Atmospheric and Oceanic Technology*, **35 (12)**, 2383–2403, doi:10.1175/JTECH-D-17-0198.1, URL http://journals.ametsoc.org/doi/10.1175/JTECH-D-17-0198.1.

Accepted for publication in *Journal of Physical Oceanography*. DOI 10.1175/JPO-D-19-0188.1.

McDougall, T. J., and P. Barker, 2011: Getting started with TEOS-10 and the Gibbs Seawater (GSW) Oceanographic Toolbox. SCOR/IAPSO WG127, 28 pp.

Pickart, R. S., and M. A. Spall, 2007: Impact of Labrador Sea Convection on the North Atlantic Meridional Overturning Circulation. *Journal of Physical Oceanography*, **37 (9)**, 2207–2227, doi:10.1175/JPO3178.1, URL http://journals.ametsoc.org/doi/abs/10.1175/JPO3178.1.

Pickart, R. S., F. Straneo, and G. W. Moore, 2003: Is Labrador Sea Water formed in the Irminger basin? *Deep-Sea Research Part I: Oceanographic Research Papers*, **50 (1)**, 23–52, doi:10.1016/S0967-0637(02)00134-6.

Polyakov, I. V., and Coauthors, 2004: Variability of the Intermediate Atlantic Water of the Arctic Ocean over the Last 100 Years. *Journal of Climate*, **17 (23)**, 4485–4497, doi:10.1175/JCLI-3224.1, URL http://journals.ametsoc.org/doi/10.1175/JCLI-3224.1.

Schott, F. A., R. Zantopp, L. Stramma, M. Dengler, J. Fischer, and M. Wibaux, 2004: Circulation and Deep-Water Export at the Western Exit of the Subpolar North Atlantic. *Journal of Physical Oceanography*, **34 (4)**, 817–843, doi:10.1175/1520-0485(2004)034⟨0817:CADEAT⟩2.0.CO;2.

Shu, Q., Q. Wang, J. Su, X. Li, and F. Qiao, 2019: Assessment of the Atlantic water layer in the Arctic Ocean in CMIP5 climate models. *Climate Dynamics*, **53 (9-10)**, 5279–5291, doi:10.1007/s00382-019-04870-6, URL http://link.springer.com/10.1007/s00382-019-04870-6.

Smeed, D. A., and Coauthors, 2014: Observed decline of the Atlantic meridional overturning circulation 2004-2012. *Ocean Science*, **10 (1)**, 29–38, doi:10.5194/os-10-29-2014.

Talley, L. D., and M. S. McCartney, 1982: Distribution and Circulation of Labrador Sea Water. *Journal of Physical Oceanography*, **12 (11)**, 1189–1205, doi:10.1175/1520-0485(1982)

Accepted for publication in *Journal of Physical Oceanography*. DOI 10.1175/JPO-D-19-0188.1.

012⟨1189:DACOLS⟩2.0.CO;2, URL http://journals.ametsoc.org/doi/abs/10.1175/1520-0485%281982%29012%3C1189%3ADACOLS%3E2.0.CO%3B2.

Thornalley, D. J. R., and Coauthors, 2018: Anomalously weak Labrador Sea convection and Atlantic overturning during the past 150 years. *Nature*, **556 (7700)**, 227–230, doi:10.1038/s41586-018-0007-4, URL http://www.nature.com/articles/s41586-018-0007-4.

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.

Yashayaev, I., 2007: Hydrographic changes in the Labrador Sea, 1960–2005. *Progress in Oceanography*, **73 (3-4)**, 242–276, doi:10.1016/j.pocean.2007.04.015, URL https://linkinghub.elsevier.com/retrieve/pii/S0079661107000857.

Yashayaev, I., M. Bersch, and H. M. van Aken, 2007: Spreading of the Labrador Sea Water to the Irminger and Iceland basins. *Geophysical Research Letters*, **34 (10)**, 1–8, doi:10.1029/2006GL028999.

Yashayaev, I., and B. Dickson, 2008: Transformation and Fate of Overflows in the Northern North Atlantic. *Arctic–Subarctic Ocean Fluxes*, Springer Netherlands, Dordrecht, 505–526, doi:10.1007/978-1-4020-6774-7_22, URL http://link.springer.com/10.1007/978-1-4020-6774-7_22.

Yashayaev, I., and J. W. Loder, 2016: Recurrent replenishment of Labrador Sea Water and associated decadal-scale variability. *Journal of Geophysical Research: Oceans*, **121 (11)**, 8095–8114, doi:10.1002/2016JC012046, URL http://doi.wiley.com/10.1002/2016JC012046.

Yashayaev, I., and J. W. Loder, 2017: Further intensification of deep convection in the Labrador Sea in 2016. *Geophysical Research Letters*, **44 (3)**, 1429–1438, doi:10.1002/2016GL071668.

Accepted for publication in *Journal of Physical Oceanography*. DOI 10.1175/JPO-D-19-0188.1.

Zhang, R., 2010: Latitudinal dependence of Atlantic meridional overturning circulation (AMOC) variations. *Geophysical Research Letters*, **37 (16)**, doi:10.1029/2010GL044474, URL http://doi.wiley.com/10.1029/2010GL044474.

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.

Zou, S., and M. S. Lozier, 2016: Breaking the Linkage Between Labrador Sea Water Production and Its Advective Export to the Subtropical Gyre. *Journal of Physical Oceanography*, **46 (7)**, 2169–2182, doi:10.1175/jpo-d-15-0210.1.

Accepted for publication in *Journal of Physical Oceanography*. DOI 10.1175/JPO-D-19-0188.1.

LIST OF TABLES

Accepted for publication in *Journal of Physical Oceanography*. DOI 10.1175/JPO-D-19-0188.1.

TABLE 1. Summary of model parameters varied in RTHIM ensembles. Note that v_{ref} smoothing, surface flux error, and T and S error only apply to runs using the EN4 dataset. The ERA-Interim surface flux errors were calculated by first creating arrays of random numbers between +0.5 and -0.5 with the same dimensions as the surface flux arrays. These were then multiplied by a scalar $40\,Wm^2$ in the case of the heat flux and $8{\times}10^{-9}\,m\,s^{-1}$ in the case of the freshwater flux, creating arrays of random errors of $\pm \leq 20\,Wm^2$ and $\pm \leq 4{\times}10^{-9}\,m\,s^{-1}$ with zero mean which were then added to their respective surface fluxes. The EN4 T and S errors were calculated in a similar way, except that the multipliers for the random number arrays were arrays with the same dimensions as the T and S fields, comprising of the uncertainties on T and S included with the EN4 product. Each ensemble using the EN4 dataset contains one run where errors have been added to the surface fluxes and one where errors have been added to the T and S fields as described.

Parameter	Description	Values
v_{ref} smoothing	Boxcar smoothing over 2n+1 grid cells of v_{ref} initial condition	Integers 1-12 inclusive
$\boldsymbol{F}$ smoothing	2D boxcar smoothing of diffusive flux tensor initial condition	Integers 1-4 inclusive
$W1$	Weighting factor applied to total net transport constraint	1, 10, 50, 100
W2	Weighting factor on transport through OSNAP West constraint	0.0005, 0.001, 0.005, 0.01, 0.05, 0.1
D	Vertical diffusivity used in calculating $\boldsymbol{F}$ initial condition	$10^{-4}\,m^2\,s^{-1}$, $10^{-5}\,m^2\,s^{-1}$
K	Lateral diffusivity used in calculating $\boldsymbol{F}$ initial condition	$100\,m^2\,s^{-1}$, $200\,m^2\,s^{-1}$, $500\,m^2\,s^{-1}$
TS grid size	Number of nodes in the T and S ranges constructing the grid	15 x 15, 16 x 16, 17 x 17
Trend term	Inclusion of the trend term in the volume budget	Yes / No
Surface flux error	Random errors within bounds added to surface fluxes	Heat $\pm \leq 20\,Wm^2$, freshwater $\pm \leq 4{\times}10^{-9}\,m\,s^{-1}$
T and S error	Random errors within bounds added to interior T and S	Uncertainty limits from EN4 dataset

Accepted for publication in *Journal of Physical Oceanography*. DOI 10.1175/JPO-D-19-0188.1.

TABLE 2. MOC (top), heat (middle) and freshwater (bottom) transports for all 3 years of RTHIM solutions and the OSNAP observations, measured across the whole OSNAP section (-All), OSNAP West (-W), and OSNAP East (-E). For the RTHIM ensembles the value for each metric from the best fit solution to the ensemble mean is given followed by the ensemble range in brackets as (lower bound, upper bound); for the observations the 2015 mean and uncertainty range are given. RTHIM solution metrics for 2015 where the ensemble range and observational uncertainty do not overlap are italicised. The 2015 metrics also appear in graphical form on Fig. 7.

	EN4			ECCO			Obs
	2013	2014	2015	2013	2014	2015	2015
MOC-All (Sv)	14.9 (11.1, 16.6)	16.0 (12.0, 17.8)	17.4 (14.2, 23.0)	19.5 (18.4, 20.6)	18.6 (16.6, 23.0)	*21.4 (18.4, 22.8)*	14.9 (13.7,16.0)
MOC-W (Sv)	3.9 (3.1, 7.6)	4.8 (3.7, 8.0)	3.2 (2.1, 6.5)	6.5 (6.3, 6.7)	7.3 (6.9, 8.3)	*7.5 (7.1, 7.7)*	1.6 (1.1, 2.1)
MOC-E (Sv)	19.0 (15.4, 20.5)	16.2 (12.8, 18.5)	19.2 (15.1, 25.0)	15.2 (14.5, 16.0)	14.3 (13.3, 16.1)	16.2 (13.3, 17.3)	15.1 (14.1, 16.0)
MHT-All (PW)	0.48 (0.43, 0.52)	0.45 (0.39, 0.49)	0.46 (0.41, 0.59)	0.52 (0.50, 0.55)	0.47 (0.44, 0.52)	0.48 (0.41, 0.50)	0.47 (0.45, 0.49)
MHT-W (PW)	0.11 (0.10, 0.14)	0.13 (0.11, 0.17)	0.11 (0.09, 0.14)	0.13 (0.13, 0.13)	0.13 (0.13, 0.14)	0.13 (0.13,0.13)	0.14 (0.13, 0.15)
MHT-E (PW)	0.38 (0.29, 0.42)	0.32 (0.23, 0.37)	0.35 (0.27, 0.47)	0.39 (0.37, 0.42)	0.34 (0.31, 0.39)	0.35 (0.28, 0.37)	0.33 (0.31, 0.35)
MFT-All (Sv)	-0.35 (-0.37, -0.33)	-0.31 (-0.33, -0.29)	-0.30 (-0.37, -0.29)	-0.38 (-0.39, -0.36)	-0.37 (-0.39, -0.35)	-0.35 (-0.36, -0.31)	-0.36 (-0.37, -0.35)
MFT-W (Sv)	-0.22 (-0.23, -0.21)	-0.19 (-0.21, -0.18)	*-0.19 (-0.20, -0.18)*	-0.20 (-0.21,-0.20)	-0.20 (-0.21, -0.17)	*-0.20 (-0.21, -0.20)*	-0.22 (-0.23, -0.21)
MFT-E (Sv)	-0.13 (-0.15, -0.11)	-0.12 (-0.14, -0.10)	-0.12 (-0.17, -0.10)	-0.17 (-0.18. -0.16)	-0.17 (-0.18, -0.15)	-0.15 (-0.16, -0.12)	-0.14 (-0.15, -0.13)

Accepted for publication in *Journal of Physical Oceanography*. DOI 10.1175/JPO-D-19-0188.1.

Table A1. Water mass definitions used in RTHIM which are plotted as boxes on Fig. 2 for EN4 and Fig. A1 for ECCO. The ranges are given as e.g. ‘lower limit, upper limit’.

Datset	Water mass	S range (g/kg)	T range (°C)
EN4	ASW1	29.76, 34.40	-4.81, -0.20
ECCO	ASW1	29.04, 34.40	-2.60, -0.20
EN4	ASW2	34.40, 34.93	-4.81, 1.47
ECCO	ASW2	34.40, 34.93	-2.60, 1.51
EN4	ADW	34.93, 37.19	-4.81, 1.47
ECCO	ADW	34.93, 36.13	-2.60, 1.51
EN4	AAW1	35.14, 35.20	1.47, 3.30
ECCO	AAW1	35.14, 35.20	1.51, 3.30
EN4	AAW2	35.20, 37.19	1.47, 3.98
ECCO	AAW2	35.20, 36.13	1.51, 4.00
EN4	LSW	35.00, 35.10	3.30, 3.98
ECCO	LSW	35.00, 35.10	3.30, 4.00
EN4	OW1	35.10, 35.20	3.30, 3.98
ECCO	OW1	35.10, 35.20	3.30, 4.00
EN4	OW2	35.05, 35.14	1.47, 3.30
ECCO	OW2	35.04, 35.14	1.51, 3.30
EN4	NAW	35.05, 37.19	3.98, 11.79
ECCO	NAW	35.04, 36.13	4.00, 13.21

Accepted for publication in *Journal of Physical Oceanography*. DOI 10.1175/JPO-D-19-0188.1.

Table B1. Definitions of row vectors defining RTHIM TS grids. Each dataset for each year has three different TS grids of varying sizes for sensitivity runs.

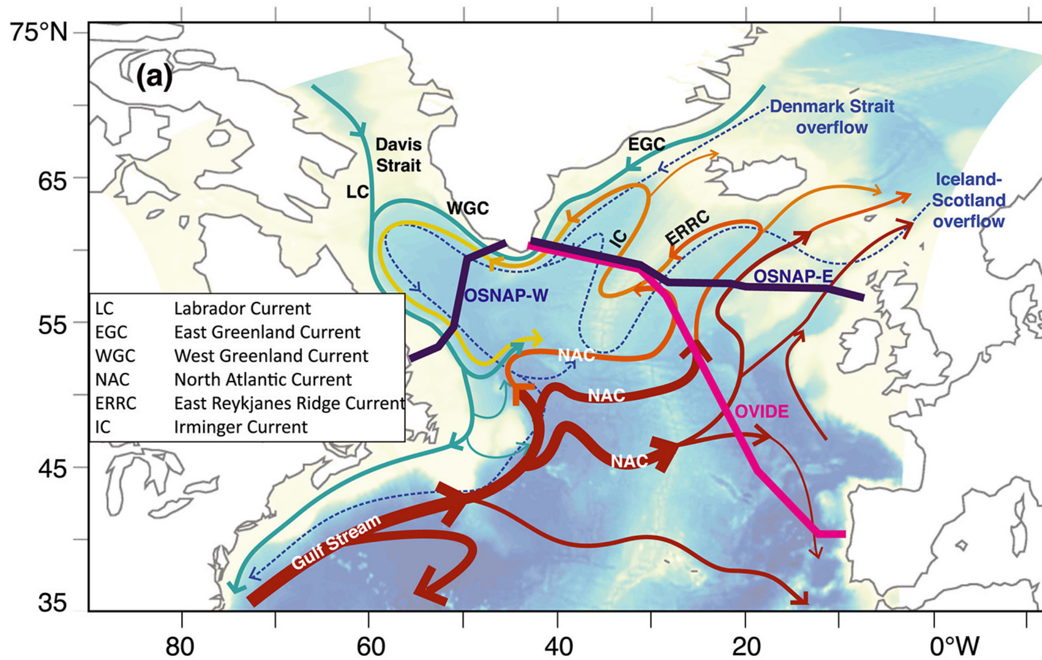
Datset	T vectors	S vectors
EN4 2013	[-5.7 -4.7 -0.6 0.2 1.1 1.9 2.1 3.1 3.6 3.9 4.1 7.8 16.3 21.4 22.4]	[2.65 3.65 26.43 33.93 34.87 34.99 35.01 35.08 35.12 35.16 35.19 35.21 36.19 39.19 40.19]
EN4 2013	[-5.7 -4.7 -0.6 0.2 1.2 1.8 2.2 3.2 3.4 3.9 4.1 6.0 7.7 16.3 21.4 22.4]	[2.65 3.65 26.05 33.93 34.87 34.99 35.01 35.07 35.13 35.15 35.19 35.21 35.59 35.81 39.19 40.19]
EN4 2013	[-5.7 -4.7 -1.4 -0.6 0.2 1.0 2.0 2.6 3.2 3.4 3.5 4.5 7.3 8.7 16.6 21.4 22.4]	[2.65 3.65 27.34 31.56 33.92 34.88 34.99 35.01 35.08 35.12 35.16 35.18 35.22 35.24 36.16 39.19 40.19]
EN4 2014	[-5.3 -4.3 -0.6 0.2 1.2 1.7 2.2 3.0 3.6 3.8 4.2 8.8 14.8 19.8 20.8]	[3.20 4.20 25.63 33.89 34.91 34.95 35.04 35.06 35.13 35.15 35.19 35.21 36.18 38.20 39.20]
EN4 2014	[-5.3 -4.3 -0.6 0.2 1.1 1.9 2.0 3.3 3.4 3.9 4.1 5.3 9.4 15.5 19.8 20.8]	[3.20 4.20 25.84 33.92 34.88 34.98 35.02 35.07 35.13 35.15 35.17 35.24 35.50 35.93 38.20 39.20]
EN4 2014	[-5.3 -4.3 -2.0 -0.6 0.2 1.1 1.9 2.4 3.0 3.6 3.7 4.3 5.7 9.9 15.4 19.8 20.8]	[3.20 4.20 26.61 32.47 33.88 34.91 34.96 35.03 35.06 35.13 35.15 35.17 35.23 35.59 35.82 38.20 39.20]
EN4 2015	[-5.3 -4.3 -0.6 0.2 1.2 1.7 2.2 3.0 3.6 3.8 4.2 8.8 14.8 19.8 20.8]	[3.20 4.20 25.63 33.89 34.91 34.95 35.04 35.06 35.13 35.15 35.19 35.21 36.18 38.20 39.20]
EN4 2015	[-5.3 -4.3 -0.6 0.2 1.1 1.9 2.0 3.3 3.4 3.9 4.1 5.3 9.4 15.5 19.8 20.8]	[3.20 4.20 25.84 33.92 34.88 34.98 35.02 35.07 35.13 35.15 35.17 35.24 35.50 35.93 38.20 39.20]
EN4 2015	[-5.3 -4.3 -2.0 -0.6 0.2 1.1 1.9 2.4 3.0 3.6 3.7 4.3 5.7 9.9 15.4 19.8 20.8]	[3.20 4.20 26.61 32.47 33.88 34.91 34.96 35.03 35.06 35.13 35.15 35.17 35.23 35.59 35.82 38.20 39.20]
ECCO 2013	[-3.1 -2.1 -0.6 0.2 1.0 2.0 2.5 3.2 3.4 3.8 4.2 8.4 15.5 21.7 22.7]	[14.15 15.15 31.43 31.46 33.93 34.87 34.99 35.01 35.07 35.13 35.15 35.19 35.21 36.36 37.36]
ECCO 2013	[-3.1 -2.1 -0.6 0.2 1.1 1.9 2.3 3.0 3.6 3.7 4.4 6.2 11.4 17.1 21.7 22.7]	[14.15 15.15 22.80 28.35 30.70 33.91 34.89 34.97 35.03 35.06 35.13 35.15 35.17 35.23 36.36 37.36]
ECCO 2013	[-3.1 -2.1 -0.0 0.2 1.1 1.9 2.4 3.1 3.5 3.9 4.1 6.7 8.9 12.4 18.3 21.7 22.7]	[14.15 15.15 22.63 27.66 31.91 32.19 33.93 34.87 34.99 35.01 35.07 35.13 35.15 35.17 35.24 36.36 37.36]
ECCO 2014	[-3.1 -2.1 -0.6 0.2 1.3 1.8 2.4 3.2 3.4 3.6 4.4 9.1 17.4 21.1 22.1]	[14.82 15.82 26.30 31.77 33.92 34.88 34.98 35.01 35.07 35.13 35.15 35.17 35.23 35.63 36.63]
ECCO 2014	[-3.1 -2.1 -0.6 0.2 1.1 1.9 2.9 3.2 3.4 3.5 4.5 6.4 10.8 16.0 21.1 22.1]	[14.82 15.82 27.57 31.17 33.93 34.87 34.99 35.01 35.07 35.12 35.16 35.18 35.22 35.35 35.63 36.63]
ECCO 2014	[-3.1 -2.1 -0.6 0.2 1.1 1.9 2.4 3.0 3.6 3.8 4.2 6.0 8.2 12.0 17.9 21.1 22.1]	[14.82 15.82 22.92 27.48 32.09 32.12 33.93 34.87 34.99 35.01 35.07 35.13 35.15 35.17 35.23 35.63 36.63]
ECCO 2015	[-3.1 -2.1 -0.6 0.2 1.3 1.8 2.4 3.2 3.4 3.6 4.4 9.1 17.4 21.1 22.1]	[14.82 15.82 26.3 31.77 33.92 34.88 34.98 35.01 35.07 35.13 35.15 35.17 35.23 35.63 36.63]
ECCO 2015	[-3.1 -2.1 -0.6 0.2 1.1 1.9 2.9 3.2 3.4 3.5 4.5 6.4 10.8 16.0 21.1 22.1]	[14.82 15.82 27.57 31.17 33.93 34.87 34.99 35.01 35.07 35.12 35.16 35.18 35.22 35.35 35.63 36.63]
ECCO 2015	[-3.1 -2.1 -0.6 0.2 1.1 1.9 2.4 3.0 3.6 3.8 4.2 6.0 8.2 12.0 17.9 21.1 22.1]	[14.82 15.82 22.92 27.48 32.09 32.12 33.93 34.87 34.99 35.01 35.07 35.13 35.15 35.17 35.23 35.63 36.63]

Accepted for publication in *Journal of Physical Oceanography*. DOI 10.1175/JPO-D-19-0188.1.

LIST OF FIGURES

Accepted for publication in *Journal of Physical Oceanography*. DOI 10.1175/JPO-D-19-0188.1.

Accepted for publication in *Journal of Physical Oceanography*. DOI 10.1175/JPO-D-19-0188.1.



941 FIG. 1. The circulation of the North Atlantic Subpolar Gyre. Warm, salty water (red arrows) flows northwards, loses heat and joins with colder, fresher waters (blue arrows) flowing southwards. Also marked is the OSNAP section (Overturning in the North Atlantic Subpolar Gyre) and the OVIDE section (Daniault et al. 2016). Figure from Holliday et al. (2018).

Accepted for publication in *Journal of Physical Oceanography*. DOI 10.1175/JPO-D-19-0188.1.

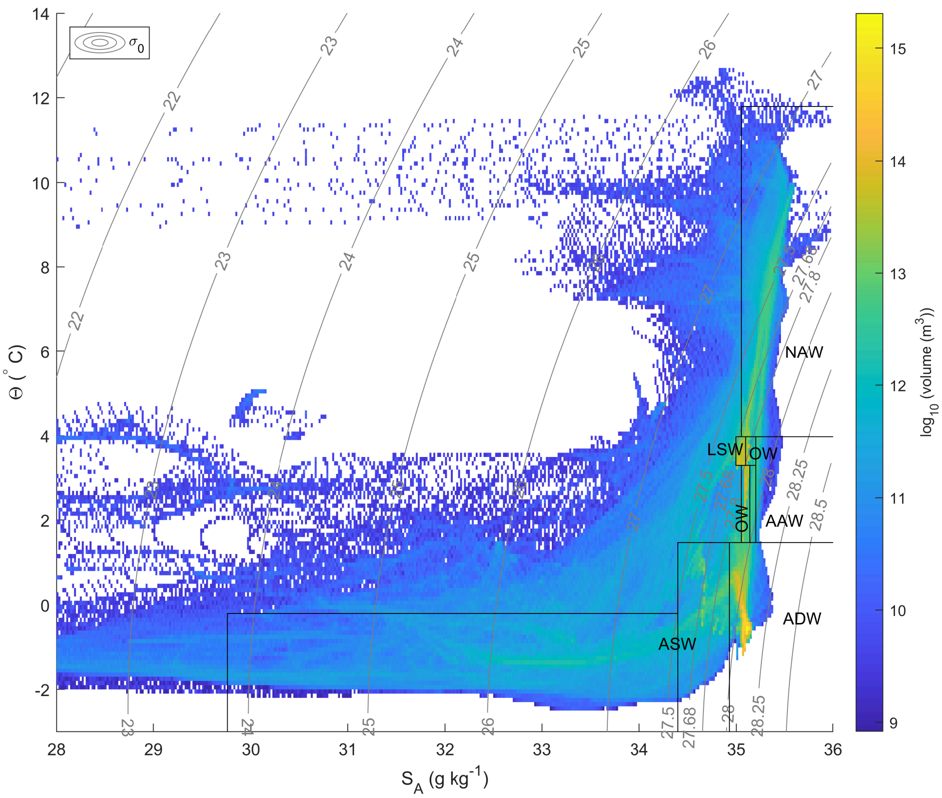
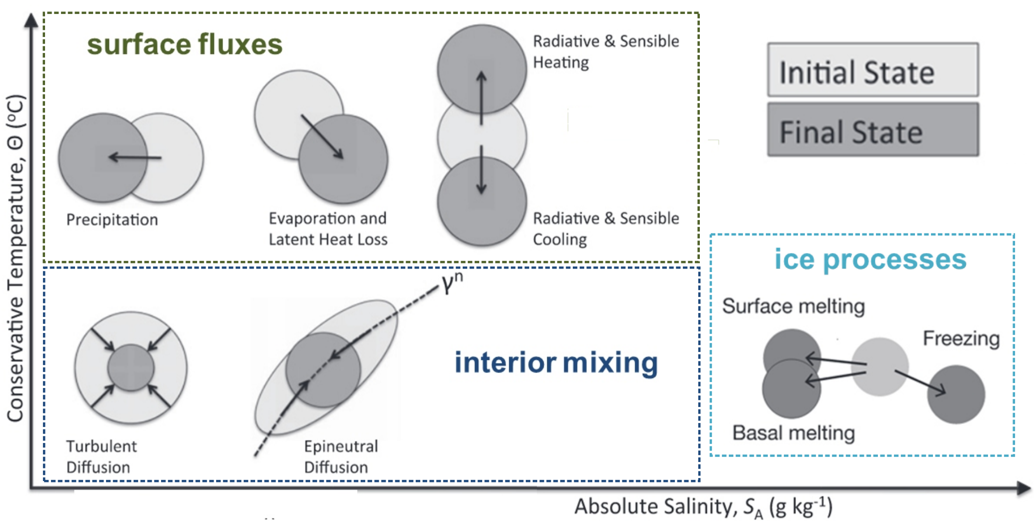


FIG. 2. Volumetric distribution of water masses calculated using the 2015 time-mean of the EN4 objective analysis (Good et al. 2013). The colours represent the volume of water between pairs of isotherms and isohalines, summed over the ocean volume bounded by the section shown on Fig. 4 (note the logarithmic colour scale). The bin sizes are $\Delta\Theta = 0.1°C$ in conservative temperature and $\Delta S_A = 0.02\,g\,kg^{-1}$ in absolute salinity. Overlaid are boxes defining water masses in TS space: North Atlantic Water (NAW), Arctic Atlantic Water (AAW), Arctic Deep Water (ADW), Arctic Surface Water (ASW), Labrador Sea Water (LSW), and overflow water (OW). Potential density contours σ_0 are also overlaid.

Accepted for publication in *Journal of Physical Oceanography*. DOI 10.1175/JPO-D-19-0188.1.



952 FIG. 3. Water mass transformation in TS space by different processes in the ocean. Surface forcing tends
953 to increase the spread of water masses in TS space; mixing tends to bring them together. γ_n indicates a neutral
954 density surface. Figure adapted from Groeskamp et al. (2014a).

Accepted for publication in *Journal of Physical Oceanography*. DOI 10.1175/JPO-D-19-0188.1.

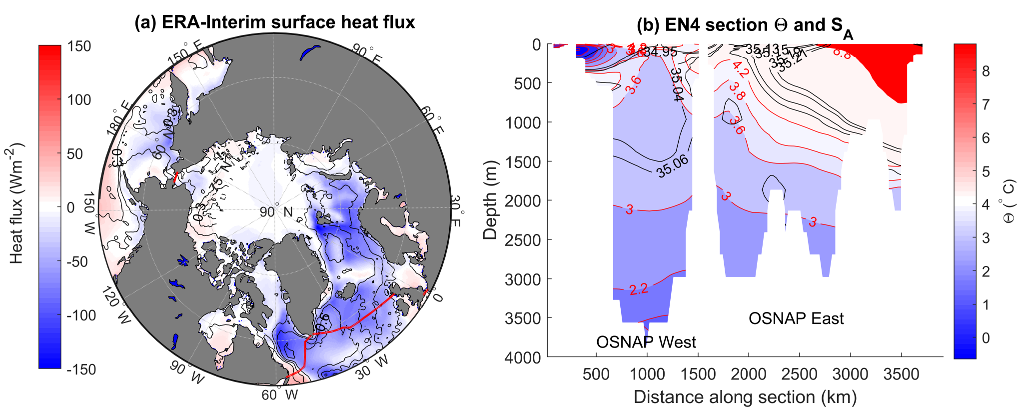


FIG. 4. (a) 2015 mean surface heat flux from ERA-Interim. Overlaid are black contours of absolute dynamic topography from altimetry for the same period. The red line shows the circumpolar section bounding our control volume, following the OSNAP section as indicated on Fig. 1 and with additional sections at Bering Strait and the English Channel to enclose the volume. (b) 2015 mean temperature (colours/red contours) and salinity (black contours) from EN4 for the section shown on (a). The section is in 3 parts: the Bering Strait (shallow section on the left), OSNAP West, and OSNAP East, separated by land (not shown). The English Channel does not appear because it is closed at our grid resolution.

Accepted for publication in *Journal of Physical Oceanography*. DOI 10.1175/JPO-D-19-0188.1.

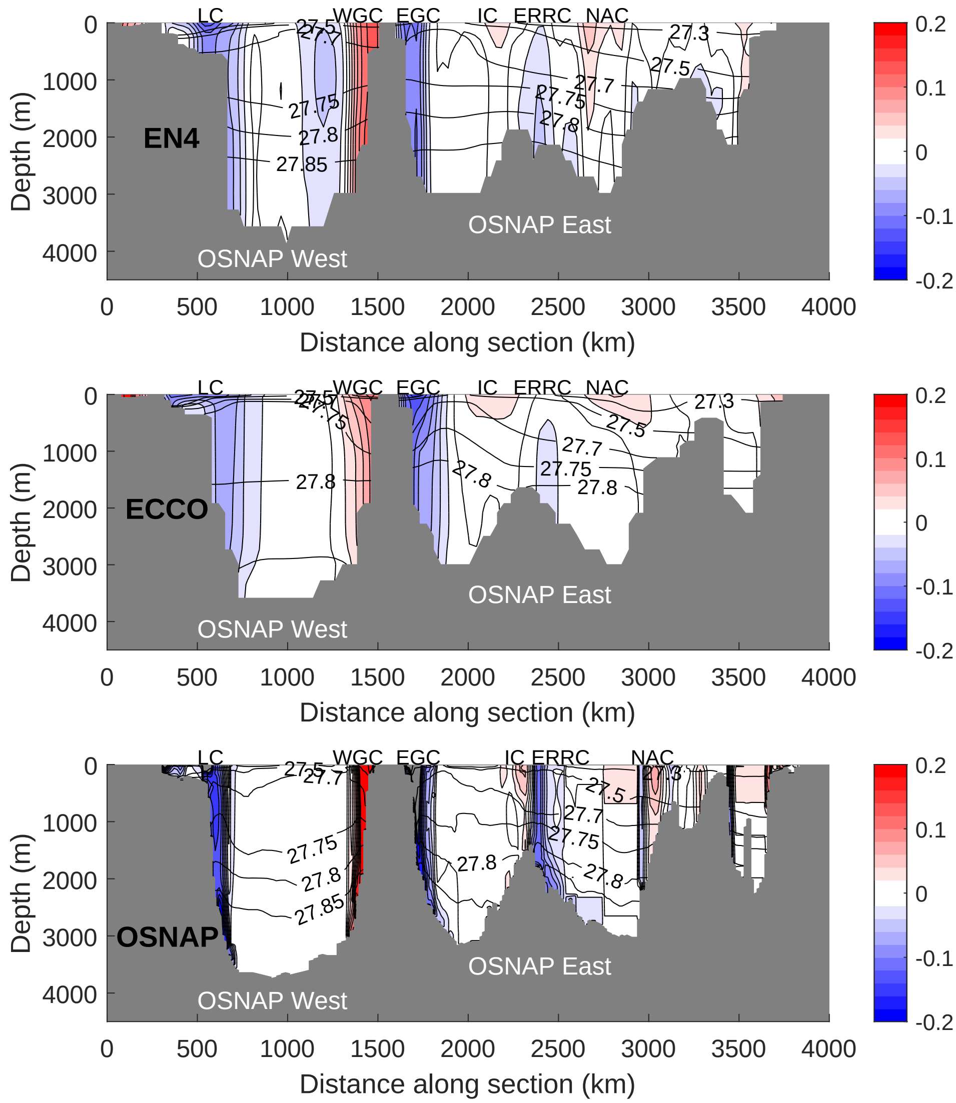
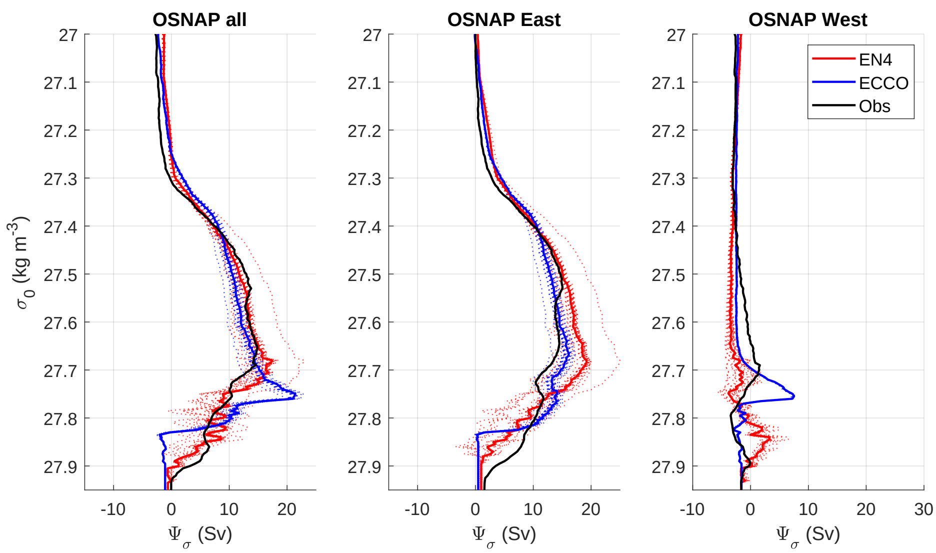


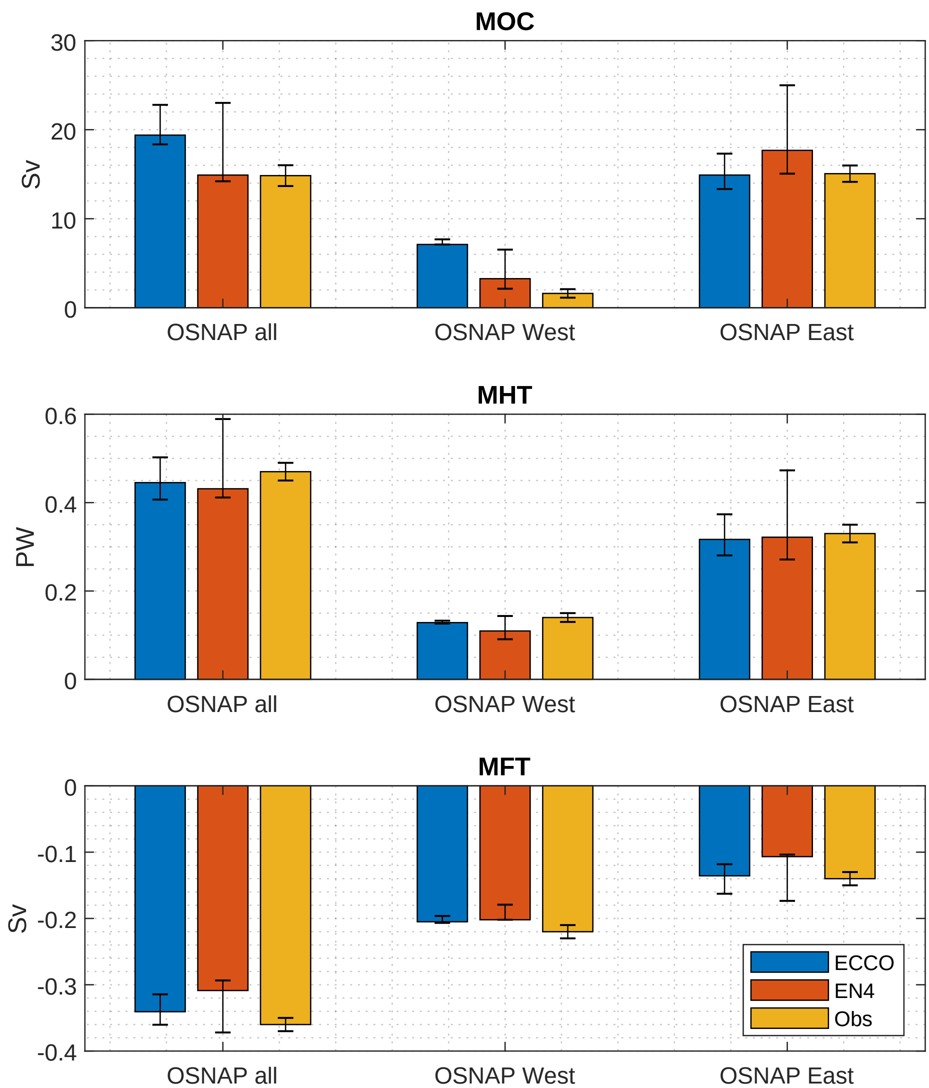
FIG. 5. 2015-mean velocities for the section outlined on Fig. 4 from RTHIM solutions using EN4 (top) and ECCO (middle) datasets, and from the OSNAP observations (bottom). Velocities in ms^{-1} are shown by the background colours with red northward flow (into the control volume) and blue southward flow (out of the volume). Density contours σ_0 are overlaid. The positions of the currents identified on Fig. 1 are labelled at the top of each plot.

Accepted for publication in *Journal of Physical Oceanography*. DOI 10.1175/JPO-D-19-0188.1.



967 FIG. 6. Overturning streamfunctions for the 2015 mean transport through the OSNAP section. Transports are integrated from lower to higher density so that the x-intercept is the total transport through each section part. Individual 'best fit' solutions from the RTHIM ensembles using the EN4 (solid red lines) and ECCO (solid blue lines) datasets are shown along with the OSNAP observations (solid black lines). The RTHIM ensemble members are also plotted as dotted lines in their respective colours for each dataset.

Accepted for publication in *Journal of Physical Oceanography*. DOI 10.1175/JPO-D-19-0188.1.



972 FIG. 7. Meridional overturning circulation (top), meridional heat transport (middle) and meridional freshwater
973 transport (bottom) through the OSNAP section for the year 2015 from RTHIM solutions based on EN4 and
974 ECCO datasets and from the OSNAP observations. For the RTHIM results, the coloured bars represent the
975 individual solution chosen to be the best fit to the ensemble mean (see section 3b), and the error bars give the
976 ensemble range. For the observations, the coloured bars and the error bars are as described in section 3.

Accepted for publication in *Journal of Physical Oceanography*. DOI 10.1175/JPO-D-19-0188.1.

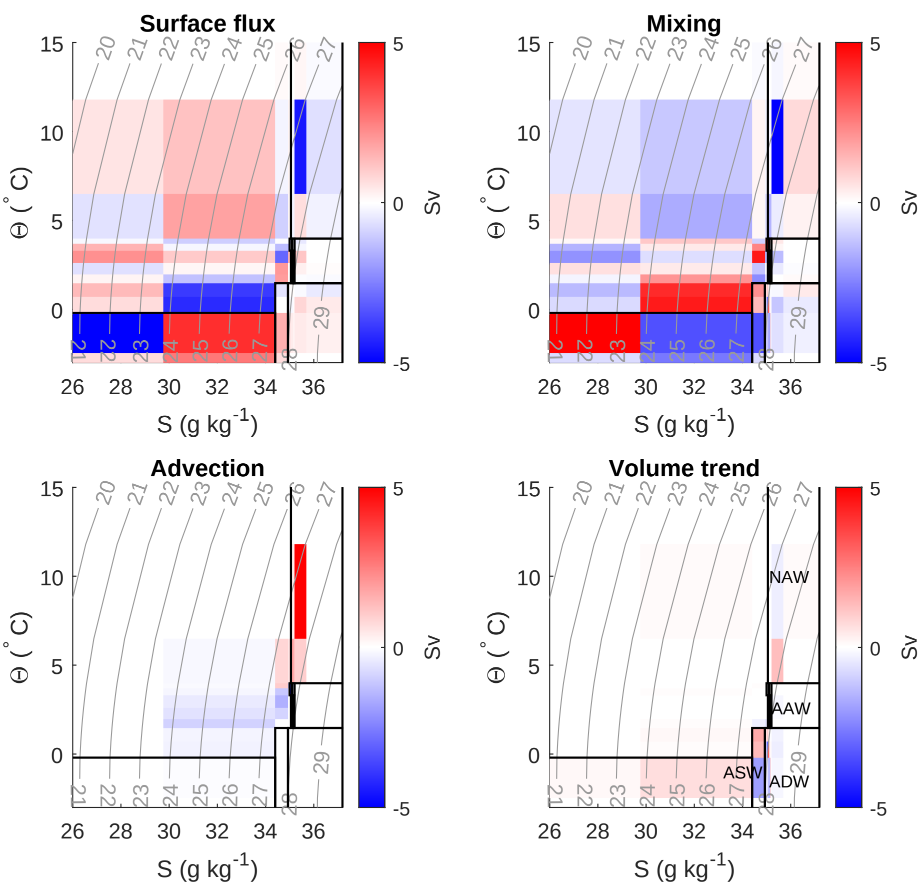
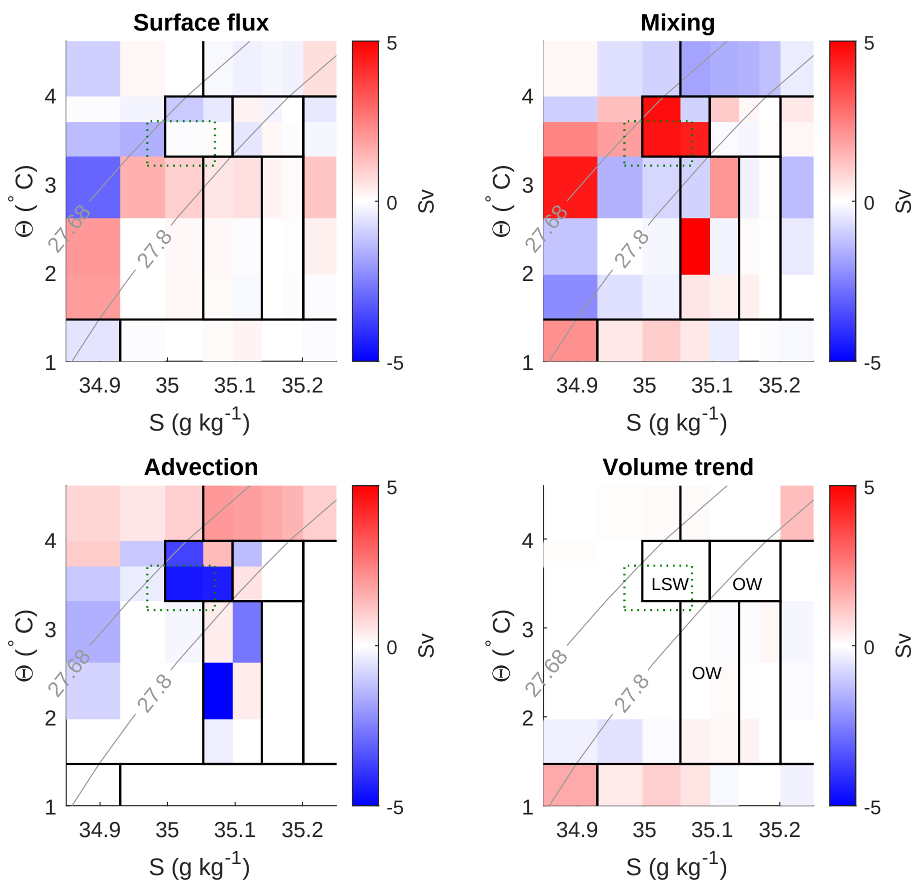


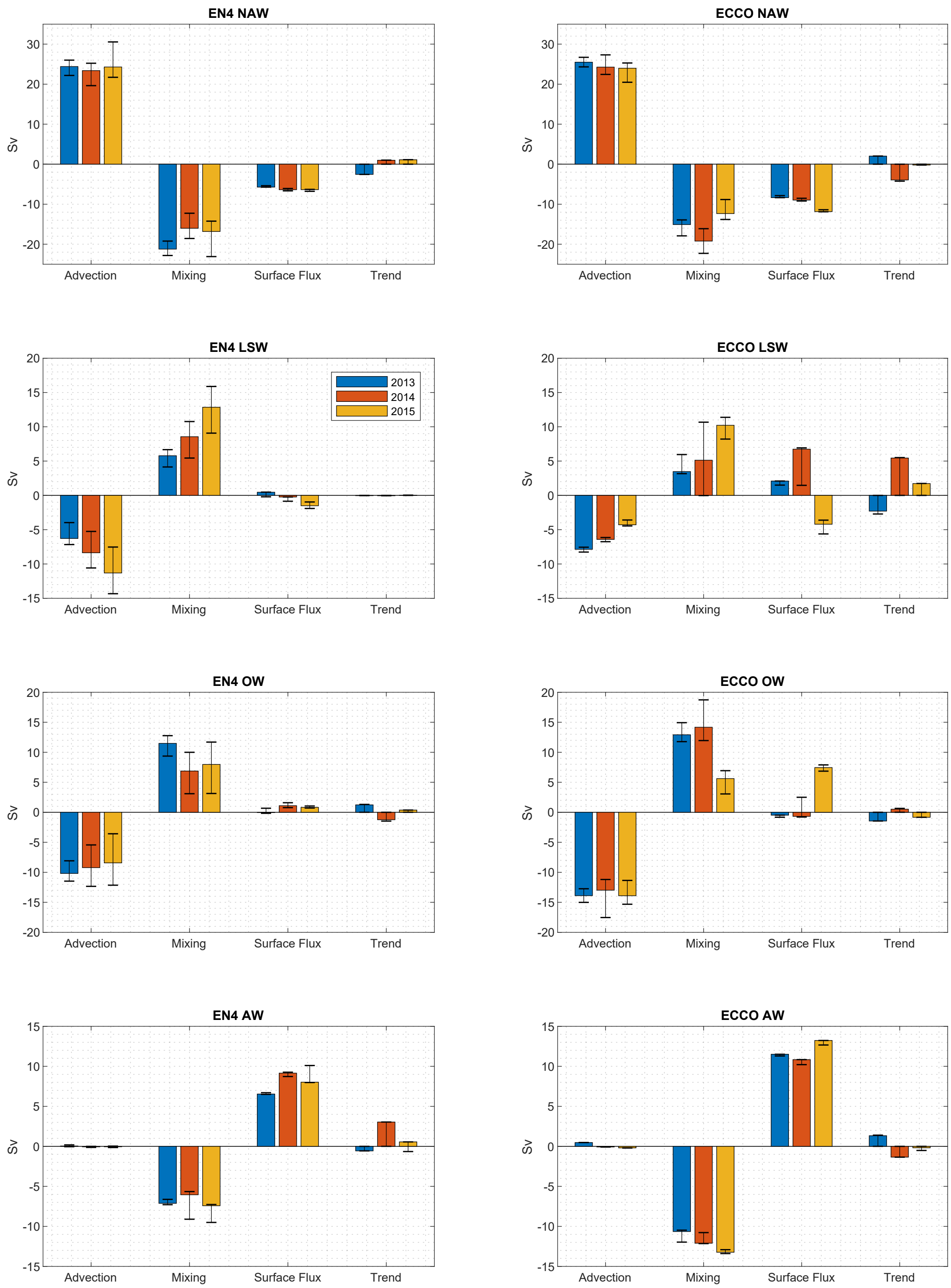
FIG. 8. Volume budget in TS coordinates for the EN4 2015 RTHIM solution. The four panels show the four terms described by Eq. 1, with regions of TS space where there is net formation by a given process shown in red, and regions where there is net destruction shown in blue. The boundaries of our important water masses as defined on Fig. 2 are shown by the black boxes, with water mass labels shown on the volume trend term (NAW = North Atlantic Water, AAW = Arctic/Atlantic Water, ADW = Arctic Deep Water, ASW = Arctic Surface Water). Grey contours are σ_0 potential density surfaces.

Accepted for publication in *Journal of Physical Oceanography*. DOI 10.1175/JPO-D-19-0188.1.



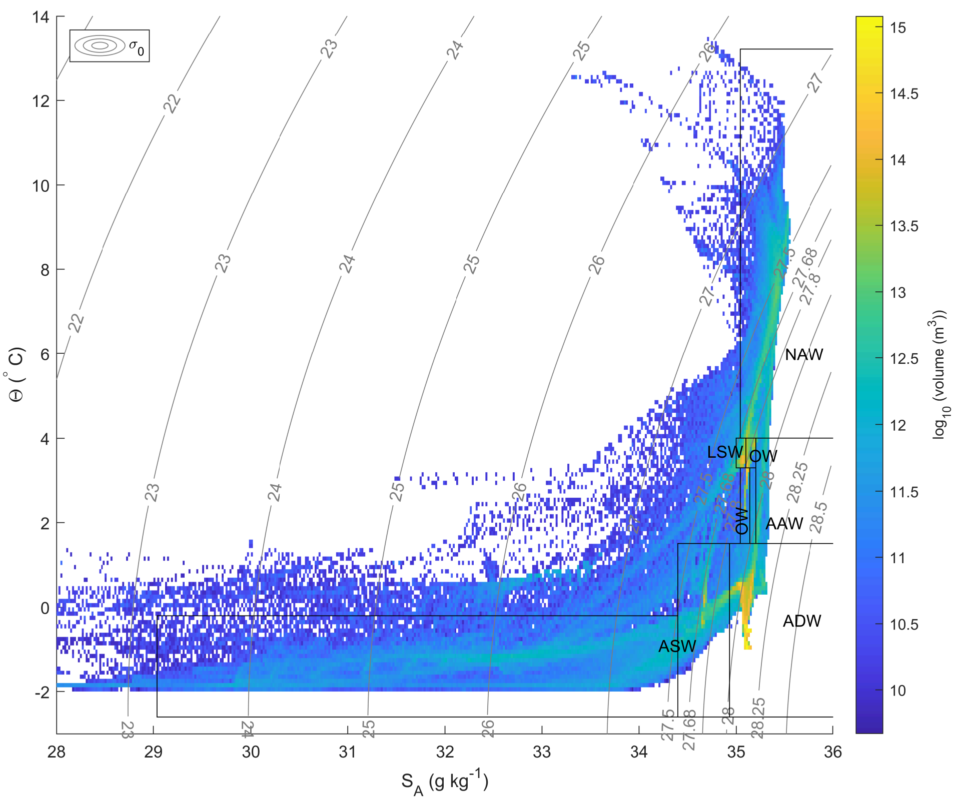
983 FIG. 9. As Fig. 8 but zoomed in on the Labrador Sea Water (LSW) and Overflow Water (OW) water masses
984 as identified on Fig. 2. An alternative definition of the LSW estimated from Fig. 2 of YL2017 is shown by the
985 dotted green boxes.

Accepted for publication in *Journal of Physical Oceanography*. DOI 10.1175/JPO-D-19-0188.1.



986 FIG. 10. Formation rates for each term in the volume budget of Eq. 1 integrated over the water masses
987 defined on Fig. 2, for RTHIM solutions applied to EN4 and ECCO datasets from 2013-2015. The bars give
988 the formation rate for the individual solution that best represents the ensemble mean, and the error bars give the
989 ensemble range.

Accepted for publication in *Journal of Physical Oceanography*. DOI 10.1175/JPO-D-19-0188.1.



990 Fig. A1. As Fig. 2 but using the 2015 time-mean of the ECCO state estimate (Forget et al. 2015; Fukumori
991 et al. 2017). Note that the water mass boundaries differ from those on Fig. 2 due to the different TS grid (see
992 Appendix B).

Accepted for publication in *Journal of Physical Oceanography*. DOI 10.1175/JPO-D-19-0188.1.

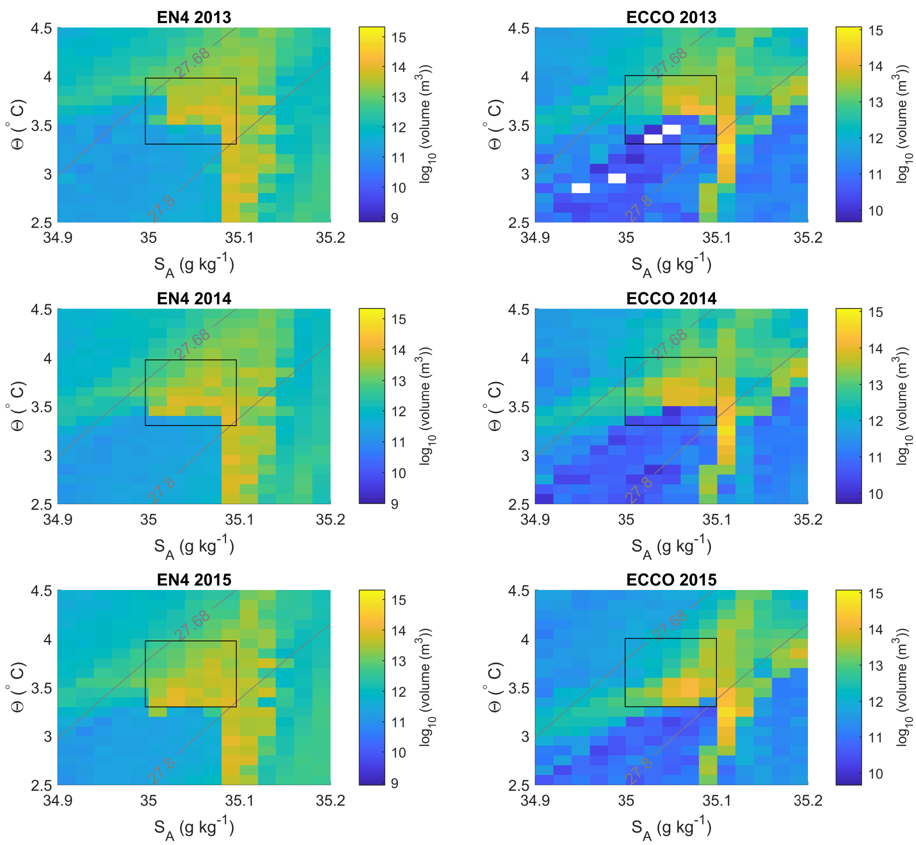


Fig. A2. Volumetric distributions for EN4 and ECCO for 2013, 2014 and 2015 in the region of TS space occupied by the LSW. LSW definition is overlayed as a black box, and grey contours show σ_0.

Accepted for publication in *Journal of Physical Oceanography*. DOI 10.1175/JPO-D-19-0188.1.

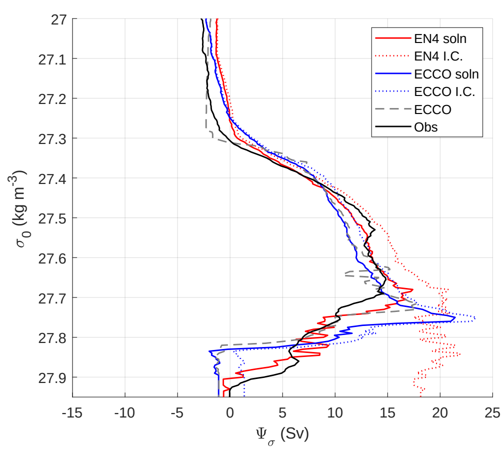


Fig. C1. Overturning streamfunctions for the whole OSNAP section for RTHIM applied to EN4 (red) and ECCO (blue) datasets. The solid lines show individual RTHIM solutions, and the dotted lines show the initial conditions calculated from geostrophic surface velocities (Eq. 3 and thermal wind shear (Eq. 2). Also plotted for comparison are equivalent streamfunctions from the OSNAP observations (black) and from the full ECCO density and velocity fields (grey dashed).

Accepted for publication in *Journal of Physical Oceanography*. DOI 10.1175/JPO-D-19-0188.1.