Assimilation of Argo Temperature and Salinity Profiles Using a Bias-Aware EnOI Scheme for the Labrador Sea

K. ANDREA SCOTT AND CHANGHENG CHEN

Department of Systems Design Engineering, University of Waterloo, Waterloo, Ontario, Canada

PAUL G. MYERS

Department of Earth and Atmospheric Sciences, University of Alberta, Edmonton, Alberta, Canada

(Manuscript received 21 December 2017, in final form 24 July 2018)

ABSTRACT

In this study, temperature and salinity profiles from Argo floats are assimilated into a coupled ice–ocean model over the North Atlantic Ocean and Arctic using an ensemble optimal interpolation (EnOI) scheme, with the aim of improving the thermohaline structure of the Labrador Sea estimated by the model. Data assimilation experiments are carried out from September 2014 to April 2015 both with and without a one-step bias correction method from the literature. It is found that assimilation of the Argo profiles reduces the errors in the model temperature and salinity when verification is done against both withheld Argo profiles and sea surface temperature from satellite data. The assimilation also leads to deeper mixed layer depth in the Labrador Sea, closer to observations shown in other studies, in particular when bias correction is used. We hypothesize that this is because the bias field leads to vertical density profiles that are less stratified, and hence requiring less energy for mixing, than when bias correction is not used.

1. Introduction

The Labrador Sea is one of the few locations in the World Ocean where deep convection occurs. This is the process by which surface waters are brought to depth through vigorous vertical mixing (Lilly et al. 1999). Deep mixing events are initiated through surface buoyancy loss during the winter months, when cold air from the North American landmass blows over the Labrador Sea. These events typically mix a patch covering a surface area of approximately 100–200 km to a depth of 2000 m (Lilly et al. 1999). After a deep mixing event, the region is restratified by convective eddies, which are eddies generated through baroclinic instability resulting from the strong density gradients that surround a patch that has undergone convective mixing (Jones and Marshall 1997; Chanut et al. 2008; Frajka-Williams et al. 2014).

Predicting the depth of convective mixing and the duration of the convective season in the Labrador Sea is important for several reasons. Deep mixing events can lead to the formation of Labrador Sea Water (LSW),

which impacts the southward return branch of the Atlantic meridional overturning circulation, an important component of the global climate system (Marshall and Schott 1999). There is recent evidence that deep convection may play an important role in the global carbon cycle, for example, because of the role of convection in atmosphere–ocean gas exchange (Koelling et al. 2017; Fröb et al. 2016), and because of the link between eddies generated as a result of convection and the spring phytoplankton bloom (Mahadevan et al. 2012).

In spite of recent advances in the sophistication of ocean models and increases in spatial resolution, it is still very difficult to capture the details of convection in the Labrador Sea. For example, the location of the convection region is sensitive to the strength and path of the Irminger eddies, which can be diagnosed through the position and magnitude of the eddy kinetic energy in their formation region off the west coast of Greenland (Chanut et al. 2008). Both the depth and the extent of convection have also been shown to be sensitive to the frequencies available in the atmospheric forcing provided to the ocean model (Holdsworth and Myers 2015), in addition to possible biases in the boundary conditions or forcing (Fenty and Heimbach 2013; Treguier et al. 2005). One

Corresponding author: Andrea Scott, ka3scott@uwaterloo.ca

method to potentially improve the representation of convection provided by a numerical ocean model is through data assimilation. This is a method in which observations are combined with a model state to obtain an improved state estimate (Lahoz et al. 2010). Previous studies have demonstrated the ability of data assimilation to produce ocean state estimates with reduced errors (e.g., Oke and Schiller 2007; Oke et al. 2008; Deng et al. 2010; Deng et al. 2011; Fu et al. 2011; Jones et al. 2012; Pan et al. 2014; Sakov and Sandery 2015). The data assimilation methodology defines these state estimates to be optimal in the sense of minimal variance, but this is true only if a set of assumptions are met. One of the key assumptions is that of unbiased errors for both the background state (which is from the prognostic ocean model) and observations. This assumption can be checked by looking at the bias in the difference between the background state and the observations. If this vector exhibits bias, then either the background state and/or the observations are biased. Typically, based on knowledge of the system, the user chooses to correct the bias in either the background state or the observations. If these biases are not corrected, then the optimal state estimate produced by the data assimilation system will also exhibit bias (Dee 2005).

Systematic model biases, such as would arise in a background state, are common in numerical realizations of the ocean (e.g., Thompson et al. 2006; Zhao and Hendon 2009; Brown et al. 2013; Weisheimer et al. 2014). Previous studies have shown that the temperature and salinity model states exhibit bias in the North Atlantic (Rattan et al. 2010; Treguier et al. 2005; Marzocchi et al. 2015). To address model bias in data assimilation, bias correction schemes can be used (Dee and da Silva 1998; Dee 2005; Chepurin et al. 2005; Lorente-Plazas and Hacker 2017). The two-stage scheme (Dee and da Silva 1998; Dee and Todling 2000) is particularly appealing, since it produces a minimum variance estimate for both the state and the bias. However, the two-stage scheme involves solving an additional equation for the bias vector that is as computationally intensive as that solved for the analysis (optimal state). A reduced version of this scheme, referred to as “one-step bias correction scheme” is widely used (Dee 2005; Deng et al. 2010, 2011). This reduced scheme updates the bias vector by using strictly the bias vector from the previous assimilation cycle and error covariance matrices (i.e., not through solving a minimization problem). In this case, assuming the error covariance matrix for the bias is proportional to that of the background state, with a small constant of proportionality, allows the bias vector to be updated in a computationally inexpensive manner (Dee 2005).

In this study, the one-step bias correction scheme (Dee 2005) is used to assimilate temperature T and salinity S profiles from Argo floats in the Labrador Sea into a coupled ice–ocean model of the North Atlantic and Arctic Oceans. The ensemble optimal interpolation (EnOI) scheme is used for the data assimilation. EnOI is widely represented in ocean data assimilation schemes because of its low computational cost (e.g., Oke et al. 2008; Oke and Schiller 2007). EnOI is a reduced version of the ensemble Kalman filter (EnKF) in the sense that it uses a static ensemble to estimate the background error covariance matrix. The background error covariance matrix is computed using anomalies of the model state from a mean state, typically from a long model run. In this way the background error covariance matrix allows for anisotropic and inhomogeneous state-dependent errors. However, the matrix is static (it is computed in advance of the assimilation) and does not evolve with the model state, as is done in the ensemble Kalman filter. Thus, there is a trade-off between the ability of the ensemble to capture “errors of the day” and the efficiency of the method.

The purpose of the present study is to investigate if the simple one-step bias correction (Dee 2005) in an EnOI system scheme is able to improve the ocean state, with emphasis on the thermohaline structure and convection processes in the Labrador Sea. The application domain, convection in the Labrador Sea, is not widely studied using data assimilation. While other bias correction schemes have been evaluated in the literature, (e.g., Reichle and Koster 2004; Lea et al. 2008; Kumar et al. 2012; Chepurin et al. 2005) and more sophisticated data assimilation methods have been used for ocean state estimation and forecasting (Pan et al. 2014; Fenty and Heimbach 2013; Sakov and Sandery 2015), the present combination was chosen for its simplicity and computational efficiency.

In the next section the ocean model configuration, observational data, and the EnOI scheme are introduced. This is followed by the design of experiments, results, discussion, and concluding remarks.

2. The data assimilation system

a. Model configuration

The model used in this study is the Nucleus for European Modelling of the Ocean (NEMO; Madec 2008), version 3.4, coupled to the Louvain-la-Neuve Sea Ice Model (LIM2; Fichefet and Morales Maqueda 1997; Bouillon et al. 2009). The Arctic and Northern Hemisphere Atlantic (ANHA) configuration used in the present study covers a subdomain of the global ORCA25 model. It

FIG. 1. The nested domain (ANHA4-SPG12) of the Arctic and North Hemisphere Atlantic (dashed lines) and the data assimilation domain (solid lines). The parent domain, ANHA4 (not shown), covers the Atlantic and Arctic Oceans from the Bering Strait to 20°S. The colors represent the depth (m).

consists of a parent grid with a resolution of $1/4^\circ$ (ANHA4), with a nested domain with a resolution of $1/12^\circ$ centered on the North Atlantic (ANHA4-SPG12, where SPG stands for subpolar gyre; black dashed box in Fig. 1). The nesting and two-way communication between the two grids is carried out using the Adaptive Grid Refinement in FORTRAN (AGRIF) toolbox (Debreu et al. 2008). The parent grid covers all of the North Atlantic Ocean north of 27°N in addition to the entire Arctic Ocean, with open boundaries at 27°N and at the Bering Strait. Boundary conditions are provided by the Global Ocean Reanalysis and Simulations (GLORYS2v3). The nested domain is centered on the Labrador Sea, with boundaries along 36°N to the south, 24°W to the east, the Hudson Strait outflow to the west, and Davis Strait to the north. There are 50 unevenly spaced levels in the vertical. The thickness of the first level is 1 m at the surface, with thickness of subsequent levels increasing with depth to a thickness of 450 m at the lowest level. Lateral mixing is carried out using a biharmonic operator with a horizontal eddy viscosity of 1×10^{10} (1.5×10^{11}) $\text{m}^4 \text{s}^{-1}$ in the nested (parent) domain. For tracer lateral diffusion, an isopycnal Laplacian operator is used with a horizontal eddy diffusivity of 50 (300) $\text{m}^2 \text{s}^{-1}$ in the nested (parent) domain. In the nested domain, the resolution is sufficiently fine to be eddy permitting over most of the region of interest, with the first Rossby radius of deformation resolved by 1–2 grid points (Dupont et al. 2015). The model configuration and model spinup is the same as that described in Müller et al. (2017).

Atmospheric forcing is provided by fields from the Global Deterministic Prediction System (GDPS) developed at the Canadian Meteorological Centre (Smith et al. 2014). The atmospheric forcing is refreshed every hour and has a spatial resolution of 33 km. The experiments were run

FIG. 2. Positions of the assimilated Argo profiles (black dots) and the withheld Argo profiles (blue and red dots). The withheld profiles are split into two groups, one that is in the central Labrador Sea (blue dots) and one that is in the coastal region (red dots). The contours in the plot are the bathymetry at 1000, 2000, and 3000 m. The thick black lines denote the data assimilation domain. The region over which the Argo profiles were obtained is shown by the light gray box.

from a spunup state and were run for two years from September 2012 to September 2014 before any data were assimilated.

b. Observational data

1) ARGO DATA

In this study, temperature T and salinity S profiles from Argo observations are assimilated for the period from 19 September 2014 to 30 April 2015. The script used to download the data was generated using the North Atlantic Argo Regional Center information and data mining tool. As the focus of the study is on deep convection, a geographic region centered at 45°W and 56.75°N with a width of 23.7° and a height of 14° was specified to obtain the Argo data. This region is shown in Fig. 2. Both real-time and delayed mode Argo profiles were downloaded from the North Atlantic Argo Regional Center and are subject to automated quality control procedures. To select the profiles for assimilation, each day the availability of Argo observations across the top layer of the model is checked. A maximum number of two profiles are selected for assimilation each day. The rest of the T/S profiles are used for evaluation of the analyses from data assimilation and are referred to as withheld data. These withheld data can be at a geographically distant location from the observations assimilated. For example, an observation can be assimilated in the central Labrador Sea, and a withheld

observation on the same day may be in the Irminger Sea. In this way the withheld data are an assessment of the ability of the model to interpolate the assimilated information in space and time. In total, 321 *T/S* profiles are assimilated and 141 profiles are withheld. Geophysical positions of the Argo profiles are shown in Fig. 2. Data that are within 100 km of any coastline are not assimilated because of large differences between the observed profiles and the background state from the model in these regions.

The spatial distribution of the assimilated dates was evenly spread throughout the region. There were two profiles available for assimilation in the month of September, while the remaining months each had between 40 and 60 profiles for assimilation. The withheld data were also evenly spread throughout the domain, with slightly more data in the Irminger Sea during October and November. Before assimilation, the Argo data are mapped to the model grid. In the horizontal a nearest neighbor interpolation is used, while in the vertical the data are binned according to the vertical grid levels of the model (Jones et al. 2012). The mean of each set of *T* and *S* observations in each bin is the *T* and *S* observation assimilated at the corresponding model grid point. The same approach was used to compare the analyses from the data assimilation with the withheld data.

2) SST

Sea surface temperature (SST) from NOAA's 0.25° daily Optimum Interpolation Sea Surface Temperature, version 2 (OISSTv2; Reynolds et al. 2007), are used to verify SST from the analyses. The NOAA product combines SST from the Advanced Very High Resolution Radiometer (AVHRR), buoys, and ship data. Because of the difference in spatial resolution between the OISSTv2 and the model grid, the model output was interpolated to the observational grid by averaging over boxes of approximately 1/4° spatial resolution.

In addition to SST, each grid point is assigned a sea ice concentration that corresponds to the median value for the week centered on the given date. Sea ice concentration data used in the SST analysis are from the NCEP operational dataset and are based on retrievals from passive microwave brightness temperatures. All points for which the sea ice concentration was greater than zero were removed before comparison with the analyses. This removes the SSTs that are close to land, particularly along the Labrador coast and around Newfoundland, with a narrower strip around Greenland. Argo data are not used in the OISSTv2 product and are hence independent data.

c. The EnOI scheme

The EnOI scheme computes the analysis state vector $\mathbf{x}^a$ using the following equation (Oke et al. 2008):

$$\mathbf{x}^a = \mathbf{x}^b + \mathbf{W}(\mathbf{y}_o - \mathbf{H}\mathbf{x}^b), \quad (1)$$

where $\mathbf{x}^a$ and $\mathbf{x}^b$ denote the analysis and background state vectors, respectively; $\mathbf{y}_o$ is the observation vector; $\mathbf{H}$ is the observation operator that maps the state space to the observation space; and $\mathbf{W}$ is the gain matrix, which is defined according to

$$\mathbf{W} = \mathbf{B}\mathbf{H}^T(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1}. \quad (2)$$

In the gain matrix, $\mathbf{B}$ is the background error covariance matrix and $\mathbf{R}$ is the observation error covariance matrix.

For the background error covariance matrix, the ocean model is used to provide information. In the EnOI method, $\mathbf{B}$ is represented by a stationary ensemble of model perturbations $\mathbf{A}'$,

$$\mathbf{B} = \frac{\alpha}{N-1} \mathbf{A}'\mathbf{A}'^T, \quad (3)$$

where N is the number of ensemble members, $\mathbf{A}'$ is a matrix for which each column is an anomaly from a mean model state (described below), and α is a tunable parameter that controls the magnitude of the background error covariance matrix entries. In this study, $\alpha = 0.05$ is used, which provides a conservative update to the model state (Oke and Schiller 2007). The state vector consists of temperature and salinity at each point in the assimilation domain. The other prognostic model variables, notably u and v , were not included because no data were assimilated for these variables and because computational expense of the data assimilation was an issue. A similar approach has been used in previous studies assimilating *T/S* data using EnOI (Jones et al. 2012). In the interest of computational efficiency, the data assimilation is carried out in a smaller region within the nested domain (the black box in Figs. 1 and 2) that extends from approximately 43° to 70.5°N and from 21° to 69°W).

1) BACKGROUND ERROR COVARIANCE MATRIX

In the EnOI approach, the background error covariance matrix is calculated using the anomaly matrix $\mathbf{A}'$. This matrix is generated by temporal filtering and sampling of a control run, which is a model run in which no data are assimilated. While there is no consensus as how to formulate $\mathbf{A}'$, in general the variability in the anomaly fields used to construct $\mathbf{A}'$ should be representative of the model errors of interest for the problem being studied. In this study Argo data are assimilated once each day. Hence, the errors should represent those that occur on a daily time scale. To represent these errors, a 2-yr control run of the same NEMO model configuration was used, spanning the period from

27 September 2012 to 28 September 2014. From this control run, 121 daily mean model states were selected by sampling every 6 days. From these, 121 anomaly fields were generated by subtracting each daily mean state from its mean. This yields an ensemble of size 121, which is consistent with the ensemble size chosen in other studies (Mignac et al. 2015; Xie and Zhu 2010; Counillon and Bertino 2009; Sakov and Sandery 2015).

2) LOCALIZATION

Because of the limited number of ensemble members, $O(100)$ in comparison to the size of the state vector, $O(10^6)$, spatial localization was used to reduce spurious long-range covariances that arise when computing the background error covariance matrix (Oke et al. 2007; Houtekamer and Mitchell 2001; Buehner 2005). This was carried out by element-wise multiplication of the background error covariance matrix by a localization matrix $\mathbf{L}$. With localization, the gain matrix, which can be obtained by substitution of Eq. (3) into Eq. (2), can be expressed as

$$\mathbf{W} = \alpha[\mathbf{L} \circ (\mathbf{A}'\mathbf{A}'^T)]\mathbf{H}^T \{ \alpha\mathbf{H}[\mathbf{L} \circ (\mathbf{A}'\mathbf{A}'^T)]\mathbf{H}^T + (N-1)\mathbf{R} \}^{-1}, \quad (4)$$

where the open circle denotes an element-by-element matrix multiplication. The i, j element of $\mathbf{L}$, L_{ij} , is the correlation coefficient between the i th and j th grid points of the model and depends only on the physical distance between the two. A Gaussian function with e -folding scale is used to calculate L_{ij} :

$$L_{ij} = \exp\left(\frac{-d_h^2}{2L_h^2} + \frac{-d_v^2}{2L_v^2}\right), \quad (5)$$

where L_h is the horizontal localization length scale, L_v is the vertical localization length scale, d_h is the horizontal distance, and d_v is the vertical distance between the i th and j th grid points. In our experiment, $L_h = 100$ km and $L_v = 750$ m were used. Initial tests with no localization in the vertical dimension showed that in this case an observation could have a significant impact across the entire ocean depth, even though the observations themselves reach a maximum depth of only 2000 m. These large increments deep in the ocean can be attributed to spurious correlations resulting from the limited sample size and are not physical. A small vertical localization length scale (50 m) was then tested. Because the model tends to be too warm and salty, in particular at the surface, this led to a large increment close to the surface and the formation of a cool, fresh layer at the surface, which shut down convection. This outcome is similar to what has been found when eddies with a cool

fresh core in the upper layer enter the convective region. Such eddies have been found to inhibit convection (Schmidt and Send 2007). A vertical localization scale of 750 m was then tested. This led to increments that appeared physically reasonable with no shut down of convection.

In the implementation of the EnOI method, Eq. (4) was modified following Zhu et al. (2011) to a form that is significantly less expensive to calculate. By shifting observation profiles to the closest model grid point during processing, the observation points are coincident with the model grid points, and the gain matrix can be written as

$$\mathbf{W} = \alpha(\mathbf{LH}^T) \circ [\mathbf{A}'(\mathbf{HA}')^T] \times \{ \alpha(\mathbf{HLH}^T) \circ [\mathbf{HA}'(\mathbf{HA}')^T] + (N-1)\mathbf{R} \}^{-1}. \quad (6)$$

3) OBSERVATION ERROR COVARIANCE MATRIX

The observation error covariance matrix $\mathbf{R}$ is assumed to be diagonal, which means spatial correlations and intervariable correlations are neglected in the observation error structure. Given that $\mathbf{R}$ is assumed diagonal, only the diagonal components, or error variances, need to be calculated. Here the observation error variance ε_o^2 specification follows that in Oke et al. (2008) and Deng et al. (2010),

$$\varepsilon_o^2 = \varepsilon_{\text{instr}}^2 + \varepsilon_{\text{RE}}^2, \quad (7)$$

where $\varepsilon_{\text{instr}}$ is the instrument error and ε_{RE} is the representation error. In Oke et al. (2008), there is an additional term to account for the error resulting from the difference between the analysis time and the observation time. This was not taken into account because the assimilation window is relatively short [1 day, as compared to 7–10 days in Oke et al. (2008)]. The instrument errors were set to $\varepsilon_{\text{instr}} = 1.0^\circ\text{C}$ for temperature and $\varepsilon_{\text{instr}} = 0.1$ for salinity at all depths. The representation error ε_{RE} is dependent on the model state following Deng et al. (2010),

$$\varepsilon_{\text{RE}} = \kappa S_{\text{mod}}, \quad (8)$$

where κ is a tunable coefficient and S_{mod} is the model standard deviation calculated from the 2-yr control run. For the results shown, $\kappa = 0.5$ for both temperature and salinity. This means that where $0.5S_{\text{mod}}$ is much greater than the instrument error, the observation error variance will be less than the background error variance and the observations will be given more weight in the analysis than the background state. However, where $0.5S_{\text{mod}}$ is much smaller than the instrument error, the observation error variance will be larger than the background

error variance, and the observations will be given less weight in the analysis than the background state. The former is more likely to occur in regions of high meso-scale variability as well as in the thermocline region, while the latter is more likely to be the case for the deeper, more quiescent, levels of the model. Note that while the instrument error used for temperature is higher than that used in previous studies (Mignac et al. 2015; Turpin et al. 2016), it is the ratio between the observation and background error variances (as well as the spatial structure of the covariance fields) that determines the impact of the observation on the state estimate. For the formulation of the observation error used here, a different value of instrument error could be used with a different value of κ to achieve similar results.

4) BIAS CORRECTION

The one-step bias correction scheme used here solves the following equations for the state and bias estimates at the k th data assimilation cycle,

$$\mathbf{X}_a^k = (\mathbf{X}_b^k - \mathbf{b}^{k-1}) + \mathbf{W}^k [\mathbf{y}_o^k - \mathbf{H}(\mathbf{X}_b^k - \mathbf{b}^{k-1})], \quad (9)$$

$$\mathbf{b}^k = \mathbf{b}^{k-1} - \gamma \mathbf{W}^k [\mathbf{y}_o^k - \mathbf{H}(\mathbf{X}_b^k - \mathbf{b}^{k-1})], \quad (10)$$

where $\mathbf{X}_a^k$, $\mathbf{X}_b^k$, $\mathbf{b}^k$, and $\mathbf{y}_o^k$ denote the analysis and background state vectors, the bias vector, and the observation vector, respectively. In this study $\gamma = 0.01$; thus, the bias estimates evolve slowly with time. For the first assimilation cycle, $\mathbf{b}^0 = 0$ is used (Deng et al. 2010, 2011).

d. Experiment design

In this study three experiments are carried out for the period from 19 September 2014 to 30 April 2015. The first experiment is a control run (CTL), and no observational data are assimilated. The second experiment (ASSIM) assimilates T and S profiles from Argo observations. The third experiment (ASSIM-BC) assimilates the same T and S profiles as in ASSIM but implements the one-step bias correction scheme. Note that both ASSIM and ASSIM-BC can be described by Eq. (9) with $\mathbf{b}^k = 0 \forall k$ for ASSIM. An assimilation window of 24 h (once per day) was used. No special treatment was found necessary to ease the addition of the increment to the background state; that is, the method did not suffer from initialization shock (Jones et al. 2012; Oke et al. 2008). This could be due to the choice of the assimilation window, which is shorter than that used in typical oceanographic data assimilation studies. A shorter assimilation window means the model is less likely to drift as far from the optimal trajectory as compared to when a longer assimilation window is used.

However, given the sparse distribution of the observations in space and time, it is more likely that the lack of initialization shock is due to the small α value used [see Eqs. (3) and (9), and the discussion in Mignac et al. (2015)].

To quantify the skill of the data assimilation schemes, the difference between the analyses and both withheld Argo profiles and independent SST observations is calculated, as described in section 3b.

3. Results

a. Bias reduction

To assess the model bias, the innovations of the data assimilation, which are defined as $\mathbf{y}_o - \mathbf{H}(\mathbf{x}_b)$, are examined (Dee 2005). The innovations for CTL, ASSIM, and ASSIM-BC averaged over the study period show negative values for temperature (Fig. 3a) and salinity (Fig. 3b) at all depths, indicating the model is too warm and salty in comparison with the observations. The innovation for temperature ranges from -1°C at around 2000 m to -3°C in the top layers, and that for salinity ranges from -0.3 at around 2000 m to -0.5 in the top layers. In general, both temperature and salinity are less biased as depth increases below the thermocline. Note that the largest biases are seen for CTL, which makes sense because this is the experiment without data assimilation. A greater magnitude of bias can be seen for ASSIM as compared with ASSIM-BC, because of the effectiveness of the bias correction scheme.

The reduction in both the bias and standard deviation of the state in comparison to the observations resulting from data assimilation can be seen in comparing $\mathbf{y}_o - \mathbf{H}(\mathbf{x}_b)$ with $\mathbf{y}_o - \mathbf{H}(\mathbf{x}_a)$ (lines with and without symbols in Fig. 3). For the temperature the biases are reduced for both ASSIM and ASSIM-BC in the upper 800 m, with a slight reduction in the standard deviation. Note that from the analysis from ASSIM-BC the bias is approximately 1.2°C below 300 m, whereas for CTL the bias is greater than this over all depths until 1200 m. For salinity the reduction in bias in the upper 200 m is less significant, and a small increase can be seen in ASSIM and ASSIM-BC when $\mathbf{y}_o - \mathbf{H}(\mathbf{x}_a)$ is compared with $\mathbf{y}_o - \mathbf{H}(\mathbf{x}_b)$ below the thermocline. The standard deviation is reduced for salinity at all model depths with data assimilation as compared to CTL.

The evolution of the horizontally averaged value of the observation minus background (for CTL) and the observations minus analyses (for ASSIM-BC) at a depth of 110 m with respect to time are shown in Figs. 4a and 4b for salinity and temperature, respectively. The depth of 110 m was chosen because the bias is significant

FIG. 3. Differences between observed and model (a) temperature and (b) salinity over the data assimilation period from 19 Sep 2014 to 30 Apr 2015. CTL (black line), lines with symbols are innovations $y_o - H(\mathbf{x}_b)$, and lines without symbols are $y_o - H(\mathbf{x}_a)$; ASSIM (blue lines) and ASSIM-BC (red lines). Solid lines denote mean values, and dashed lines denote standard deviations.

at that depth. For temperature (Fig. 4a) it can be seen that the bias is greatest at the beginning of the assimilation period for both CTL and ASSIM-BC, with a decreasing trend with respect to time. For both temperature and salinity, the difference between CTL and ASSIM-BC becomes significant at the end of January. Note that the outliers that can be seen in these plots are from Argo profiles in the southern portion of the domain, where there is significant eddy activity from the North Atlantic Current.

b. Evaluation of analyses from data assimilation

1) COMPARISON AGAINST WITHHELD ARGO DATA

To quantify the skill of the data assimilation schemes, the mean difference (MD) between the model analysis and withheld observations is calculated according to Deng et al. (2011),

$$MD = \frac{1}{n} \sum_{i=1}^n (x_a^i - y_o^i), \quad (11)$$

where x_a^i is the model analysis interpolated to the location of the i th observation, y_o^i is the withheld observation, i is the index of the observation, and n is the total number of observations. The MD between the analyses and the withheld data is shown in Fig. 5. It can be seen that the data assimilation is able to reduce the model biases, as shown in Figs. 5a and 5b for temperature and salinity, respectively, for all points in the Labrador Sea interior (blue symbols in Fig. 2). Both ASSIM and ASSIM-BC show smaller values of MD with respect to withheld data than CTL at all model levels. For temperature, ASSIM-BC shows greater reduction in MD than ASSIM as a result of the effectiveness of the bias correction scheme. For temperature it was found that the MD between the different experiments was statistically significant for all depths above 300 m. For salinity the MDs between CTL and both ASSIM and ASSIM-BC are statistically significant above 380 m, while for ASSIM and ASSIM-BC MDs are statistically significant for depths between 50 and 220 m. Mean differences for profiles closer to the coast (red symbols in Fig. 2) for temperature and salinity observations are shown in Figs. 5c and 5d, respectively. It can be seen that MDs are reduced for temperature with data assimilation, but there is no impact of data assimilation on salinity MDs for these locations.

2) COMPARISON AGAINST INDEPENDENT SST OBSERVATIONS

The bias of the difference between the SST from the observations and the SST from the analyses is shown in Fig. 6. It can be seen that there is a negative bias in the central Labrador Sea for CTL that is $\approx -1.5^\circ\text{C}$. A similar bias has been found in other studies (Dupont et al. 2015; Marzocchi et al. 2015). This bias is slightly reduced for ASSIM and is significantly reduced for ASSIM-BC (note the red tones indicate less bias). A time series of the SST bias averaged over the region north of 52°N and west of 36°W is shown in Fig. 7. The region was chosen to reduce the contribution from the southern part of the assimilation domain on the statistics, since the objective of the study was to examine

FIG. 4. Differences between observed (a) temperature and (b) salinity with respect to time at a depth of 110 m. Shown are $y_o - \mathbf{H}(\mathbf{x}_b)$ for CTL (black line) and $y_o - \mathbf{H}(\mathbf{x}_a)$ for ASSIM-BC (red line). Results are similar if a different model level is chosen, with the magnitude of the differences generally decreasing with depth.

the convective region. It can be seen that ASSIM-BC has the lowest bias over the duration of the experiment, and the difference between the bias from ASSIM and ASSIM-BC is sustained at approximately half a degree through March and April. The standard deviation of the differences between the SST from the analyses and that from the observations was also checked. It was found there were no significant differences between the experiments.

c. MLD changes

In this section, we examine how assimilation of Argo temperature and salinity profiles changes the mixed layer depth (MLD). The MLD is defined as the depth at which the potential density difference relative to the value at a reference vertical level near the surface exceeds a threshold value, which is chosen to be 0.01 kg m^{-3} . The MLD is calculated using the code provided by Holte and Talley (2009). We note that this definition has been found to lead to deeper MLDs than other definitions in regions with deep convection as a result of temperature–salinity compensation. We found

the conclusions of our study were the same when other definitions of the MLD were examined.

The average MLD over the period from 1 January 2015 to 31 March 2015 for each experiment is shown in Fig. 8. It can be seen in comparing Fig. 8a with Figs. 8b and 8c that assimilation of the Argo profiles of temperature and salinity leads to deeper convection in the central Labrador Sea. The average MLD from the CTL run is around 1100 m, while ASSIM-BC exhibits deeper convection, with depths in excess of 1300 m, over a larger area. The difference between the MLDs from ASSIM-BC and CTL is shown in Fig. 8d. Fröb et al. (2016) calculated the MLD using a semisubjective method for Argo profiles from February 2015 to April 2015 and showed that the deepest MLD is around 1800 m. Yashayaev and Loder (2016) showed that the deepest MLD in the winter of 2014/15 reaches 1700 m, defined as the 75th percentile of the pycnostad's deepest extent of in situ observations and Argo profiles. Note that ASSIM and ASSIM-BC also show deeper convection along the path of the East Greenland Current and in the Irminger Sea, which is in better agreement with observations (Fröb et al. 2016) than CTL.

FIG. 5. MD between (a),(c) temperature and (b),(d) salinity from the analyses and withheld Argo profiles. MD is averaged over the data assimilation period from 19 Sep 2014 to 30 Apr 2015 and over all times and locations where a withheld Argo profile is available. Withheld profiles in the central Labrador Sea (blue dots in Fig. 2) are shown in (a) and (b), and withheld profiles in the coastal regions (red dots in Fig. 2) are shown in (c) and (d). CTL (black lines), ASSIM (blue lines), and ASSIM-BC (red lines).

The time evolution of the MLD averaged over the rectangular box in Fig. 8a is shown in Fig. 9. For clarity, the MLDs shown are a rolling average over the week centered on the date indicated. It can be seen that data assimilation leads to deeper convection depth throughout the convection season. While the onset of convection occurs at approximately the same time of the year for each experiment, ASSIM and ASSIM-BC show the rate of MLD deepening is faster with data assimilation. The bias correction scheme of ASSIM-BC leads to deeper convection all through the convection season.

Plots of the potential density, temperature, and salinity averaged over the box in Fig. 8a are shown in Fig. 10. It can be seen that changes in the potential density, which lead to the increased rate of deepening of the MLD with data assimilation, are due to changes in both the potential temperature and salinity.

4. Discussion

Without data assimilation, the model is biased toward predicting a warmer and saltier Labrador Sea than is

FIG. 6. Difference between SST from OISSTv2 and data assimilation analyses averaged with respect to time from 14 Sep 2014 to 30 Apr 2015. The areas in white have a nonzero value of sea ice concentration for at least one day of the year. Note that warmer tones denote less bias. The bias in the central part of the Labrador Sea is reduced with data assimilation, and the reduction is more significant for ASSIM-BC than for ASSIM.

FIG. 7. SST bias as a function of time between analyses and OISSTv2 (OBS) over the ANHA4-SPG12 domain north of 52°N and west of 36°W. OBS-CTL (thick line), OBS-ASSIM (thin line), and OBS-ASSIM-BC (thick line with symbols). Note that the bias is reduced significantly for ASSIM-BC.

seen in the Argo data [see Figs. 5 and 6, and the discussion in Rattan et al. (2010) and Marzocchi et al. (2015)]. Based on this observation, two data assimilation experiments were carried out, both with and without a

one-step bias correction scheme from the literature (Dee 2005). This scheme was effective in reducing the model bias and improving the analyses when they are compared to both withheld and independent data, and

FIG. 8. MLD averaged over the period 1 Jan 2015–31 Mar 2015. MLD is computed as the depth at which the density difference from the surface exceeds the threshold value of 0.01 kg m^{-3} . Note in (d) that the MLD for ASSIM-BC is deeper in the central part of the Labrador Sea and in the Irminger Sea, and shallower in the southern part of the Labrador Sea, as compared to the experiment without data assimilation (CTL). The black box is used to highlight differences between MLDs for the various experiments (see discussion in text).

FIG. 9. Time evolution of area-averaged MLDs. The average is taken within the rectangular box shown in Fig. 8a. For clarity, a rolling average is also taken over the 7 days centered on the indicated date. CTL (thick line), ASSIM (thin line), and ASSIM-BC (thick line with symbols). Note that the onset of convection is less gradual for the runs with data assimilation.

led to a deeper mixed layer depth during the convective season as compared to the experiment without bias correction. To understand this further, we note that for the one-step bias correction method, the bias does not evolve with a prognostic equation between assimilation cycles as does the state, which is updated by the ocean model physics. The bias field is updated when a new Argo observation is assimilated through a comparison of the model state with the observations. Because of the use of a spatially correlated weight matrix, the bias update is then spread horizontally and vertically, and is able to influence the model over a larger spatial domain than the set of points at which observations are assimilated. This behavior is analogous to the spreading of the innovation when a spatially correlated background error covariance matrix is used (Buehner 2005). For example, the bias fields at a depth of 110 m for temperature and salinity are shown in Fig. 11 for 30 March 2015. The locations of the Argo profiles assimilated up to and including the specified date are shown with the small black dots. Note that the assimilation of a profile does not necessarily make a significant contribution to the bias field; for example, if there is not a systematic difference between the model state and the observation, then over time the impact of an individual observation on the bias field will be reduced.

The bias has the effect of locally reducing the temperature and salinity of the model state, which may make it more susceptible to deep mixing. To investigate this, we looked for days when the mixed layer depth exhibited evidence of local deep mixing events for which the depth of mixing was different between the experiments with and without bias correction. As an example, the MLD is shown in Fig. 12 for 15 February 2015 for the region outlined with the black box in Fig. 11a. This box was chosen because it is the largest area within the

convection region where the magnitude of both the temperature and salinity bias fields are significant from the end of January until the end of the experiment. It can be seen there is a region for ASSIM-BC (Fig. 12b) where the MLD is deeper than that for ASSIM (Fig. 12a). We also note that the density contours show evidence of strong lateral mixing. An investigation of the potential density profiles before the mixing event indicated that ASSIM-BC exhibited profiles with a less steep slope than ASSIM, indicating they may be more favorable to deep mixing. To quantify this further, we calculated the convective energy that must be removed to allow convection to a given depth, which is defined as (Holdsworth and Myers 2015)

$$CE(h) = \frac{g}{A} \iint_A \left[h \rho_{\text{pot}}(h) - \int_0^h \rho_{\text{pot}}(z) dz \right] dA, \quad (12)$$

and is closely related to the concept of convection resistance (Frajka-Williams et al. 2014). The calculation of the convective energy was carried out at each latitude and longitude location in the black box in Fig. 11a for which there was an increase in MLD from 14 to 15 February of at least 200 m. The probability density functions of the convective energy for both ASSIM and ASSIM-BC are shown in Figs. 13a and 13b, respectively. This calculation used the potential density on 14 February, the day prior to the deep mixing event. Note there are no negative values for convective energy because the mixing parameterizations used in NEMO support only stably stratified density profiles (unstable profiles are immediately mixed). It can be seen in Fig. 13 that the probability of low convective energy is higher for ASSIM-BC than for ASSIM. Similar results were found on other days when deep mixing events were noted to be different between ASSIM and ASSIM-BC. In addition, the results

FIG. 10. Time evolution of (a),(b) potential density, (c),(d) potential temperature, and (e),(f) salinity averaged over the rectangular box shown in Fig. 8a. The MLD (without averaging; green lines) for (left) CTL and (right) ASSIM-BC.

were not sensitive to the exact threshold used for the difference in MLD before and after mixing (thresholds larger than 200 m were also tested and yielded similar results).

5. Conclusions

This study assimilates Argo observations of temperature and salinity into the Labrador Sea in a coupled ice–ocean model of the Arctic and northern Atlantic. A

static background error covariance matrix is computed from a 2-yr free run for the EnOI scheme. The findings were as follows: 1) The EnOI schemes with and without bias correction were successful in improving the model state estimate in comparison with independent data. It should be noted that a vertical localization scale of 750 m was key to this improvement. When a much smaller vertical localization scale (e.g., 50 m) was used, this led to a strongly stratified top layer, which prohibited deep convection. 2) The differences between model

FIG. 11. (a) Temperature and (b) salinity bias fields at a depth level of 110 m on 30 Mar 2015. The black dots indicate the position of the Argo observations assimilated up to 30 Mar 2015. The black box is used to highlight MLD differences between ASSIM and ASSIM-BC (see discussion in text).

predictions and withheld observations of temperature and salinity are smaller with data assimilation than without. 3) The SST bias in the central part of the Labrador Sea is reduced with data assimilation. With bias correction this reduction is statistically significant. 4) MLDs are deeper and closer to other studies in the model runs with data

assimilation (1300 m for ASSIM vs 1100 m for CTL). Other studies (e.g., Fröb et al. 2016; Yashayaev and Loder 2016) show the deepest mixed layer depth in the Labrador Sea is around 1700–1800 m for the winter of 2014/15. 5) The one-step bias correction scheme leads to deeper MLD (1400 m) than when bias correction is not used.

FIG. 12. MLD for (a) ASSIM and (b) ASSIM-BC on 15 Feb 2015 for the region outlined with the black box in Fig. 11a. Also shown are the potential density contours at a depth level of 110 m. The location of a nearby Argo profile that was assimilated on a previous day is also shown by the black dot. However, it should be noted that similar differences between mixing events were also found in this region on days without a coincident Argo observation.

FIG. 13. Probability density function of the convective energy for (a) ASSIM and (b) ASSIM-BC for 15 Feb 2015 for all points in the region outlined with the black box in Fig. 10 for which the MLD deepened by at least 200 m between 14 and 15 Feb 2015.

Considering the sparsity of the Argo data and the relatively short horizontal localization length scales (100 km), it may seem surprising that the data assimilation is able to have an impact. Over the time scales considered in the present study, it is not likely that this

impact is related to the advective nature of the bias (Rattan et al. 2010). It is more likely that the impact is related to the changes in temperature and salinity profiles because of the bias field. In particular, when bias correction is used, this reduces the convective energy required for mixing. Finally, note that other dates were found that exhibited deeper mixing with bias correction, regardless of whether an observation was assimilated on that date. Therefore, the exact timing and location of the observations assimilated may be less critical than having a sufficient number of floats to sample the convection region and to generate a bias field that is able to reduce the convective energy required for deep mixing.

Acknowledgments. This study was supported by the National Science and Engineering Research Council of Canada through the Climate Change and Atmospheric Research Program. The study was carried out as part of the VITALS project, studying ventilation, interactions, transports across the Labrador Sea. We gratefully acknowledge the computing resources provided by Compute Canada. We would also like to thank G. Smith for the CGRF forcing fields, made available from Environment and Climate Change Canada; and G. Garric for the GLORYS2v3 output. Argo float data are managed by the Global Ocean Data Assimilation Experiment (<http://www.usgoda.org/argo/argo.html>) with documentation available (<http://doi.org/10.17882/42182>). We sincerely thank the reviewers for their constructive comments, which significantly improved the quality of the manuscript.

REFERENCES

- Bouillon, S., M. Morales Maqueda, V. Legat, and T. Fichefet, 2009: An elastic–viscous–plastic sea ice model formulated on Arkawa B and C grids. *Ocean Modell.*, **27**, 174–184, <https://doi.org/10.1016/j.ocemod.2009.01.004>.
- Brown, J. N., and Coauthors, 2013: Implications of CMIP3 model biases and uncertainties for climate projections in the western tropical Pacific. *Climatic Change*, **119**, 147–161, <https://doi.org/10.1007/s10584-012-0603-5>.
- Buehner, M., 2005: Ensemble-derived stationary and flow-dependent background-error covariances: Evolution in a quasi-operational NWP setting. *Quart. J. Roy. Meteor. Soc.*, **131**, 1013–1043, <https://doi.org/10.1256/qj.04.15>.
- Chanut, J., B. Barnier, W. Large, L. Debreu, T. Penduff, J. Molines, and P. Mathiot, 2008: Mesoscale eddies in the Labrador Sea and their contribution to convection and restratification. *J. Phys. Oceanogr.*, **38**, 1617–1643, <https://doi.org/10.1175/2008JPO3485.1>.
- Chepurin, G. A., J. A. Carton, and D. Dee, 2005: Forecast model bias correction in ocean data assimilation. *Mon. Wea. Rev.*, **133**, 1328–1342, <https://doi.org/10.1175/MWR2920.1>.
- Counillon, F., and L. Bertino, 2009: Ensemble optimal interpolation: Multivariate properties in the Gulf of Mexico. *Tellus*, **61A**, 296–308, <https://doi.org/10.1111/j.1600-0870.2008.00383.x>.

- Debreu, L., C. Voulard, and E. Blayo, 2008: AGRIF: Adaptive grid refinement in Fortran. *Comput. Geosci.*, **34**, 8–13, <https://doi.org/10.1016/j.cageo.2007.01.009>.
- Dee, D. P., 2005: Bias and data assimilation. *Quart. J. Roy. Meteor. Soc.*, **131**, 3323–3343, <https://doi.org/10.1256/qj.05.137>.
- , and A. M. da Silva, 1998: Data assimilation in the presence of forecast bias. *Quart. J. Roy. Meteor. Soc.*, **124**, 269–295, <https://doi.org/10.1002/qj.49712454512>.
- , and R. Todling, 2000: Data assimilation in the presence of forecast bias: The GEOS moisture analysis. *Mon. Wea. Rev.*, **128**, 3268–3282, [https://doi.org/10.1175/1520-0493\(2000\)128<3268:DAITPO>2.0.CO;2](https://doi.org/10.1175/1520-0493(2000)128<3268:DAITPO>2.0.CO;2).
- Deng, Z., Y. Tang, and G. Wang, 2010: Assimilation of Argo temperature and salinity profiles using a bias-aware localized EnKF system for the Pacific Ocean. *Ocean Modell.*, **35**, 187–205, <https://doi.org/10.1016/j.ocemod.2010.07.007>.
- , —, and H. J. Freeland, 2011: Evaluation of several model error schemes in the EnKF assimilation: Applied to Argo profiles in the Pacific Ocean. *J. Geophys. Res.*, **116**, C09027, <https://doi.org/10.1029/2011JC006942>.
- Dupont, F., and Coauthors, 2015: A high-resolution ocean and sea ice modelling system for the Arctic and North Atlantic oceans. *Geosci. Model Dev.*, **8**, 1577–1594, <https://doi.org/10.5194/gmd-8-1577-2015>.
- Fenty, I., and P. Heimbach, 2013: Coupled sea–ice ocean state estimation in the Labrador Sea. *J. Phys. Oceanogr.*, **43**, 884–904, <https://doi.org/10.1175/JPO-D-12-065.1>.
- Fichefet, T., and M. Morales Maqueda, 1997: Sensitivity of a global sea ice model to the treatment of ice thermodynamics and dynamics. *J. Geophys. Res.*, **102**, 12 609–12 646, <https://doi.org/10.1029/97JC00480>.
- Frajka-Williams, E., P. Rhines, and C. Eriksen, 2014: Horizontal stratification during deep convection in the Labrador Sea. *J. Phys. Oceanogr.*, **44**, 220–228, <https://doi.org/10.1175/JPO-D-13-069.1>.
- Fröb, F., A. Olsen, K. Våge, G. Moore, I. Yashayaev, E. Jeansson, and B. Rajasakaren, 2016: Irminger Sea deep convection injects oxygen and anthropogenic carbon to the ocean interior. *Nat. Commun.*, **7**, 13244, <https://doi.org/10.1038/ncomms13244>.
- Fu, W., J. She, and S. Zhuang, 2011: Application of an Ensemble Optimal Interpolation in a North/Baltic Sea model: Assimilating temperature and salinity profiles. *Ocean Modell.*, **40**, 227–245, <https://doi.org/10.1016/j.ocemod.2011.09.004>.
- Holdsworth, A., and P. Myers, 2015: The influence of high-frequency atmospheric forcing on the circulation and deep convection of the Labrador Sea. *J. Climate*, **28**, 4980–4996, <https://doi.org/10.1175/JCLI-D-14-00564.1>.
- Holte, J., and L. Talley, 2009: A new algorithm for finding mixed layer depths with application to Argo data and Subantarctic Mode Water formation. *J. Atmos. Oceanic Technol.*, **26**, 1920–1939, <https://doi.org/10.1175/2009JTECH0543.1>.
- Houtekamer, P., and L. Mitchell, 2001: A sequential ensemble Kalman filter for atmospheric data assimilation. *Mon. Wea. Rev.*, **129**, 123–137, [https://doi.org/10.1175/1520-0493\(2001\)129<0123:ASEKFF>2.0.CO;2](https://doi.org/10.1175/1520-0493(2001)129<0123:ASEKFF>2.0.CO;2).
- Jones, E. M., P. R. Oke, F. Rizwi, and L. M. Murray, 2012: Assimilation of glider and mooring data into a coastal ocean model. *Ocean Modell.*, **47**, 1–13, <https://doi.org/10.1016/j.ocemod.2011.12.009>.
- Jones, H., and J. Marshall, 1997: Restratification after deep convection. *J. Phys. Oceanogr.*, **27**, 2276–2287, [https://doi.org/10.1175/1520-0485\(1997\)027<2276:RADC>2.0.CO;2](https://doi.org/10.1175/1520-0485(1997)027<2276:RADC>2.0.CO;2).
- Koelling, J., D. Wallace, U. Send, and J. Karstensen, 2017: Intense oceanic update of oxygen during 2014–2015 winter convection in the Labrador Sea. *Geophys. Res. Lett.*, **44**, 7855–7864, <https://doi.org/10.1002/2017GL073933>.
- Kumar, S. V., R. H. Reichle, K. W. Harrison, C. D. Peters-Lidard, S. Yatheendradas, and J. A. Santanello, 2012: A comparison of methods for a priori bias correction in soil moisture data assimilation. *Water Resour. Res.*, **48**, W03515, <https://doi.org/10.1029/2010WR010261>.
- Lahoz, W., B. Khattatov, and R. Ménard, 2010: Data assimilation and information. *Data Assimilation: Making Sense of Observations*, W. Lahoz, B. Khattatov, and R. Ménard, Eds., Springer, 3–12, https://doi.org/10.1007/978-3-540-74703-1_1.
- Lea, D. J., J.-P. Drecourt, K. Haines, and M. J. Martin, 2008: Ocean altimeter assimilation with observational- and model-bias correction. *Quart. J. Roy. Meteor. Soc.*, **134**, 1761–1774, <https://doi.org/10.1002/qj.320>.
- Lilly, J., P. Rhines, M. Visbeck, R. Davis, J. Lazier, F. Schott, and D. Farmer, 1999: Observing deep convection in the Labrador Sea during winter 1994/95. *J. Phys. Oceanogr.*, **29**, 2065–2098, [https://doi.org/10.1175/1520-0485\(1999\)029<2065:ODCITL>2.0.CO;2](https://doi.org/10.1175/1520-0485(1999)029<2065:ODCITL>2.0.CO;2).
- Lorente-Plazas, R., and J. P. Hacker, 2017: Observation and model bias estimation in the presence of either or both sources of error. *Mon. Wea. Rev.*, **145**, 2683–2696, <https://doi.org/10.1175/MWR-D-16-0273.1>.
- Madec, G., 2008: NEMO ocean engine. Note due Pôle de Modélisation de L'Institut Pierre-Simon Laplace, 357 pp.
- Mahadevan, A., E. D'Asaro, C. Lee, and M. J. Perry, 2012: Eddy-driven stratification initiates North Atlantic spring phytoplankton blooms. *Science*, **337**, 54–58, <https://doi.org/10.1126/science.1218740>.
- Marshall, J., and F. Schott, 1999: Open-ocean convection: Observations, theory, and models. *Rev. Geophys.*, **37**, 1–64, <https://doi.org/10.1029/98RG02739>.
- Marzocchi, A., J. J.-M. Hirschi, N. Holliday, S. Cunningham, A. Blaker, and A. Coward, 2015: The North Atlantic subpolar circulation in an eddy-resolving global ocean model. *J. Mar. Syst.*, **142**, 126–143, <https://doi.org/10.1016/j.jmarsys.2014.10.007>.
- Mignac, D., C. A. S. Tanajura, A. N. Santana, L. N. Lima, and J. Xie, 2015: Argo data assimilation into HYCOM with an EnOI method in the Atlantic Ocean. *Ocean Sci.*, **11**, 195–213, <https://doi.org/10.5194/os-11-195-2015>.
- Müller, V., D. Kieke, P. Myers, C. Pennelly, and C. Martens, 2017: Temperature flux carried by individual eddies across 47°N in the Atlantic Ocean. *J. Geophys. Res. Oceans*, **122**, 2441–2464, <https://doi.org/10.1002/2016JC012175>.
- Oke, P. R. and A. Schiller, 2007: Impact of Argo, SST, and altimeter data on an eddy-resolving ocean reanalysis. *Geophys. Res. Lett.*, **34**, L19601, <https://doi.org/10.1029/2007GL031549>.
- , P. Sakov, and S. P. Corney, 2007: Impacts of localisation in the EnKF and EnOI: Experiments with a small model. *Ocean Dyn.*, **57**, 32–45, <https://doi.org/10.1007/s10236-006-0088-8>.
- , G. B. Brassington, D. A. Griffin, and A. Schiller, 2008: The Bluelink ocean data assimilation system (BODAS). *Ocean Modell.*, **21**, 46–70, <https://doi.org/10.1016/j.ocemod.2007.11.002>.
- Pan, C., L. Zheng, R. H. Weisberg, Y. Liu, and C. E. Lembke, 2014: Comparisons of different ensemble schemes for glider data assimilation on West Florida Shelf. *Ocean Modell.*, **81**, 13–24, <https://doi.org/10.1016/j.ocemod.2014.06.005>.
- Rattan, S., P. Myers, A.-M. Treguier, S. Theetten, A. Biastock, and C. Böning, 2010: Towards an understanding of Labrador Sea salinity drift in eddy-permitting simulations. *Ocean Modell.*, **35**, 77–88, <https://doi.org/10.1016/j.ocemod.2010.06.007>.
- Reichle, R. H. and R. D. Koster, 2004: Bias reduction in short records of satellite soil moisture. *Geophys. Res. Lett.*, **31**, L19501, <https://doi.org/10.1029/2004GL020938>.

- Reynolds, R. W., T. M. Smith, C. Liu, D. B. Chelton, K. S. Casey, and M. G. Schlax, 2007: Daily high-resolution-blended analyses for sea surface temperature. *J. Climate*, **20**, 5473–5496, <https://doi.org/10.1175/2007JCLI1824.1>.
- Sakov, P., and P. A. Sandery, 2015: Comparison of EnOI and EnKF regional ocean reanalysis systems. *Ocean Modell.*, **89**, 45–60, <https://doi.org/10.1016/j.ocemod.2015.02.003>.
- Schmidt, S., and U. Send, 2007: Origin and composition of seasonal Labrador Sea freshwater. *J. Phys. Oceanogr.*, **37**, 1445–1454, <https://doi.org/10.1175/JPO3065.1>.
- Smith, G., F. Roy, P. Mann, F. Dupont, F. Brasnett, J.-F. Lemieux, S. Laroche, and S. Blair, 2014: A new atmospheric dataset for forcing ice-ocean models: Evaluation of reforecasts using the Canadian global deterministic prediction system. *Quart. J. Roy. Meteor. Soc.*, **140**, 881–894, <https://doi.org/10.1002/qj.2194>.
- Thompson, K. R., D. G. Wright, Y. Lu, and E. Demirov, 2006: A simple method for reducing seasonal bias and drift in eddy resolving ocean models. *Ocean Modell.*, **13**, 109–125, <https://doi.org/10.1016/j.ocemod.2005.11.003>.
- Treguier, A., S. Theetten, E. Chassignet, T. Penduff, R. Smith, L. Talley, J. Beismann, and C. Böning, 2005: The North Atlantic subpolar gyre in four high-resolution models. *J. Phys. Oceanogr.*, **35**, 757–774, <https://doi.org/10.1175/JPO2720.1>.
- Turpin, V., E. Remy, and P.-Y. Le Traon, 2016: How essential are Argo observations to constrain a global ocean data assimilation system? *Ocean Sci.*, **12**, 257–274, <https://doi.org/10.5194/os-12-257-2016>.
- Weisheimer, A., S. Corti, T. Palmer, and F. Vitart, 2014: Addressing model error through atmospheric stochastic physical parametrizations: Impact on the coupled ECMWF seasonal forecasting system. *Philos. Trans. Roy. Soc. London*, **372A**, 20130290, <https://doi.org/10.1098/rsta.2013.0290>.
- Xie, J., and J. Zhu, 2010: Ensemble optimal interpolation schemes for assimilating Argo profiles into a hybrid coordinate ocean model. *Ocean Modell.*, **33**, 283–298, <https://doi.org/10.1016/j.ocemod.2010.03.002>.
- Yashayaev, I., and W. Loder, 2016: Recurrent replenishment of Labrador Sea Water and associated decadal-scale variability. *J. Geophys. Res. Oceans*, **121**, 8095–8114, <https://doi.org/10.1002/2016JC012046>.
- Zhao, M., and H. H. Hendon, 2009: Representation and prediction of the Indian Ocean dipole in the POAMA seasonal forecast model. *Quart. J. Roy. Meteor. Soc.*, **135**, 337–352, <https://doi.org/10.1002/qj.370>.
- Zhu, J., F. Zheng, and X. Li, 2011: A new localization implementation scheme for ensemble data assimilation of non-local observations. *Tellus*, **63A**, 244–255, <https://doi.org/10.1111/j.1600-0870.2010.00486.x>.