

Simulations of Hydrographic Properties in the Northwestern North Atlantic Ocean in Coupled Climate Models

M. F. DE JONG

Royal Netherlands Institute for Sea Research, Den Burg, Netherlands

S. S. DRIJFHOUT AND W. HAZELEGER

Royal Netherlands Meteorological Institute, De Bilt, Netherlands

H. M. VAN AKEN

Royal Netherlands Institute for Sea Research, Den Burg, Netherlands

C. A. SEVERIJNS

Royal Netherlands Meteorological Institute, De Bilt, Netherlands

(Manuscript received 28 January 2008, in final form 7 October 2008)

ABSTRACT

The performance of coupled climate models (CCMs) in simulating the hydrographic structure and variability of the northwestern North Atlantic Ocean, in particular the Labrador and Irminger Seas, has been assessed. This area plays an important role in the meridional overturning circulation. Hydrographic properties of the preindustrial run of eight CCMs used in the Intergovernmental Panel on Climate Change (IPCC) Fourth Assessment Report (AR4) are compared with observations from the World Ocean Circulation Experiment Repeat Section 7 (WOCE AR7). The mean and standard deviation of 20 yr of simulated data are compared in three layers, representing the surface waters, intermediate waters, and deep waters. Two models simulate an extremely cold, fresh surface layer with model biases down to -1.7 psu and -4.0°C , much larger than the observed ranges of variability. The intermediate and deep layers are generally too warm and saline, with biases up to 0.7 psu and 2.8°C . An analysis of the maximum mixed layer depth shows that the low surface salinity is related to a convective regime restricted to the upper 500 dbar. Thus, intermediate water formed by convection is partly replaced by warmer water from the south. Model biases seem to be caused by the coupling to the atmospheric component of the CCM. Model drift during long spinup periods allows the initially small biases in water mass characteristics to become significant. Biases that develop in the control run are carried over to the twentieth-century runs, which are initialized from the control runs.

1. Introduction

The performance of coupled ocean–atmosphere climate models (CCMs) in simulating the present climate has improved greatly during recent years. This is partly because of increased model resolution and a better understanding of physical processes. However, the resolution of the ocean component of global models is often still limited to about $1^\circ \times 1^\circ$ (Table 1), which is about

$90 \text{ km} \times 90 \text{ km}$ at 60°N . At this resolution many small-scale ocean processes and boundary currents, which play an important role in the northwestern North Atlantic Ocean, cannot be resolved. Surface fluxes, which have a poor observational coverage over the ocean, have a major role in local adjustment of water masses and the formation of mode waters (Brambilla and Talley 2008). The surface waters are furthermore dependent on a combination of sea ice melt and advection. Convection plumes ($\sim 1 \text{ km}$) and convection areas ($\sim 100 \text{ km}$) are crucial in intermediate water mass formation (Marshall and Schott 1999). Mesoscale eddies ($\sim 10\text{--}50 \text{ km}$), which are a large source of variability (Volkov 2005), bring boundary current water to the centers of the basins and

Corresponding author address: M. F. de Jong, Royal Netherlands Institute for Sea Research, Landsdiep 4, 1797 SZ, Den Burg, Netherlands.
E-mail: jong@nioz.nl

TABLE 1. Information of global climate models, reanalyses, and the stand-alone ocean model of which data were used in the analysis. The resolution of the ocean component of the models is denoted in degrees lon \times degrees lat \times the number of vertical levels. The spinup time is denoted in years.

Global climate model	Resolution	Spinup	Initial conditions	References
Bjerknes Centre for Climate Research Bergen Climate Model version 2.0 (BCCR-BCM2.0)	$1.5 \times 1.5 \times 35$	80	Levitus (1982)	Furevik et al. (2003)
CNRM-CM3	$2 \times 2 \times 31$	70	Previous model	Salas-Mélia et al. (2005)
GFDL CM2.0	$1 \times 1 \times 50$	300	Levitus (1992)	Delworth et al. (2006)
IPSL-CM4	$2 \times 2 \times 31$	330	Levitus (1992)	Marti et al. (2005)
MIROC3.2(hires)	$0.28 \times 0.19 \times 47$	109	Previous model	Hasumi and Emori (2004)
MIROC3.2(medres)	$1.4 \times 1.4 \times 43$	109	Previous model	Hasumi and Emori (2004)
UKMO HadCM3	$1.25 \times 1.25 \times 20$	300	Levitus (1995)	Gordon et al. (2000)
UKMO HadGEM1	$1 \times 1 \times 40$	85	Levitus et al. (1998)	Johns et al. (2004)
Ocean model	Resolution	Surface flux	Initial conditions	References
NEMO-OPA 2.0	$2 \times 2 \times 31$	CORE, ERA-40	Previous model	Madec (2008)
Reanalysis	Resolution	Surface flux	Adjustment	References
ECCO SIO 1	$1 \times 1 \times 23$	National Center for Atmospheric Research (NCAR)	Initialization, flux	Köhl et al. (2003)
SODA 1.4.2	$0.5 \times 0.5 \times 40$	ERA-40	Ocean properties	Carton et al. (2005)

facilitate restratification after convection. The deep currents are gravity flows originating from the overflows, which are shallow (650- and 850-m depth) and narrow enough to fall between model grid points. All of these processes are difficult to include correctly into models and have a relatively large influence on the (local) ocean state and variability.

The northwestern North Atlantic Ocean is an important area in constituting the meridional overturning circulation (MOC) and the climate of the North Atlantic region (Marshall et al. 2001; van Aken 2007). This makes it a particularly interesting area to assess model performance in simulating the local ocean state and its variability. This study aims to assess the ability of the present generation of CCMs used for the Intergovernmental Panel on Climate Change (IPCC) Fourth Assessment Report (AR4; Meehl et al. 2007; Solomon et al. 2007) to simulate the hydrography of the central Labrador and Irminger Seas, similar to the study of Sloyan and Kamenkovich (2007) of water masses in the Southern Ocean. A model–observation and model–model intercomparison will highlight model biases and indicate points on which models might be improved. For this purpose we use the preindustrial simulation of the CCMs. A comparison with the twentieth-century simulation has also been made, but we will show that these results do not significantly deviate from those of the preindustrial simulation. Furthermore, two ocean reanalyses [Estimating the Circulation and Climate of the Ocean (ECCO) and Simple Ocean Data Assimilation (SODA)] and an ocean model driven by observed fluxes [Nucleus for European Modelling of the Ocean–Océan

Paralléléisé (NEMO-OPA)] are compared to the observations and the coupled models to gain insight into the possible causes of simulation biases. The mean state as well as the variability of the potential temperature, salinity, potential density, and stratification is investigated in 3 layers: the surface layer (0–200 dbar), the intermediate Labrador Sea Water layer (750–1250 dbar), and the deep Northeast Atlantic Deep Water layer (2000–2500 dbar). Details of the analyzed observational and model time series are given in section 2. A short description of the observed hydrography and causes of its variability are contained in section 3. Section 4 deals with the analysis of the observed and simulated data and the use of error scores based on the analysis by Schmittner et al. (2005). Section 5 will begin with a general overview of the model results and continues with the differences between the preindustrial and twentieth-century simulation and the influence of the seasonal cycle on the assessment of model performance. This section continues with the results of the preindustrial simulation–observation comparison. Subsequently, the results of the model–model intercomparison are discussed in section 6. The final two sections, 7 and 8, comprise the discussion of possible causes of model biases and conclusions.

2. Data

a. Observational time series

The observational data used for this study originate from the hydrographic sections [the Atlantic Repeat line 7 (AR7)] through the Labrador and Irminger Seas. During

the World Ocean Circulation Experiment (WOCE) Hydrographic Program, from 1990 to 1997, these were surveyed annually in order to study whether the WOCE one-time survey was representative for the overall ocean hydrographic structure in the 1990s. The (near) annual survey of this line was continued after the termination of WOCE as a contribution to the Climate Variability and Predictability Program (part of the World Climate Research Programme). The western part of the AR7 section, the AR7W line between Labrador and Greenland, has been surveyed by the Bedford Institute of Oceanography, Nova Scotia (BIO). The resulting dataset along the AR7W line was extended backward to 1938 by I. Yashayaev of the BIO, using available historical hydrographic data. Near-annual observations are available from 1949 onwards. The eastern part of the AR7 section, the AR7E line from Greenland through the Irminger Sea and ultimately ending on the Irish continental shelf, has been surveyed near annually since 1990 until the present by the Institute for Sea Research, Hamburg (IfMH), and the Royal Netherlands Institute for Sea Research, Texel (NIOZ), with one survey in 1991 by the Southampton Oceanography Centre, presently the National Oceanography Centre, Southampton (NOC).

Hydrographic observations were selected from the boxes in the central Labrador and Irminger Seas, shown in Fig. 1. These boxes do not include the slope regions but only the deepest parts of the basin where dense overflow waters are found as well as the areas where deep convection has been observed or is expected (Marshall and Schott 1999; Pickart et al. 2003). The profiles within these boxes were isopycnally averaged per survey in order to generate a single profile per survey, representing the characteristic properties in the center of that basin. This is done by robust (isopycnal) averaging as described by Yashayaev (2007) using density intervals of 0.005 kg m^{-3} . From these mean survey profiles of temperature and salinity, potential temperature θ , potential density ρ relative to the surface, the potential density anomaly σ_0 , and Brunt–Väisälä frequency N were calculated. The resulting series of hydrographic property profiles were regridded at equal time intervals using a kriging technique. This prevented disproportional contributions of years with multiple hydrographic surveys. The resulting final equidistant time series contained 75 time steps for the Labrador Sea (from 58 surveys) and 25 time steps for the Irminger Sea (from 16 surveys). The timing of the original surveys is shown by the symbols along the top axis of the panels in Fig. 2. We have to note here that, since very few research cruises in the Irminger Sea and Labrador Sea were carried out in winter because of the adverse weather and sea ice conditions, these time series have a bias toward the

summer half year. The hydroographies of the Labrador and Irminger Seas and their temporal variability have been analyzed in detail by Yashayaev (2007), Hendry et al. (2007), and Yashayaev et al. (2007). A description of the observed hydrography follows in section 3.

b. Simulated time series

To compare the hydrographic observations with CCM model simulations, 20 yr of monthly temperature and salinity data were obtained from the preindustrial simulations (Picntrl runs) from each of the eight models (Table 1). These data are available from the World Climate Research Program (WCRP) Coupled Model Intercomparison Project 3 multimodel dataset (CMIP3; Meehl et al. 2007). These model simulations are part of the extensive model dataset used for the IPCC AR4 report (Solomon et al. 2007) and represent the state of the art in climate modeling. The selection of the models was based on the availability of ocean volume data and the absence of flux adjustment in the models. The models that exhibited either a too strong [$> 30 \text{ Sv}$; ($1 \text{ Sv} \equiv 10^6 \text{ m}^3 \text{ s}^{-1}$)] or too weak ($< 10 \text{ Sv}$) Atlantic MOC in the analysis of Schmittner et al. (2005) were not included in the selection. For the preindustrial simulations these models use modern-day topography and ice sheet coverage, a constant solar constant of 1365 W m^{-2} , and greenhouse gas concentrations at the low levels from before the industrial revolution. Most of the models are initialized from an ocean at rest, with the temperature and salinity distribution set at values from one of the Levitus datasets (Levitus and Boyer 1994; Levitus et al. 1994, 1998). Three of the models (Table 1) are initialized with the final ocean state of a previous version of the model. The models are spun up for 70–330 yr to ensure that remaining trends are small. The results from these preindustrial control runs are assumed to give a good assessment of internal model differences in the ocean mean hydrographic state and variability without trends caused by anthropogenic climate forcing.

To guarantee that the model biases found in the preindustrial simulations are not the results of the difference between preindustrial CO_2 forcing and current climate, a comparison with the twentieth-century control simulations (20C3M runs) has also been made. Most of these 20C3M runs are initialized with the results of the preindustrial control simulation, thus possible biases in the preindustrial simulations have an effect on the 20C3M results. The applied solar forcing in the 20C3M control runs is based on a reconstructed time series or a constant value, both within the range $1365\text{--}1370 \text{ W m}^{-2}$. CO_2 and aerosol concentrations are based on annual

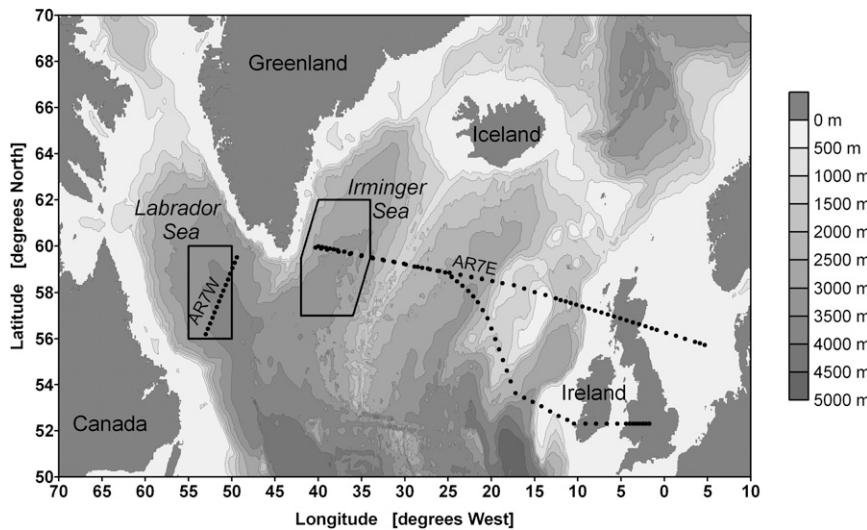


FIG. 1. Topography of the northern North Atlantic Ocean with contour lines at 500-m intervals. The black dots represent examples of locations of hydrographic stations. The areas in the Labrador and Irminger Seas in which the data are compared are enclosed by the thick black lines.

observations and no volcanic variability. The 20C3M simulations generally run from 1850 to 2000. For this study the last 20 yr (1980–2000), which coincide with the observations, were analyzed.

A simulation originating from an ocean- and sea ice-only model driven by realistic surface forcing was added to the dataset to determine whether the coupling to other models (most important the atmospheric and sea ice components of the CCM) has a significant effect on the performance of the ocean model. The ocean model that was chosen is the NEMO–OPA model version 2.0 (Madec 2008; <http://www.locean-ipsl.upmc.fr/NEMO/>) from the ORCA group of ocean models. This model is very similar to the ocean component included in some of the CCMs [ORCA in L’Institut Pierre-Simon Laplace Coupled Model, version 4 (IPSL CM4) and OPA 8.1 in Centre National de Recherches Météorologiques Coupled Global Climate Model, version 3 (CNRM-CM3)]. The NEMO–OPA simulation uses the downward shortwave and longwave radiation forcing from the Coordinated Ocean and Sea Ice Reference Experiment (CORE), also used in the experiments of the DRAKKAR group. For the precipitation a modified CORE field is used in which precipitation is reduced by 15%–20% northward of 30°N. The turbulent fluxes are calculated from temperature, humidity, and wind fields from 40-yr European Centre for Medium-Range Weather Forecasts (ECMWF) Re-Analysis (ERA-40) combined with sea surface temperature generated by the ocean model. No restoring surface conditions are used. The NEMO–OPA ocean model includes the Louvain-la-Neuve Sea Ice Model [LIM; also LIM in

IPSL CM4 and Global Experimental Leads and Sea Ice for Atmosphere and Ocean (GELATO) 2 in CNRM-CM3]. The ocean model has been run twice from the beginning of 1958 to the end of 2001. The second run was initialized on the final conditions of the first run. The 20 simulated years between 1980 and 2000 of the second run were used for the comparison.

Finally, data from two ocean reanalyses were also analyzed. These reanalysis data were obtained from the ECCO project [Scripps Institute of Oceanography (SIO) version 1, from <http://www.ecco-group.org>] and the SODA project (version 1.4.2, downloaded from http://apdrc.soest.hawaii.edu/w_data/ocean3.htm) and consist of 11 and 20 yr of monthly data, respectively (Table 1). The reanalysis data represents the “best possible” fit of a CCM to basin-scale observations, meaning that one should not expect the eight CCMs or the ocean model to perform better than the reanalysis. The ECCO and SODA ocean reanalyses both use atmospheric reanalyses as surface forcing for the ocean general circulation model, albeit from different origins (Table 1). Both reanalyses do not include sea ice models. The simulated ocean properties are adjusted to match observed values from hydrographic and satellite datasets as closely as possible. The ECCO model uses an iteration of model runs, adjusting the initial state and surface forcing after every model run, until a best fit to the observations is obtained (Köhl et al. 2003). The SODA model adjusts the simulated ocean properties to the observed values during the model run, resulting in a best fit after a single model run (Carton et al. 2005).

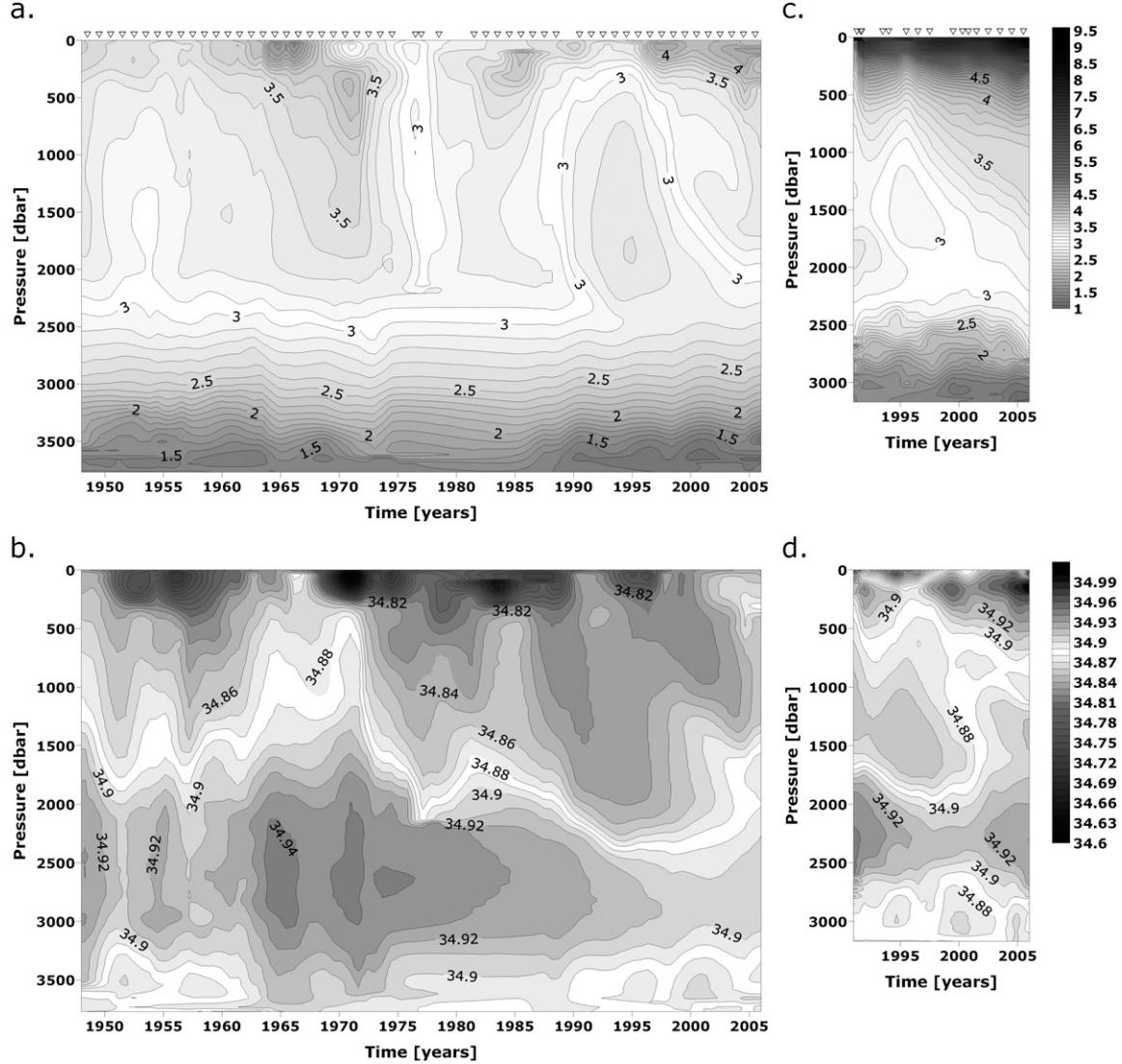


FIG. 2. Observed changes over time in the hydrography of (a), (b) the Labrador Sea and (c), (d) the Irminger Sea. The displayed hydrographic variables are potential temperature in (a) and (c) and salinity in (b) and (d). The timing of the individual surveys is indicated by the triangles along the top axis of (a) and (c).

3. The observed hydrography

This section serves as a short summary of the various water masses observed in the Labrador and Irminger Seas, their character in terms of salinity and temperature and main causes of variability. For a more thorough description of the local water masses we refer to the selection of papers referenced in this section. The observed hydrographic series are illustrated by time-pressure sections of salinity and potential temperature showing the hydrographic structure and variability over the last few decades (Fig. 2). The surface waters in the Labrador Sea originate from the West Greenland Current, the Labrador Current, and Davis Strait Through-

flow. These relatively cold and fresh surface waters are brought to the center of the basin by baroclinic eddies (Lilly et al. 2003; Katsman et al. 2004). Relatively cold and fresh surface waters at the western border of the Irminger Sea originate from the East Greenland Current (Holliday et al. 2007) and only rarely reach the center of this basin. The variability in the surface waters is caused by air-sea interaction, meltwater from sea ice, and changes in advective patterns.

Between the surface water and the base of the winter mixed layer resides the Subpolar Mode Water (SPMW; McCartney and Talley 1982; Brambilla and Talley 2008). The SPMW is formed by air-sea interaction in branches that originate in the North Atlantic Current and thus has

a relatively warm and saline character (Brambilla et al. 2008). Presumably the SPMW is modified in the Irminger Sea by convection (Pickart et al. 2003; Bacon et al. 2003; Pickart et al. 2008). The Irminger Current transports the SPMW around the southern tip of Greenland to the Labrador Sea.

Intense surface cooling over the Labrador Sea initiates deep convection, which reaches down to between 500 and 2400 m depending on preconditioning and surface forcing (Marshall and Schott 1999; Lazier et al. 2002; Haine et al. 2008; Yashayaev et al. 2008). During periods of strong convection, like the early 1990s, the Labrador Sea Water (LSW) forms a large volume of homogeneous water, which is relatively cold and fresh because of the incorporation of fresh, strongly cooled surface waters (Yashayaev et al. 2007). The volume of LSW is drained by advection southward in the upper layer of the deep western boundary current and eastward to the Irminger Sea (which takes 1–2 yr), the Iceland Basin, and Rockall Trough.

Between 2000 and 3200 dbar a high-salinity core with well-developed temperature stratification is observed in the Irminger Sea and the Labrador Sea. This water mass is the Northeast Atlantic Deep Water (NEADW). Its salinity is influenced by varying ratios of entrainment of relatively fresh LSW and relatively saline Icelandic Slope Water (van Aken and de Boer 1995; Yashayaev et al. 2007).

Denmark Strait Overflow Water (DSOW) can be found below the NEADW. It is cold because of its origin in the Greenland Sea (Tanhua et al. 2005) and relatively fresh with respect to the overlying NEADW. Variability in DSOW is caused by variability in the overflow and entrainment of SPMW and NEADW.

4. Analysis and means of comparison

The salinity and temperature volume data generated by the CCMs, the ocean-only model, and the reanalyses were processed as follows. Simulated θ and S profiles at all time steps and grid points were interpolated vertically at 10-dbar intervals between the nonequidistant simulated depths. The derived hydrographic variables, the potential density σ_0 , and the Brunt–Väisälä frequency squared (proportional to the density gradient), were calculated from the interpolated θ and S profiles. Three layer averages, for the surface layer (0–200 dbar), the LSW layer (750–1250 dbar), and the NEADW layer (2000–2500 dbar), were obtained by averaging horizontally between grid points and vertically between the layer boundaries. A linear fit was used to detrend the layer averages, thus removing model drift and long-term (> 20 yr) variations. The resulting layer averages were used to compare the 20-yr mean and variability of the

simulations to the observations. Time- and area-averaged vertical profiles of hydrographic properties were also obtained to gain insight into the vertical structure of the simulated oceans and the depth integrated properties.

Layer averages from the observational time series were constructed similarly. These layer averages and standard deviations of the observational time series in the Irminger and Labrador Seas are given in Table 2. Although the salinity is based on a ratio of conductivity and is therefore dimensionless, we employ the practical salinity unit (psu) to consistently report biases in hydrographic variable with their units. Note the small standard deviations in this table. It shows that the quantitative range of variability is small compared to differences between the various water masses and similar water masses in the two basins.

As mentioned before, only a few of the survey cruises were carried out in winter and spring. Therefore the observations tend to have a bias toward the late summer part of the year. For better comparison with these observations, two simulated time series with a summer bias were constructed. The first summer series consists only of the simulations for the months August–October, in which most survey cruises took place and sea surface temperature (SST) is maximal. This series will give an indication of the magnitude of the seasonal cycle. The second, more inhomogeneous summer series was constructed by randomly subsampling from each year in the original 20-yr series. The timing of these subsamples is chosen by a random generator, such that it has a normal distribution around the late summer months. This last series is expected to resemble the variability of the inhomogeneously distributed observations more closely than the annual and August–October series.

The original θ and S profiles at their simulated depths were used to calculate σ_0 profiles for all time steps and grid points. From these profiles the mixed layer depth was derived using a delta criterion (Donners et al. 2005) of $\Delta\sigma_0 = 0.125 \text{ kg m}^{-3}$. The base of the mixed layer is defined as the depth at where the potential density difference with respect to the surface exceeds $\Delta\sigma_0$. The maximum mixed layer depth per winter is derived from the grid point with the deepest mixed layer. To reduce influences of differences in topography parameterization and horizontal resolution a fraction of convective volume was also determined. This is based on the fraction of available water column taking part in the mixed layer per grid point and the fraction of surface area within the analyzed box that the grid point represents. Summed over all grid points this gives a percentage of the volume within the analyzed box that is taking part in the mixed layer. Since the convective formation of LSW is a very important process for the

TABLE 2. Observed hydrographic properties in the three layers in the Labrador and Irminger Seas. The denoted values are the mean and standard deviation of the potential temperature (θ), salinity (S), potential density anomaly (σ_0), and the stratification or Brunt-Väisälä frequency squared (N^2).

Labrador Sea	θ (°C)	S (psu)	σ_0 (kg m $^{-3}$)	N^2 (s $^{-2}$)
0–200 dbar	3.59 ± 0.31	34.80 ± 0.05	27.67 ± 0.03	2.4·10 $^{-6}$ ± 2.2·10 $^{-6}$
750–1250 dbar	3.20 ± 0.17	34.86 ± 0.01	27.75 ± 0.01	4.4·10 $^{-7}$ ± 1.2·10 $^{-7}$
2000–2500 dbar	3.06 ± 0.11	34.91 ± 0.02	27.80 ± 0.01	1.0·10 $^{-6}$ ± 1.8·10 $^{-7}$
Irminger Sea	θ (°C)	S (psu)	σ_0 (kg m $^{-3}$)	N^2 (s $^{-2}$)
0–200 dbar	6.09 ± 0.34	34.91 ± 0.02	27.46 ± 0.04	3.0·10 $^{-5}$ ± 6.6·10 $^{-6}$
750–1250 dbar	3.35 ± 0.09	34.88 ± 0.005	27.77 ± 0.01	7.3·10 $^{-7}$ ± 1.2·10 $^{-7}$
2000–2500 dbar	2.99 ± 0.05	34.92 ± 0.01	27.84 ± 0.005	1.5·10 $^{-6}$ ± 2.8·10 $^{-7}$

local hydrography, differences in simulated mixed layer depth may explain biases found at the level of LSW.

To quantify the performance of the models we have assigned error scores based on the skill scores proposed by Schmittner et al. (2005). The error scores are determined for each hydrographic variable per layer per basin. They are obtained as follows: The simulation bias, or the difference between the simulated and observed mean, is normalized by dividing by the standard deviation of the observations σ_{obs} to obtain a positive normalized root-mean-square error of the simulation $\text{nrms}_{\text{mean}}$:

$$\text{nrms}_{\text{mean}} = \sqrt{\frac{\text{bias}_{\text{sim}}^2}{\sigma_{\text{obs}}^2}}.$$

With these normalized root-mean-square errors of the model performance we define an overall error score S_k as the overall averaged $\text{nrms}_{\text{mean}}$. Thus, a high error score indicates weak model performance.

5. Results of CCM–observation comparison

a. General results

Water mass analysis (as applied in section 3) is often based on combinations of temperature and salinity minima or maxima seen in a temperature–salinity plot (θ – S plot). The observed mean θ – S profile in the Irminger Sea (Fig. 3) shows the warm and saline Subpolar Mode Water, two salinity minima for the Labrador Sea Water density classes of 1994 and 2000, the saline Northeast Atlantic Water, and the fresher, cold Denmark Strait Overflow Water. However, both panels of Fig. 3 show that a description of the local simulated hydrography by using the traditional water masses is not suitable. The SODA reanalysis shows a similar θ – S profile with two salinity maxima (resembling SPMW and NEADW) and a salinity minimum (LSW), albeit at higher temperatures, but gives no indication of the

presence of cold DSOW. These water masses are not so easily recognized in the θ – S profiles of the ECCO reanalysis and NEMO–OPA ocean model. The dissimilarity between the observed θ – S profiles and the θ – S profiles simulated by the coupled climate models (right panel of Fig. 3) makes the comparison in terms of water masses even more difficult. We can, however, make a comparison in terms of vertical profiles of hydrographic properties.

The time-mean vertical profiles of salinity in the Labrador Sea (Fig. 4) illustrate the overall tendency of the coupled models to have a very low surface salinity and a high salinity below 500 dbar. However, the relatively high depth-average salinity, indicative of salt content (Fig. 4), shows that the low surface salinity does not compensate the saline deeper layers in most of the models. Also, the much larger-than-observed salinity gradient over the upper 1000 dbar strongly increases the stability of this layer in the models.

Most of the Labrador Sea potential temperature profiles (Fig. 4) provide a negative contribution to the stability of the upper 1000 dbar by having cold surface water over a warm deep layer. In the next section we will show that the stabilizing effect of the salinity is much larger than the destabilizing contribution of the potential temperature profiles. The depth-average potential temperature in the preindustrial runs of the coupled models is between 0.6° and 2.5°C higher than observed, indicating that the local heat content is very high in these simulations.

b. Preindustrial run versus twentieth-century control run

Both the Picntrl runs and 20C3M runs show a very similar tendency of the depth-averaged salinity and potential temperature in the Labrador Sea (Fig. 4). All of the coupled models have a too high depth-averaged potential temperature in the Picntrl runs. Three of the coupled models shows decrease in potential temperature

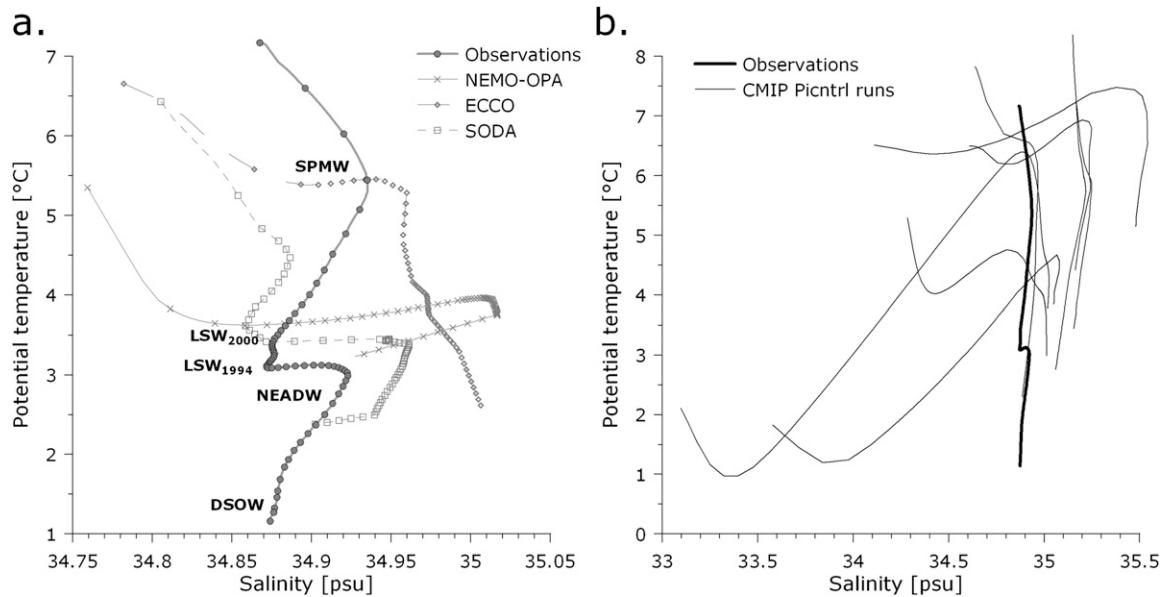


FIG. 3. Potential temperature–salinity plots of the mean profiles in the Irminger Sea. (a) θ -S profiles of the observations, ocean model, and ocean reanalyses. The local water masses, SPMW, two classes of LSW formed in the years 1994 and 2000, NEADW, and DSOW are indicated by their acronyms. (b) θ -S profiles of all preindustrial runs (thin lines) and the observed profile (thick lines). Note the difference in scale, especially for the salinity axis.

over a large part of the deeper water column in the 20C3M run, with one of them [Geophysical Fluid Dynamics Laboratory Climate Model version 2.0 (GFDL CM2.0)] simulating lower-than-observed temperatures within a large part of the deeper water column. Five of the models show an increase in the depth-average potential temperature with respect to the preindustrial run, thus also increasing the bias with respect to the observations.

The salinity profiles are very similar in both runs, simulating relative freshwater over a deep saline water column. Four of the models show a decrease in depth-averaged salinity in the 20C3M run compared to the Picntral run, while the other four show an increase in salinity. The largest increase in depth-averaged salinity (0.24 psu) is exhibited by the GFDL CM2.0 model. This coupled model had a very fresh mean surface salinity in the preindustrial run (32.42 psu against 34.80 psu observed), but it has a mean surface salinity of 34.50 psu in the twentieth-century run. Despite this improvement in the surface layer, the salinity profile below 250 dbar, which was already too saline in the Picntral run, approximately doubled its bias in the 20C3M run.

In the remaining part of the discussion of the results we will focus mostly on the preindustrial runs. Since most of the twentieth-century runs are initialized on the ocean state of the preindustrial control runs, biases in the 20C3M run will include the results of existing biases in the initialization. This is contrary to the preindustrial

runs, which are mostly initialized with an ocean state very similar to our observations, namely one of the Levitus datasets. Therefore, most insight into the original simulation biases and possible causes are more likely to be gained from the comparison between the preindustrial runs and the observations.

c. Effects of the seasonal cycle and inhomogeneous sampling

The difference between the annual layer averages and the late summer layer averages is most pronounced in the surface layer. Because of the local atmospheric seasonal cycle, especially solar radiation, there is an increase in temperature in spring and summer. The excess in precipitation decreases salinity in summer. The effect of summer heating and freshening is most pronounced in the August–October summer series. The temperature increases over the upper 200 dbar of the Labrador and Irminger Seas are both about 0.7°C (not shown) with respect to the annual series. The inhomogeneous summer series shows a maximum increase of 0.4°C when the sample distribution is centered on September. Both increases in temperature are confined to the upper layer by the stratification; differences between the annual and summer mean of the intermediate and deep layer are $< 0.01^{\circ}\text{C}$ for both summer series and are not significant. The salinity of the surface layer generally decreases by about 0.05 psu for the August–October

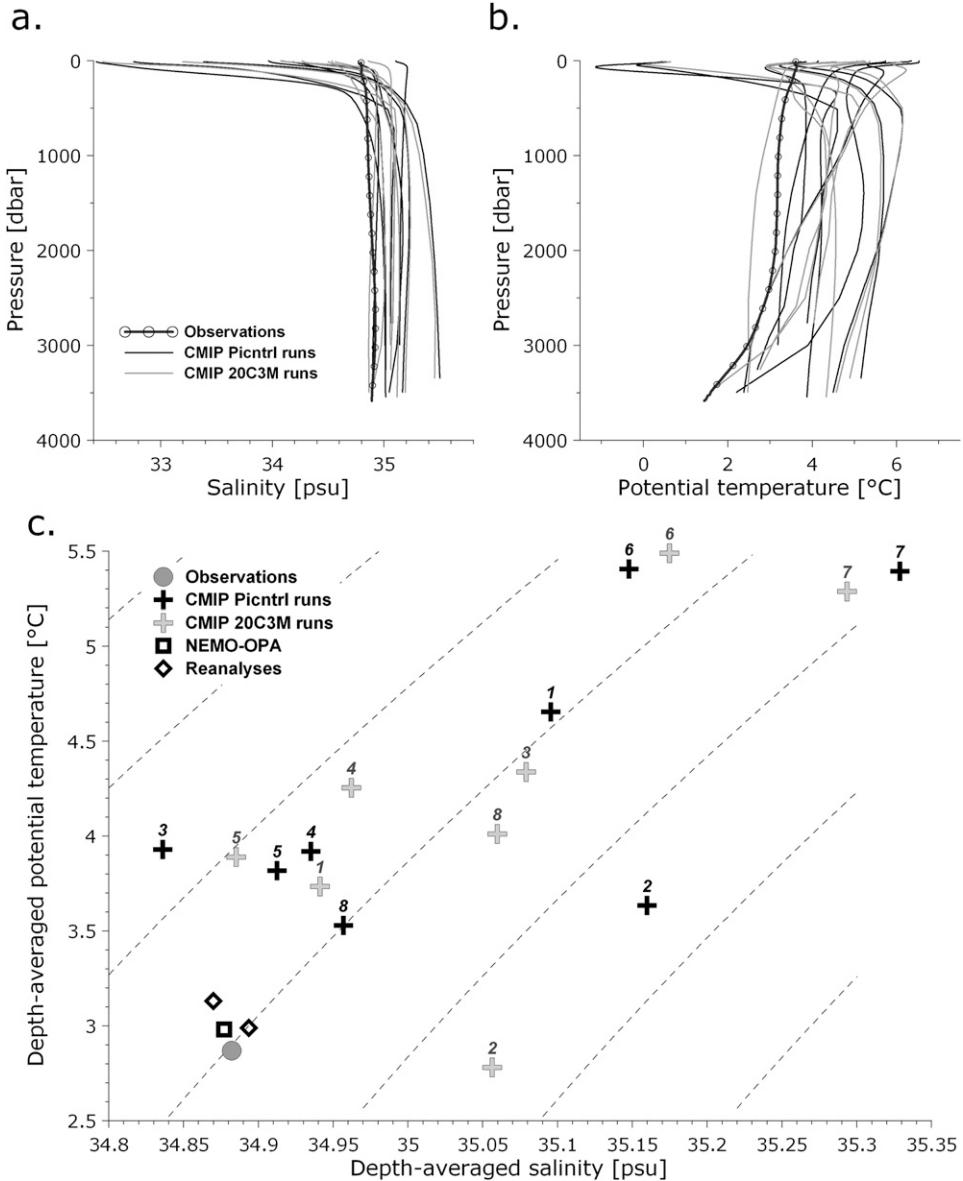


FIG. 4. Vertical profiles of (a) salinity and (b) potential temperature from the observations and simulations in the Labrador Sea. The preindustrial runs are drawn in black; the twentieth-century runs are drawn in gray. (c) The depth-averaged potential temperature and salinity for the observations, models, and reanalyses in the Labrador Sea. The numbers next to the symbols correspond to the model number in Figs. 5–8. Dashed lines of equal density anomaly are drawn at 0.1 kg m⁻³ intervals. The observations are approximately on the 27.8 kg m⁻³ isopycnal.

series and 0.03 psu for the inhomogeneous summer series. Salinity changes in the intermediate layer are about 0.001 psu and are also not significant. Although most of the observations were made in summer, consideration of the simulated summer months does not improve model performance. On the contrary, the increase in temperature and decrease in salinity strengthens the biases seen in the annual averages.

d. Mean and variability of 20 yr of CCM simulations per layer

The 20-yr mean values of the preindustrial coupled model simulations for the surface, intermediate, and deep layer are graphically presented in Figs. 5–8 for all discussed hydrographic properties. The result for the Labrador and Irminger Seas are discussed per layer.

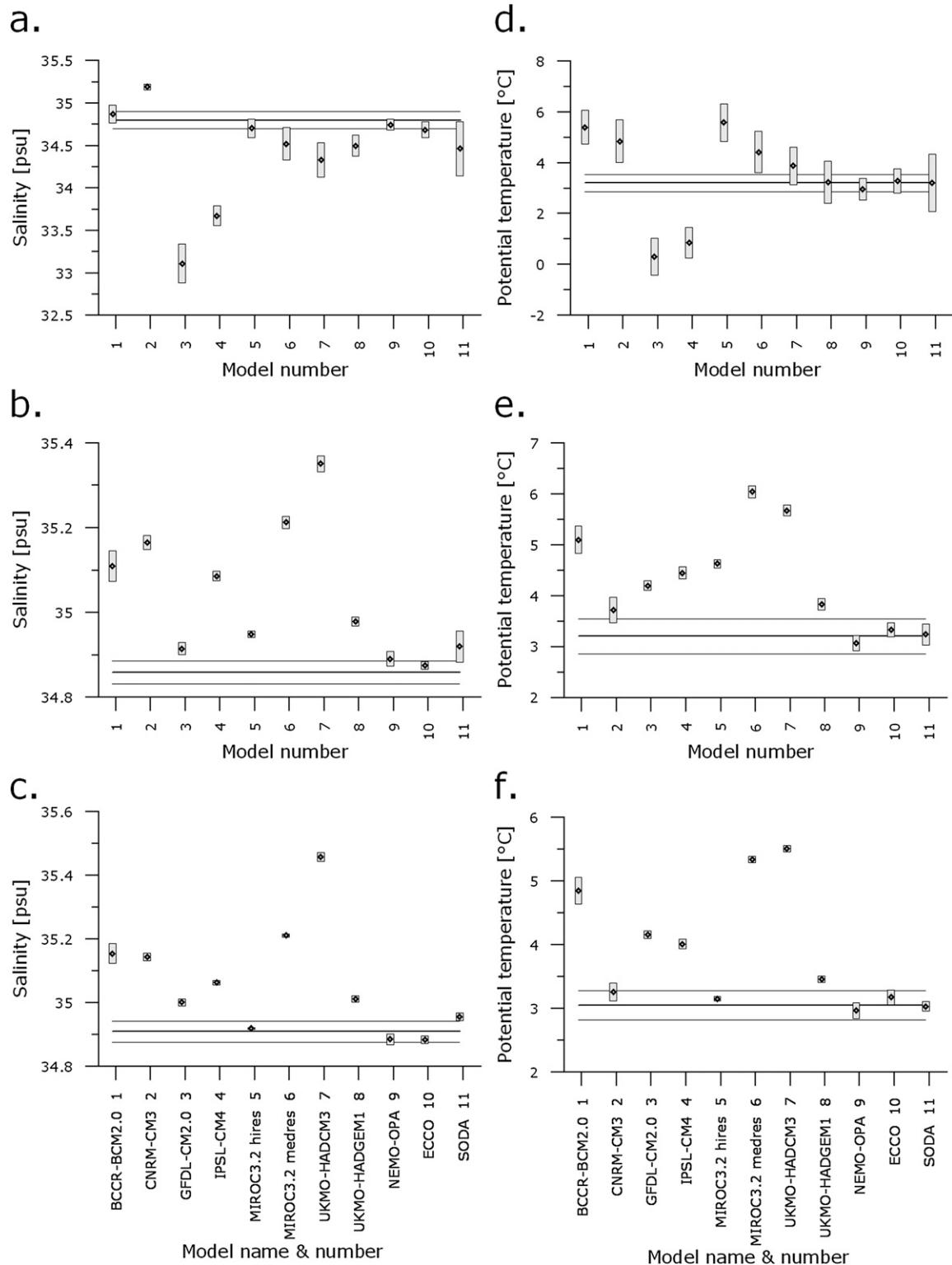


FIG. 5. (a)–(c) Salinity and (d)–(f) potential temperature in the Labrador Sea. Layer averages are shown for the surface in (a) and (d), the intermediate layer in (b) and (e), and the deep layer in (c) and (f). Black dots represent the means of the simulations. The rectangles are the means \pm one std dev. The horizontal lines indicate the observed layer mean (black) \pm two std dev (gray).

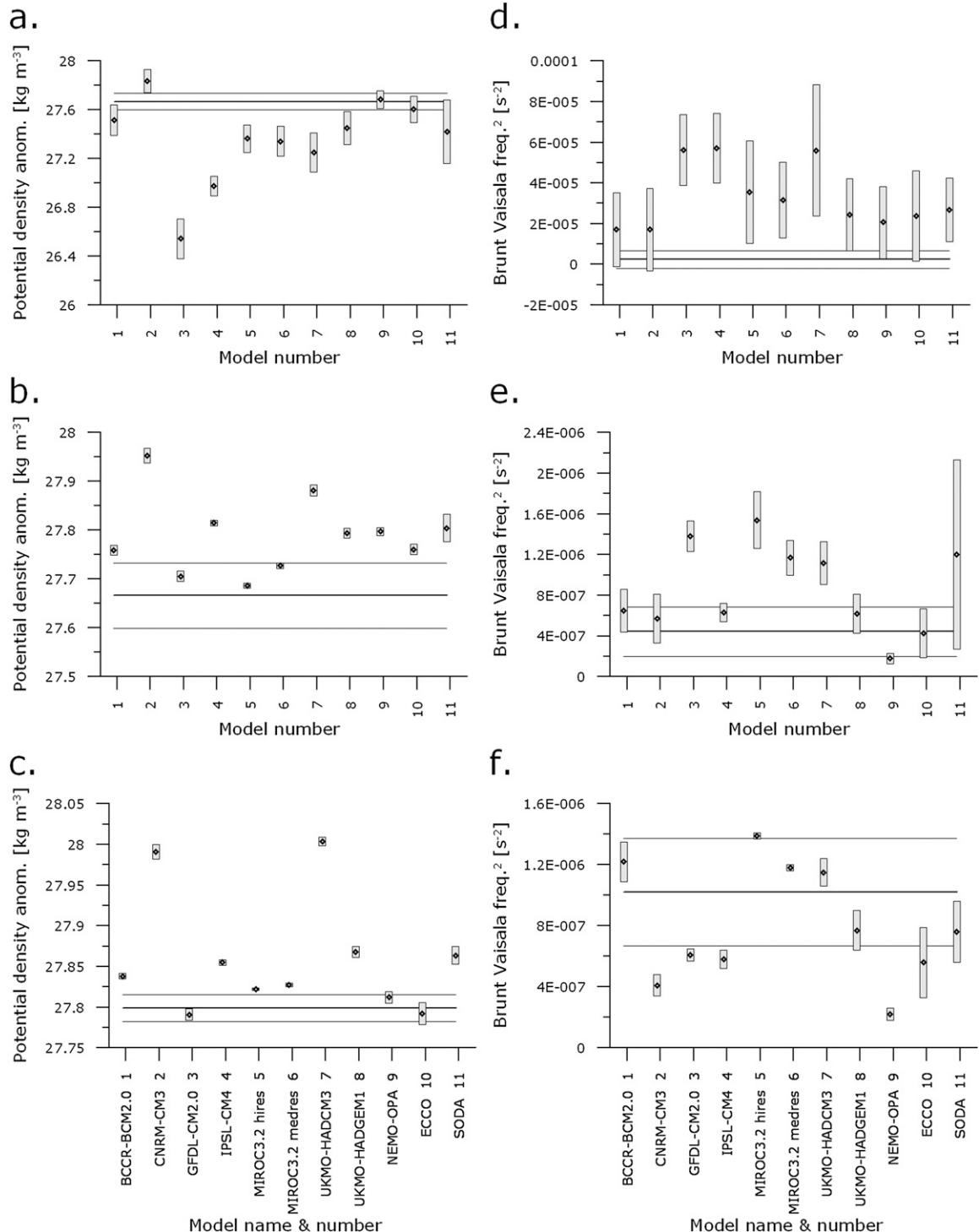


FIG. 6. (a)–(c) Potential density anomaly and (d)–(f) the Brunt–Väisälä frequency squared in the Labrador Sea. Layer averages are shown for the surface in (a) and (d), the intermediate layer in (b) and (e), and the deep layer in (c) and (f). Black dots represent the means of the simulations. The rectangles are the means \pm one std dev. The horizontal lines indicate the observed layer mean (black) \pm two std dev (gray).

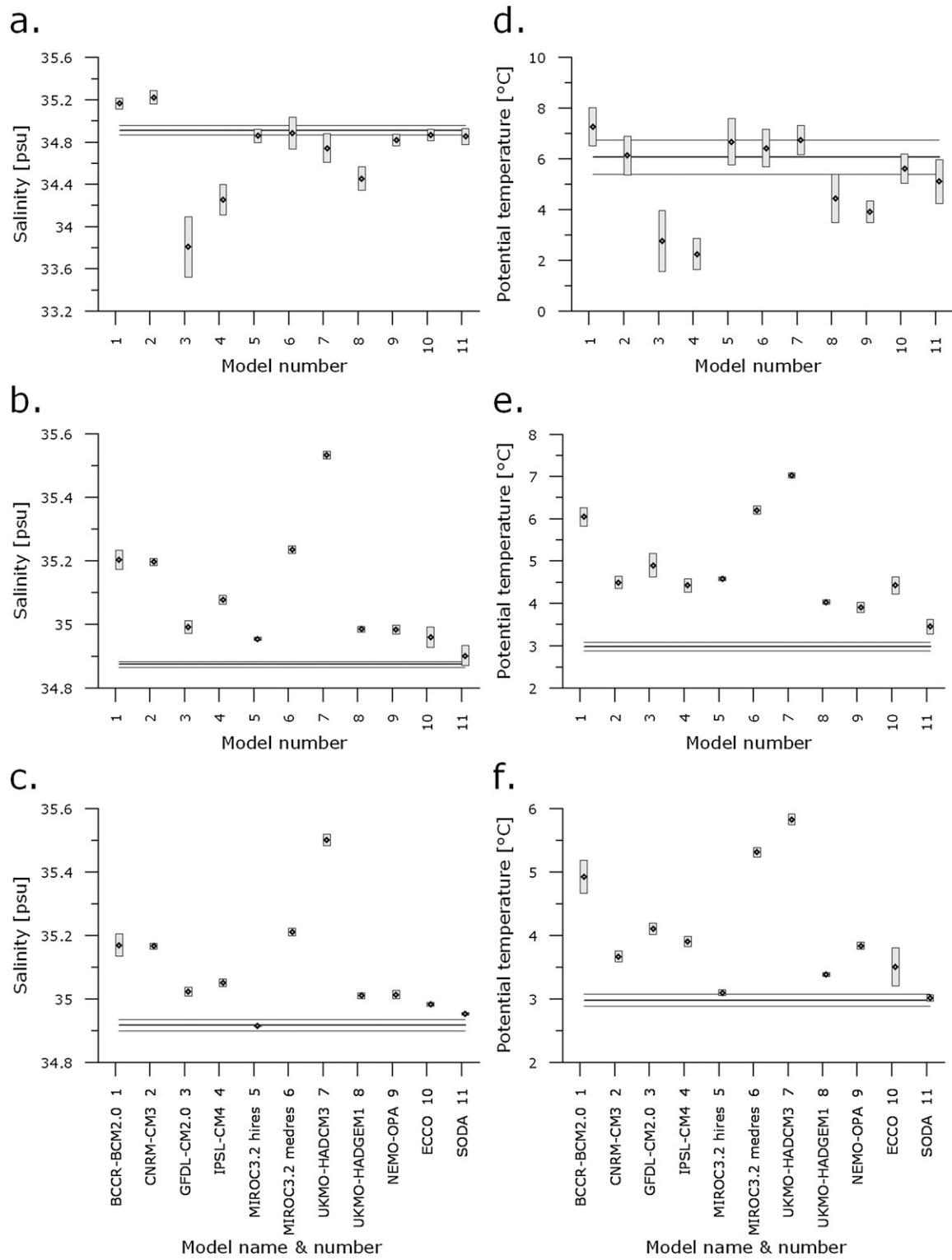


FIG. 7. As in Fig. 5, but for the Irminger Sea.

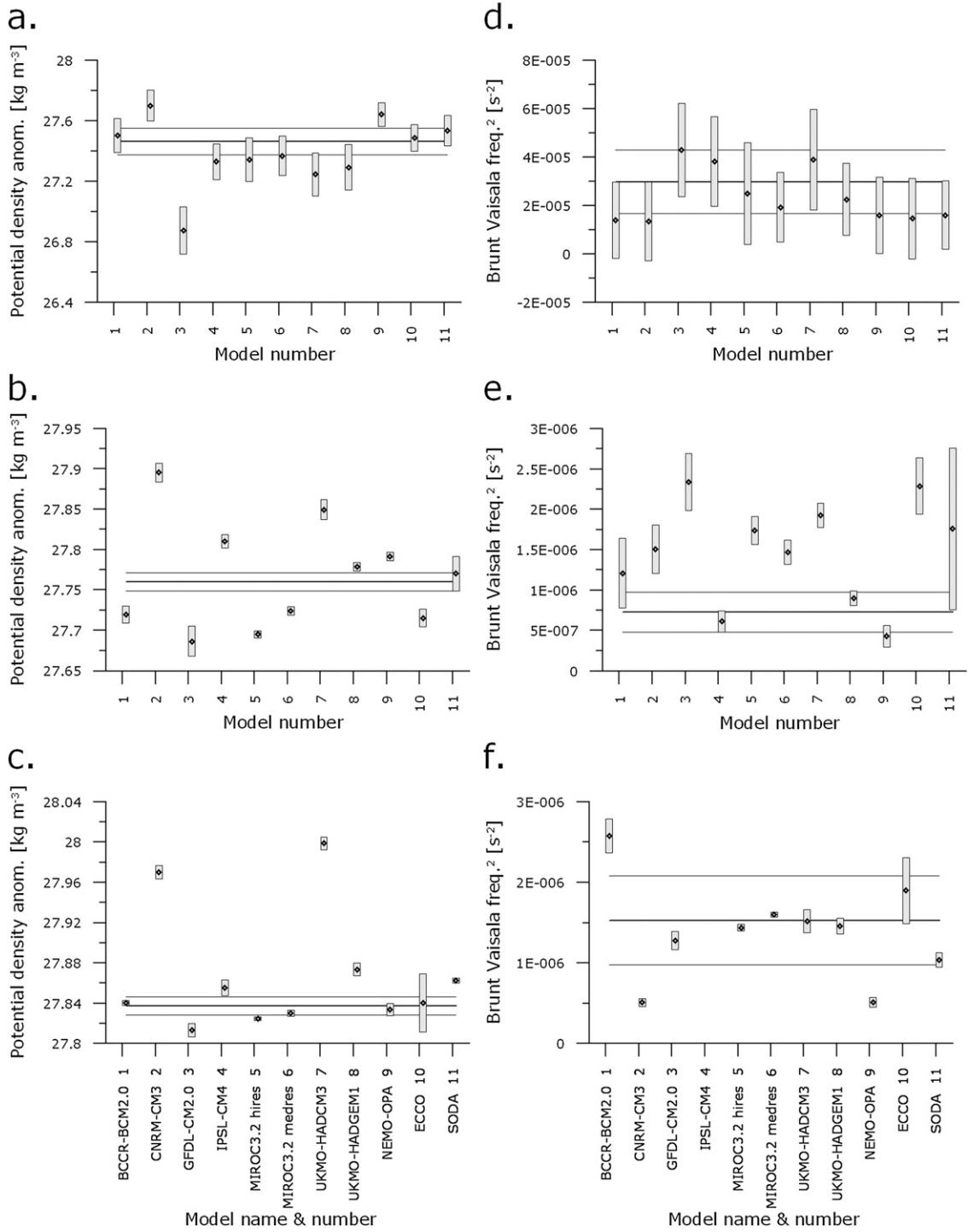


FIG. 8. As in Fig. 6, but for the Irminger Sea.

1) THE SURFACE LAYER

All four figures reveal similar model errors for the Labrador Sea and the Irminger Sea, typical for basins connected by advection pathways (Yashayaev et al.

2007). Thus, simulations with too low (or high) values for the mean hydrographic parameters in the Labrador Sea generally have also too low (or high) values in the Irminger Sea. Two models, GFDL CM2.0 and IPSL CM4, have extremely low salinities and potential temperat-

in the surface layer, between 0.7 and 1.7 psu below the mean observed salinities, and between 2.3° and 3.8°C below the observed temperatures. These model biases are orders of magnitude larger than the observed range of variability (Table 2). The salinity and temperature biases do not compensate in density. The simulated surface density is generally much lower than observed, especially in the Labrador Sea. Overall the surface density is too low for all but one CCM simulation, related to the lower-than-observed surface salinities. The low surface density causes a higher simulated stability or density gradient over the surface layer in the Labrador Sea, expressed by the Brunt–Väisälä frequency squared or N^2 . In the Irminger Sea, the stability of the surface layer in the model simulations is close to the observations.

2) THE INTERMEDIATE LAYER

The intermediate layer from 750 to 1250 dbar is representative for the LSW in the Labrador and Irminger Seas. In most model simulations this layer is much too warm and saline, with biases up to 2.8°C and 0.5 psu in the Labrador Sea and 4.0°C and 0.7 psu in the Irminger Sea. This suggests that in most models the formation of LSW in winter is not well represented. Both the salinity and temperature difference between the surface and intermediate layer and between intermediate properties in the Labrador and Irminger Seas is much larger than observed. The very high intermediate temperatures in the Model for Interdisciplinary Research on Climate 3.2, medium-resolution version [MIROC3.2(medres)] model and the Met Office (UKMO) Third Hadley Centre Coupled Ocean–Atmosphere GCM (HadCM3) models suggest that in these models the intermediate layers have the properties of SPMW, rather than LSW. Two models from the same institutes, but with higher resolution—MIROC3.2(hires) and UKMO Hadley Centre Global Environmental Model version 1 (HadGEM1)—perform better. In general, the model errors in temperature and salinity of the intermediate layer seem to compensate in density, leading to a mixed behavior of the intermediate density, resulting in 65% of the models with a too high and 35% with a too low intermediate density. Similar to the surface layer, the stability of the intermediate part of the water column is systematically too high, an indication of low convective activity, with larger errors in the Labrador Sea than in the Irminger Sea.

3) THE DEEP LAYER

The model results from the deep NEADW layer between 2000 and 2500 dbar show mixed results. While a

few CCM simulations show realistic NEADW temperatures, most CCMs generate relatively high deep temperatures, up to 2.7°C too high. Apart from the MIROC3.2(hires), the CCMs have a higher-than-observed salinity in the NEADW layer. Also for this layer, temperature and salinity errors compensate in density for several models, but not for all. Especially the CNRM-CM3 and UKMO HadCM3 models have considerably higher NEADW densities, both in the Labrador Sea and in the Irminger Sea. The relative error in N^2 in the NEADW layer is small compared with the shallower layers.

The quantitative results of the analysis are compiled in Table 3, which contains the “error score” per variable summed over the three layers. In the Irminger Sea the largest contribution (50%–60% averaged over all models) to the error score is by the intermediate layer. This layer represents the core of the water formed in the Labrador Sea. The surface layer contributes 10%–20% and the deep layer 20%–30%. In the Labrador Sea the contribution of the layers to the final error score is not the same for all variables. Nearly 70% of the error score of N^2 is contained in the surface layer, because of the high simulated stability. The other 30% is distributed over the intermediate (20%) and deep layer (10%). Temperature skills scores are distributed equally over the three layers. Salinity and potential density have their largest error scores in respectively the intermediate (40%) and deep layer (45%).

4) VARIABILITY

Figure 9 shows the relative magnitude of the variability in the simulated hydrographic properties in the Labrador Sea. The lengths of the vertical bars represent the standard deviation of the inhomogeneous summer series, normalized by the standard deviation of the observations. The standard deviation of the inhomogeneous series is independent of the summer month on which this series normal distribution is centered. Its magnitude over all layers is about 97% of the standard deviation of the annual series for all hydrographic variables. The largest decrease in variability is in the surface layer; the standard deviation of the inhomogeneous series is approximately 90% of the standard deviation of the annual series. Since the variability in the intermediate and deep layer is mostly interannual, the difference in variability in the two time series is small at depth.

The potential temperature, salinity, potential density anomaly, and stratification exhibit a similar behavior in the three layers (Fig. 9). The variability in the surface layer is too large for nearly all models. The models exhibit

TABLE 3. Normalized RMSEs, or the simulation bias normalized by the standard deviation of the observations, for the simulated 20-yr mean hydrography in the Labrador Sea and the Irminger Sea. The error values for each variable are the means over the three analyzed layers. Total error scores (S_k) over all variables and both basins are displayed in the rightmost column.

Model	Labrador Sea				Irminger Sea				S_k
	S	θ	σ_0	N^2	S	θ	σ_0	N^2	
BCCR-BCM2.0	12	13	4	3	38	36	3	3	14
CNRM-CM3	15	5	12	4	38	15	20	4	14
GFDL CM2.0	14	11	12	11	29	24	11	5	15
IPSL CM4	16	10	11	10	30	20	5	3	13
MIROC3.2(hires)	3	8	4	9	7	12	6	3	6
MIROC3.2(medres)	17	15	5	7	38	39	3	3	16
UKMO HadCM3	26	13	15	10	73	48	19	4	26
UKMO HadGEM1	7	3	6	4	18	12	5	1	7
NEMO-OPA 2.0	2	1	2	5	13	14	4	3	5
ECCO SIO 1	2	1	2	4	9	14	3	5	5
SODA 1.4.2	5	0	6	6	4	4	3	4	4

a mixed behavior for the variability in the intermediate layer, with temperature variations mostly underestimated. In the deepest layer the coupled models show too low variability for all hydrographic variables. The combination of too high surface variability and low intermediate and deep variability suggests that the direct ocean response to the local atmospheric variability is confined to a relatively thin upper layer because of the too strong stratification. Consequently, the variability in the deeper layers is underestimated.

e. Mean and variability of 200 yr of CCM simulations

Two 200-yr model simulations were used in order to investigate the influence of longer-term variability, such as the variability caused by decadal oscillations of the North Atlantic Oscillation (Hurrell 1995), on the 20-yr mean values. The two models selected for this analysis are CNRM-CM3 and IPSL CM4, which both include an ocean model based on NEMO (section 1) and run at the same resolution. These ocean models are coupled to different atmosphere and sea ice models. The two models exhibited quite different results in the simulation of the Labrador and Irminger Seas (Figs. 5–8). The variability in the 200-yr simulated time series is compared to the observational dataset from the Labrador Sea. The processing of the 200-yr datasets is identical to the processing of the 20-yr datasets, which was explained before. The mean and standard deviation of the 20- and 200-yr time series of the hydrographic variables are very similar (less than 10% difference). Also, the variability in the time series is mostly seasonal and interannual to decadal, not multidecadal, thus explaining the relatively small difference between the standard deviations of the 20- and 200-yr series. The simulated hydrographic properties do not improve significantly by extending the dataset; thus we conclude that the weak performance of the IPSL CM4 model is not due to an unlucky selection of the relatively short segment of simulated data. It can be expected that this is also true for the other model simulations.

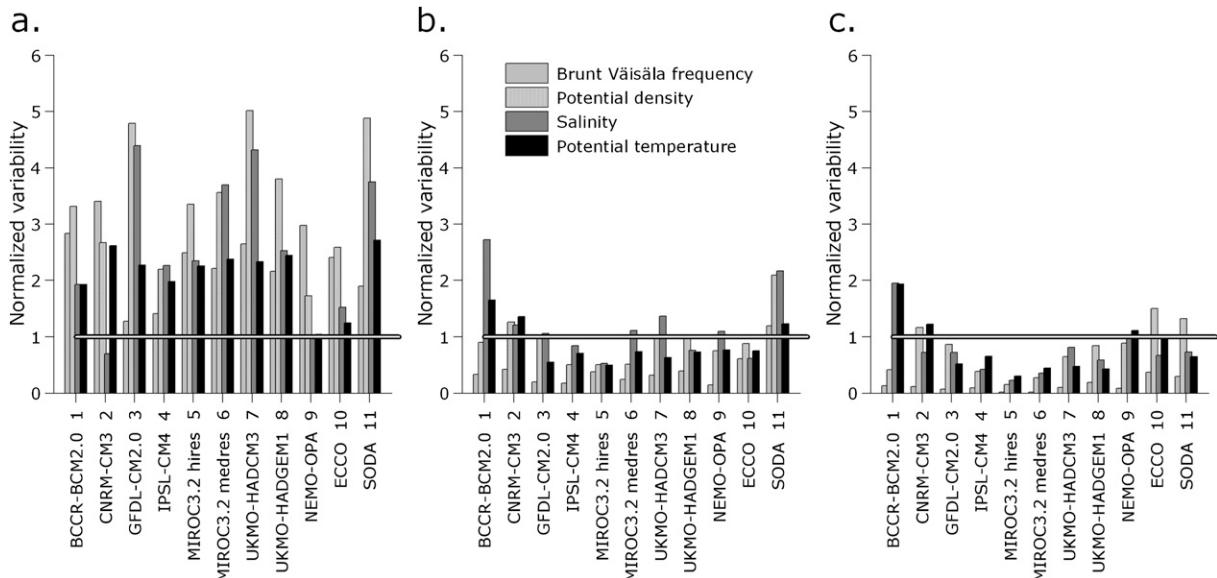


FIG. 9. Normalized variability of hydrographic properties in the Labrador Sea for (a) the surface layer, (b) the intermediate layer, and (c) the deep layer. The vertical bars represent the standard deviation of the inhomogeneous series derived from the simulations divided by the standard deviation of the observations. The thick horizontal bar represents the value at which the simulated variability is equal to the observed variability.

6. Results of the CCM–reanalysis/ocean model comparison

As expected, the reanalysis models and the NEMO–OPA ocean model perform much better than the coupled climate models in simulating both the depth-averaged properties (Fig. 3) and the layer average of hydrographic parameters (Figs. 5–8). This is expressed in low error scores in Table 3. The error scores obtained in the Labrador Sea are better than those obtained in the Irminger Sea. The potential temperature and salinity of the intermediate (LSW) layer and the deep (NEADW) layer in the Irminger Sea seem to be subject to problems similar as in the CCMs. Both the reanalyses and the NEMO–OPA model display a somewhat too high Labrador Sea surface stratification, but the NEMO–OPA ocean model performs slightly better than the reanalyses for the intermediate layer stratification (Fig. 8). The overflow waters are restricted to a shallow layer above the bottom and are not as cold as observed. The differences between the LSW in the Labrador Sea and the homogeneous water mass in the Irminger Sea are larger than observed, up to 1.1°C and 0.1 psu, but smaller than those of the coupled models. It is interesting to note that the ocean-only model performs equally well as the reanalyses, despite its low resolution. The $2^{\circ} \times 2^{\circ}$ resolution of the ocean model does not allow eddies to exist. ECCO ($1^{\circ} \times 1^{\circ}$) and SODA ($0.5^{\circ} \times 0.5^{\circ}$) have better resolutions, but are still not eddy permitting (which is reached at $0.25^{\circ} \times 0.25^{\circ}$ or better). This suggests that the eddy fluxes are not the main cause for the bias at coarse resolution. The rather large surface variability in the reanalyses could be the result of fitting ocean properties to (high temporal resolution) satellite data. The NEMO–OPA ocean model has a somewhat smaller variability. This could be related to the intensity of deep convection in the ocean model that will be discussed in the next section.

7. Discussion

The overall results of the coupled ocean–atmosphere climate simulations, for either 20 or 200 yr, are not satisfying, with model biases strongly exceeding the observed ranges of variability. In this section we discuss some possible causes for the model biases.

a. Effect of long spinup time

We have noted that the magnitude of the model biases is related to the duration of the spinup time of the model run. The three CCMs with the longest spinup periods also obtained the highest error scores for the mean hy-

drography for the Labrador Sea. Considering the five models initialized on a Levitus dataset, the correlation (R^2) between Labrador Sea overall error scores and spinup time is 0.7. During the spinup period the initial ocean state, which was originally based on observed values and thus would have presumably scored reasonably well in our study, is adjusted to equilibrium with the simulated atmosphere. Both the compensation of a “large” initial imbalance with the atmosphere and the accumulation of small numerical errors—by, for example, spurious diapycnal diffusion—drive the final ocean simulation away from the observations. Models with a larger initial imbalance usually require longer spinup periods for stabilization.

b. Mixed layer depth and intensity

As mentioned before, convection is a very important process in the northwestern North Atlantic but also a very difficult process to include in low-resolution models. The much stronger-than-observed surface layer stratification in several models and large simulation biases in the intermediate and deep layer suggest that the simulated convection is insufficient. The mixed layer depth, obtained as described in section 4, is used to investigate the convective activity of the models in the study area. The mean of the 20-yr maximum mixed layer depth series for the Labrador Sea are displayed against the mean maximum convective volume within the analyzed box (Fig. 10). The maximum mixed layer depth gives an indication of the depth to which the hydrographic properties are directly affected by local convection. The maximum convective volume gives an indication of the intensity of the convection and the amount of convectively formed water. The range of reached convection depths is denoted by the bars, giving some indication of the magnitude of the variability. Winter mixed layers in the Labrador Sea are about 500 m in years of weak convection, while the maximum observed convection depths is 2400 m or 70% of the water column (Yashayaev et al. 2008). Convective volumes are very difficult to estimate from observations as is explained by Haine et al. (2008).

The two extremes of convective regimes in the Labrador Sea stand out in the results. First, both coupled models which simulated very low surface salinities and temperatures, GFDL CM2.0 and IPSL CM4, exhibit very shallow convection (< 600 dbar) combined with a low convective volume. Thus, the improvement in surface salinity in the twentieth-century simulation of GFDL CM2.0, as described in section 5b, is most likely related to the large increase in mixed layer depth (Table 4). Second, very deep mixed layers filling a large part of

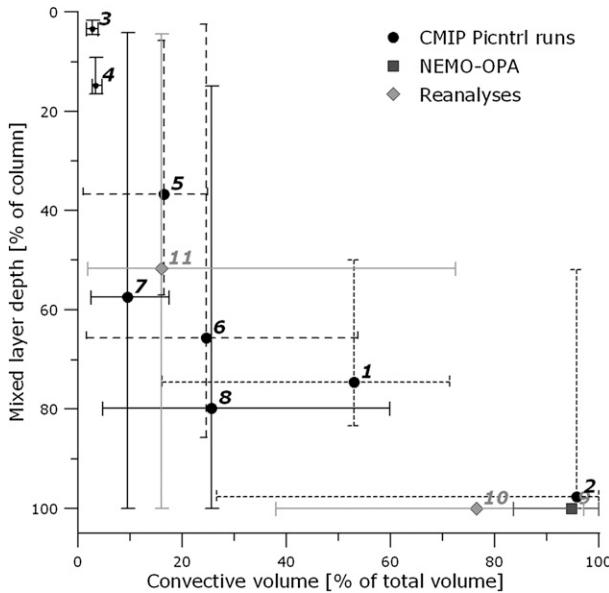


FIG. 10. Mean maximum mixed layer depth vs mean convective volume in the Labrador Sea. Shown are the CMIP coupled model simulations, the NEMO-OPA ocean model and the reanalyses. The symbols represent the means and the error bars represent the ranges (min to max) of annual maximum mixed layer depths in the 20-yr simulations. The numbers next to the symbols correspond to the model numbers in Figs. 5–8.

the volume of the basin, as seen in the CNRM-CM3 simulations, correspond to high overall salinity and potential density. In these simulations the dense and saline deep waters (such as NEADW) are mixed into the entire water column, thus increasing density/salinity and decreasing stratification over all layers.

The reanalyses and ocean model also show a large range of convective activity. Convection in SODA reaches down to 1700 m on average but displays a large variability. This is closer to the observed convection than the bottom-reaching convection of ECCO and NEMO-OPA. This may explain why the T - S profiles of SODA corresponded more closely to the observed profiles.

Five of the coupled models exhibit stronger convection in the Irminger Sea, with respect to the Labrador Sea, in the preindustrial run. In the twentieth-century run this is only seen in three of the coupled models. From the observations it is not quite so obvious that the convection in the Irminger Sea should be so strong, although some studies have shown that deep convection is possible in the southwestern Irminger Basin (Bacon et al. 2003; Pickart et al. 2003, 2008). But it is not uncommon for models to compensate shallow Labrador convection by convective formation of intermediate water in the Irminger Sea and in the central subpolar gyre (e.g., Drijfhout et al. 2007).

TABLE 4. Mean of 20-yr maximum mixed layer depth series in fraction (%) of the water column. Values are shown for the pre-industrial run (Picntrl) and the twentieth-century run (20C3M) in the Labrador Sea and the Irminger Sea. The ocean model, NEMO-OPA, and the ECCO and SODA reanalyses are based on twentieth-century forcing and ocean observations and therefore have no values for the preindustrial period.

Model	Labrador Sea		Irminger Sea	
	Picntrl	20C3M	Picntrl	20C3M
BCCR-BCM2.0	75	69	98	95
CNRM-CM3	98	97	99	91
GFDL CM2.0	3	96	29	74
IPSL CM4	15	14	61	38
MIROC3.2(hires)	37	44	35	38
MIROC3.2(medres)	66	65	60	64
UKMO HadCM3	57	43	63	56
UKMO HadGEM1	80	90	28	45
NEMO-OPA 2.0	—	100	—	100
ECCO SIO 1	—	100	—	89
SODA 1.4.2	—	52	—	76

Strong local adjustment of water masses that entered the Irminger Sea may explain the larger-than-observed differences between hydrographic properties of Labrador and Irminger water masses. Interestingly, the IPSL CM4 model exhibits much too low surface salinity in the Irminger Sea despite the fact that the simulated convection is quite strong and the stratification is closer to the observed values.

c. Surface fluxes and coupling to atmospheric or sea ice model

Convective mixing depths in the Labrador and Irminger Seas can only partially explain the biases of the coupled models. IPSL CM4 shows similar salinity biases over the Labrador and Irminger Seas despite the large difference in convection strength. This suggests that the shallow Labrador convection is the result of the too strong surface stratification rather than the too strong stratification being the result of shallow convection. Important contributors to surface stratification are air-sea surface fluxes. The small biases observed in the NEMO-OPA simulations suggest that the surface fluxes are very important. The final error score of NEMO-OPA is comparable to those of the ocean reanalyses, without incorporating ocean observations like these reanalyses. The large gap between observed and modeled depth-averaged salinity and temperature (Fig. 4) thus seem to be closed by the employment of close-to-observed surface fluxes. The low surface salinity seen in many of the models indicates that the hydrological cycle may be the leading contributor among the surface fluxes. Josey and Marsh (2005) and Myers et al. (2007) showed that positive anomalies in precipitation-evaporation play

a significant role in the freshening of both the eastern Subpolar Gyre and the Labrador Sea. However, both the analysis of precipitation over the two basins and of the zonal mean precipitation between 40° and 70°N showed that simulations with a too fresh surface layer do not have too much precipitation over the North Atlantic. Possibly a too small amount of evaporation, advection of freshwaters, or too much melting sea ice is causing the surface salinity bias. Because of limited resources these processes fall outside the scope of this study. The exact cause for the observed salinity bias in CCMs is left as an interesting subject for future research.

8. Conclusions

This study aims to assess the performance of coupled climate models in simulating the local hydrography of two basins in the North Atlantic Ocean: the Labrador and Irminger Seas. Many of the constituents of NADW, which forms the southward deep branch of the MOC, are either formed or adjusted in these basins. Some of the local processes, for example, deep convection and entrainment over the overflows, cannot yet be explicitly represented in ocean models and have been parameterized. Although coupled climate models are by their definition meant to simulate the global climate as accurately as possible and not merely a relatively small area such as the northwestern North Atlantic Ocean, it is nevertheless an important area for the heat and freshwater transport in the MOC and the climate in the North Atlantic region.

The reported biases in both the preindustrial and twentieth-century simulations strongly exceed the ranges of observed hydrographic properties. Vertical profiles of salinity and potential temperature show that the water column below 500 dbar is generally too warm and too saline. Especially in the Labrador Sea, the surface layers are characterized by low salinity, leading to a too strong surface stratification. In some of the models this situation is not unlike a permanent “Great Salinity Anomaly” as was seen circulating the North Atlantic in the 1960s and 1970s (Dickson et al. 1988), which arrested convection in the Labrador Sea. Because of the strong stratification, communication of atmospheric variability to deeper ocean layers is inhibited and surface variability is strengthened. The analysis of two 200-yr simulated datasets show that the previous results are not an artifact of the arbitrary selection of a 20-yr segment from a time series containing multidecadal oscillations. Both the mean and the variability of the simulated hydrographic properties change only minimally when the analyzed time series are extended from

20 to 200 yr. It is expected that this is also true for the other models.

The size of the biases in the simulations is related to the length of the spinup time of the model and most likely to the initial imbalance between model components. During the spinup the simulated ocean state is slowly pushed away from the initial (observed) ocean state and small numerical errors are accumulated.

A large range of convective regimes can be found in the coupled model simulations. Very shallow convective regimes are related to large negative salinity biases in the surface layer. In contrast, deep and strong convective regimes are related to high salinity and density over the entire water column. In between these extremes the correlation of convective depth with hydrographic biases is ambiguous. Possibly the deep saline and warm layer observed in the simulations is formed by convection in the Irminger Sea, which is stronger than expected in the simulations. Low surface salinities and the reasonably good performance of the ocean model NEMO-OPA, which uses observed ocean fluxes, suggest that biases in sea surface fluxes are also contributing to the ocean simulation biases. Sensitivity studies regarding the full hydrological cycle and sea ice are needed to improve the freshwater balance in the northwestern North Atlantic Ocean.

Acknowledgments. This research received financial support from the Dutch BSIK-CS1 project.

The authors kindly thank I. Yashayaev, who provided the dataset containing the hydrographic profiles from the Labrador Sea, and J. Meincke, D. Quadfasel, and M. Bersch who made the nonpublic data from the Irminger Sea available.

The authors acknowledge the modeling groups for making their simulations available for analysis, the Program for Climate Model Diagnosis and Intercomparison (PCMDI) for collecting and archiving the CMIP3 model output, and the WCRP’s Working Group on Coupled Modelling (WCCM) for organizing the model data analysis activity. The WCRP CMIP3 multimodel dataset is supported by the Office of Science, U.S. Department of Energy. The NEMO-OPA model was provided by the DRAKKAR Group.

The ECCO state estimates were provided by the Consortium for Estimating the Circulation and Climate of the Ocean funded by the National Oceanographic Partnership Program (NOPP).

The SODA estimates were obtained from the data server at the Asia-Pacific Data-Research Center via their Web site at <http://apdrc.soest.hawaii.edu/>.

We thank the anonymous reviewers, whose remarks and suggestions have strongly improved the earlier manuscript.

REFERENCES

- Bacon, S., W. J. Gould, and Y. Jia, 2003: Open-ocean convection in the Irminger Sea. *Geophys. Res. Lett.*, **30**, 1246, doi:10.1029/2002GL016271.
- Barnier, B., and Coauthors, 2007: Eddy-permitting ocean circulation hindcasts of past decades. *CLIVAR Exchanges*, No. 42, International CLIVAR Project Office, Southampton, United Kingdom, 8–10.
- Brambilla, E., and L. D. Talley, 2008: Subpolar Mode Water in the northeastern Atlantic: 1. Averaged properties and mean circulation. *J. Geophys. Res.*, **113**, C04025, doi:10.1029/2006JC004062.
- , —, and P. E. Robbins, 2008: Subpolar Mode Water in the northeastern Atlantic: 2. Origin and transformation. *J. Geophys. Res.*, **113**, C04026, doi:10.1029/2006JC004063.
- Carton, J. A., B. S. Giese, and S. A. Grodsky, 2005: Sea level rise and the warming of the oceans in the Simple Ocean Data Assimilation (SODA) ocean reanalysis. *J. Geophys. Res.*, **110**, C09006, doi:10.1029/2004JC002817.
- Delworth, T. L., and Coauthors, 2006: GFDL's CM2 global coupled climate models. Part I: Formulation and simulation characteristics. *J. Climate*, **19**, 643–674.
- Dickson, R. R., J. Meincke, S.-A. Malmberg, and A. J. Lee, 1988: The “great salinity anomaly” in the northern North Atlantic 1968–1982. *Prog. Oceanogr.*, **20**, 103–151.
- Donners, J., S. S. Drijfhout, and W. Hazeleger, 2005: Water mass transformation and subduction in the South Atlantic. *J. Phys. Oceanogr.*, **35**, 1841–1860.
- Drijfhout, S., W. Hazeleger, F. Selten, and R. Haarsma, 2007: Future changes in internal variability of the Atlantic meridional overturning circulation. *Climate Dyn.*, **30**, 407–419, doi:10.1007/s00382-007-0297-y.
- Furevik, T., M. Bentsen, H. Drange, I. K. T. Kindem, N. G. Kvamstø, and A. Sorteberg, 2003: Description and evaluation of the Bergen climate model: ARPEGE coupled with MICOM. *Climate Dyn.*, **21**, 27–51.
- Gordon, C., C. Cooper, C. A. Senior, H. Banks, J. M. Gregory, T. C. Johns, J. F. B. Mitchell, and R. A. Wood, 2000: The simulation of SST, sea ice extents and ocean heat transport in a version of the Hadley Centre coupled model without flux adjustments. *Climate Dyn.*, **16**, 147–168.
- Haine, T., and Coauthors, 2008: North Atlantic Deep Water formation in the Labrador Sea, recirculation through the subpolar gyre, and discharge to the subtropics. *Arctic-Subarctic Ocean Fluxes: Defining the Role of the Northern Seas in Climate*, R. Dickson, J. Meincke, and P. Rhines, Eds., Springer, 653–702.
- Hasumi, H., and S. Emori, Eds., 2004: K-1 coupled GLM (MIROC) description. Center for Climate System Research K-1 Tech. Rep. 1, 39 pp.
- Hendry, R. M., H. van Aken, and I. Yashayaev, 2007: Monitoring the ventilation of the Irminger and Labrador Seas. *CLIVAR Exchanges*, No. 40, International CLIVAR Project Office, Southampton, United Kingdom, 25–27.
- Holliday, N. P., A. Meyer, S. Bacon, S. G. Alderson, and B. de Cuevas, 2007: Retroflection of part of the east Greenland current at Cape Farewell. *Geophys. Res. Lett.*, **34**, L07609, doi:10.1029/2006GL029085.
- Hurrell, J. W., 1995: Decadal trends in the North Atlantic Oscillation regional temperatures and precipitation. *Science*, **269**, 676–679.
- Johns, T., and Coauthors, 2004: HadGEM1—Model description and analysis of preliminary experiments for the IPCC Fourth Assessment Report. Hadley Centre Tech. Note 55, 75 pp.
- Josey, S. A., and R. Marsh, 2005: Surface freshwater flux variability and recent freshening of the North Atlantic in the eastern subpolar gyre. *J. Geophys. Res.*, **110**, C05008, doi:10.1029/2004JC002521.
- Katsman, C. A., M. A. Spall, and R. S. Pickart, 2004: Boundary current eddies and their role in the restratification of the Labrador Sea. *J. Phys. Oceanogr.*, **34**, 1967–1983.
- Köhl, A., D. Stammer, B. Cornuelle, E. Remy, Y. Lu, P. Heimbach and C. Wunsch, 2003: The global 1° WOCE synthesis: 1992–2001. The ECCO Report Series, Rep. 20, 33 pp.
- Lazier, J., R. Hendry, A. Clarke, I. Yashayaev, and P. Rhines, 2002: Convection and restratification in the Labrador Sea, 1990–2000. *Deep-Sea Res. I*, **49**, 1819–1835.
- Levitus, S., 1982: *Climatological Atlas of the World Ocean*. NOAA Prof. Paper 13, 173 pp. and 17 microfiche.
- , and T. P. Boyer, 1994: *Temperature*. Vol. 4, *World Ocean Atlas 1994*, NOAA Atlas NESDIS 4, 117 pp.
- , R. Burgett, and T. P. Boyer, 1995: *Salinity*. Vol. 3, *World Ocean Atlas 1994*, NOAA Atlas NESDIS 3, 99 pp.
- , and Coauthors, 1998: *Introduction*. Vol. 1, *World Ocean Database*, NOAA Atlas NESDIS 18, 346 pp.
- Lilly, J. M., P. B. Rhines, F. Schott, K. Lavender, J. Lazier, U. Send, and E. D'Asaro, 2003: Observations of the Labrador Sea eddy field. *Prog. Oceanogr.*, **59**, 75–176.
- Madec, G., 2008: NEMO ocean engine. Note du Pôle de Modélisation 27, Institut Pierre-Simon Laplace, 209 pp.
- Marshall, J., and F. Schott, 1999: Open-ocean convection: Observations, theory and models. *Rev. Geophys.*, **37**, 1–64.
- , and Coauthors, 2001: North Atlantic climate variability: Phenomena, impacts and mechanisms. *Int. J. Climatol.*, **21**, 1863–1898, doi:10.1002/joc.693.
- Marti, O., and Coauthors, 2005: The new IPSL climate system model: IPSL-CM4. Institut Pierre-Simon Laplace Note du Pôle de Modélisation 26, 79 pp.
- McCartney, M. S., and L. D. Talley, 1982: The Subpolar Mode Water of the North Atlantic. *J. Phys. Oceanogr.*, **12**, 1169–1188.
- Meehl, G. A., C. Covey, T. Delworth, M. Latif, B. McAvaney, J. F. B. Mitchell, R. J. Stouffer, and K. E. Taylor, 2007: The WCRP CMIP3 multimodel dataset. *Bull. Amer. Meteor. Soc.*, **88**, 1383–1394.
- Myers, P. G., S. A. Josey, B. Wheler, and N. Kuland, 2007: Interdecadal variability in Labrador Sea precipitation minus evaporation and salinity. *Prog. Oceanogr.*, **73**, 341–357, doi:10.1016/j.pocean.2006.06.003.
- Pickart, R. S., F. Straneo, and G. W. K. Moore, 2003: Is Labrador Sea water formed in the Irminger basin? *Deep-Sea Res. I*, **50**, 23–52.
- , K. Våge, G. W. K. Moore, I. A. Renfrew, M. H. Ribergaard, and H. C. Davies, 2008: Convection in the western North Atlantic sub-polar gyre: Do small scale wind events matter? *Arctic-Subarctic Ocean Fluxes: Defining the Role of the Northern Seas in Climate*, R. Dickson, J. Meincke, and P. Rhines, Eds., Springer, 629–652.
- Salas-Mélia, D., and Coauthors, 2005: Description and validation of the CNRM-CM3 global coupled model. Centre National de Recherches Météorologiques Note 103, 36 pp.
- Schmittner, A., M. Latif, and B. Schneider, 2005: Model projection of the North Atlantic thermohaline circulation for the 21st century assessed by observations. *Geophys. Res. Lett.*, **32**, L23710, doi:10.1029/2005GL024368.

- Sloyan, B. M., and I. V. Kamenkovich, 2007: Simulation of Sub-Antarctic Mode and Antarctic Intermediate Waters in climate models. *J. Climate*, **20**, 5061–5080.
- Solomon, S., D. Qin, M. Manning, M. Marquis, K. Averyt, M. M. B. Tignor, H. L. Miller Jr., and Z. Chen, Eds., 2007: *Climate Change 2007: The Physical Science Basis*. Cambridge University Press, 996 pp.
- Tanhua, T., K. Bulsiewicz, and M. Rhein, 2005: Spreading of overflow water from the Greenland to the Labrador Sea. *Geophys. Res. Lett.*, **32**, L10605, doi:10.1029/2005GL022700.
- van Aken, H. M., 2007: *The Oceanic Thermohaline Circulation: An Introduction*. Springer, 326 pp.
- , and C. J. de Boer, 1995: On the synoptic hydrography of intermediate and deep water masses in the Iceland Basin. *Deep-Sea Res. I*, **42**, 165–189.
- Volkov, D. L., 2005: Interannual variability of the altimetry-derived eddy field and surface circulation in the extratropical North Atlantic Ocean in 1993–2001. *J. Phys. Oceanogr.*, **35**, 405–426.
- Yashayaev, I., 2007: Hydrographic changes in the Labrador Sea, 1960–2005. *Prog. Oceanogr.*, **73**, 242–276.
- , M. Bersch, and H. M. van Aken, 2007: Spreading of the Labrador Sea Water to the Irminger and Iceland basins. *Geophys. Res. Lett.*, **34**, L10602, doi:10.1029/2006GL028999.
- , N. P. Holliday, M. Bersch, and H. M. van Aken, 2008: The history of the Labrador Sea Water: Production, spreading, transformation and loss. *Arctic–Subarctic Ocean Fluxes: Defining the Role of the Northern Seas in Climate*, R. Dickson, J. Meincke, and P. Rhines, Eds., Springer, 569–612.