

## Highlights

- Inter-annual to decadal variability in AMOC from CORE-II simulations is presented.
- AMOC variability shows three stages, with maximum transports in mid- to late-1990s.
- North Atlantic temporal variability features are in good agreement among simulations.
- Such agreements suggest variability is dictated by the atmospheric data sets.
- Simulations differ in spatial structures of variability due to ocean dynamics.

# North Atlantic Simulations in Coordinated Ocean-ice Reference Experiments phase II (CORE-II). Part II: Inter-Annual to Decadal Variability

Gokhan Danabasoglu<sup>a,\*</sup>, Steve G. Yeager<sup>a</sup>, Who M. Kim<sup>b</sup>, Erik Behrens<sup>c</sup>, Mats Bentsen<sup>d</sup>, Daohua Bi<sup>e</sup>, Arne Biastoch<sup>c</sup>, Rainer Bleck<sup>f,g</sup>, Claus Böning<sup>c</sup>, Alexandra Bozec<sup>h</sup>, Vittorio M. Canuto<sup>g</sup>, Christophe Cassou<sup>i</sup>, Eric Chassignet<sup>h</sup>, Andrew C. Coward<sup>j</sup>, Sergey Danilov<sup>k</sup>, Nikolay Diansky<sup>l</sup>, Helge Drange<sup>m</sup>, Riccardo Farneti<sup>n</sup>, Elodie Fernandez<sup>i,o</sup>, Pier Giuseppe Fogli<sup>p</sup>, Gael Forget<sup>q</sup>, Yosuke Fujii<sup>r</sup>, Stephen M. Griffies<sup>s</sup>, Anatoly Gusev<sup>l</sup>, Patrick Heimbach<sup>q</sup>, Armando Howard<sup>g,t</sup>, Mehmet Ilicak<sup>d</sup>, Thomas Jung<sup>k</sup>, Alicia R. Karspeck<sup>a</sup>, Maxwell Kelley<sup>g</sup>, William G. Large<sup>a</sup>, Anthony Leboissetier<sup>g</sup>, Jianhua Lu<sup>h</sup>, Gurvan Madec<sup>u</sup>, Simon J. Marsland<sup>e</sup>, Simona Masina<sup>p,v</sup>, Antonio Navarra<sup>p,v</sup>, A. J. George Nurser<sup>j</sup>, Anna Pirani<sup>w</sup>, Anastasia Romanou<sup>g,x</sup>, David Salas y Méliá<sup>y</sup>, Bonita L. Samuels<sup>s</sup>, Markus Scheinert<sup>c</sup>, Dmitry Sidorenko<sup>k</sup>, Shan Sun<sup>f</sup>, Anne-Marie Treguier<sup>z</sup>, Hiroyuki Tsujino<sup>r</sup>, Petteri Uotila<sup>e,aa</sup>, Sophie Valcke<sup>i</sup>, Aurore Voldoire<sup>y</sup>, Qiang Wang<sup>k</sup>, Igor Yashayaev<sup>ab</sup>

<sup>a</sup>National Center for Atmospheric Research (NCAR), Boulder, CO, USA

<sup>b</sup>Texas A & M University, College Station, TX, USA

<sup>c</sup>Helmholtz Center for Ocean Research, GEOMAR, Kiel, Germany

<sup>d</sup>Uni Research Climate, Bjerknes Centre for Climate Research, Bergen, Norway

<sup>e</sup>Centre for Australian Weather and Climate Research, a partnership between CSIRO and the Bureau of Meteorology, Commonwealth Scientific and Industrial Research Organisation (CSIRO), Melbourne, Australia

<sup>f</sup>NOAA Earth System Research Laboratory, Boulder, CO, USA

<sup>g</sup>NASA Goddard Institute for Space Studies (GISS), New York, NY, USA

<sup>h</sup>Center for Ocean-Atmospheric Prediction Studies (COAPS), Florida State University, Tallahassee, FL, USA

<sup>i</sup>Centre Européen de Recherche et de Formation Avancée en Calcul Scientifique (CERFACS), Toulouse, France

<sup>j</sup>National Oceanography Centre Southampton (NOCS), Southampton, UK

<sup>k</sup>Alfred Wegener Institute for Polar and Marine Research (AWI), Bremerhaven, Germany

<sup>l</sup>Institute of Numerical Mathematics, Russian Academy of Sciences, Moscow, Russia

<sup>m</sup>Geophysical Institute, University of Bergen and Bjerknes Centre for Climate Research, Bergen, Norway

<sup>n</sup>International Centre for Theoretical Physics (ICTP), Trieste, Italy

<sup>o</sup>Mercator-Ocean, Toulouse, France

\*Corresponding author. Tel.: +1 303 497 1604; fax: +1 303 497 1700.

Email address: gokhan@ucar.edu (Gokhan Danabasoglu)

<sup>p</sup> *Centro Euro-Mediterraneo sui Cambiamenti Climatici (CMCC), Bologna, Italy*

<sup>q</sup> *Massachusetts Institute of Technology, Cambridge, MA, USA*

<sup>r</sup> *Meteorological Research Institute (MRI), Japan Meteorological Agency, Tsukuba, Japan*

<sup>s</sup> *NOAA Geophysical Fluid Dynamics Laboratory (GFDL), Princeton, NJ, USA*

<sup>t</sup> *Medgar Evers College of the City University of New York, Brooklyn, NY, USA*

<sup>u</sup> *IPSL/LOCEAN, CNRS-IRD-UPMC, Paris, France*

<sup>v</sup> *Istituto Nazionale di Geofisica e Vulcanologia (INGV), Bologna, Italy*

<sup>w</sup> *International CLIVAR Project Office, ICTP, Trieste, Italy*

<sup>x</sup> *Columbia University, New York, NY, USA*

<sup>y</sup> *Centre National de Recherches Météorologiques (CNRM-GAME), Toulouse, France*

<sup>z</sup> *Laboratoire de Physique des Océans, UMR 6523, CNRS-Ifremer-IRD-UBO, IUEM, Plouzane, France*

<sup>aa</sup> *Finnish Meteorological Institute, Helsinki, Finland*

<sup>ab</sup> *Bedford Institute of Oceanography, Fisheries and Oceans Canada, Dartmouth, Canada*

---

## Abstract

Simulated inter-annual to decadal variability and trends in the North Atlantic for the 1958–2007 period from twenty global ocean – sea-ice coupled models are presented. These simulations are performed as contributions to the second phase of the Coordinated Ocean-ice Reference Experiments (CORE-II). The study is Part II of our companion paper (Danabasoglu et al., 2014) which documented the mean states in the North Atlantic from the same models. A major focus of the present study is the representation of Atlantic meridional overturning circulation (AMOC) variability in the participating models. Relationships between AMOC variability and those of some other related variables, such as subpolar mixed layer depths, the North Atlantic Oscillation (NAO), and the Labrador Sea upper-ocean hydrographic properties, are also investigated. In general, AMOC variability shows three distinct stages. During the first stage that lasts until the mid- to late-1970s, AMOC is relatively steady, remaining lower than its long-term (1958–2007) mean. Thereafter, AMOC intensifies with maximum transports achieved in the mid- to late-1990s. This

enhancement is then followed by a weakening trend until the end of our integration period. This sequence of low frequency AMOC variability is consistent with previous studies. Regarding strengthening of AMOC between about the mid-1970s and the mid-1990s, our results support a previously identified variability mechanism where AMOC intensification is connected to increased deep water formation in the subpolar North Atlantic, driven by NAO-related surface fluxes. The simulations tend to show general agreement in their representations of, for example, AMOC, sea surface temperature (SST), and subpolar mixed layer depth variabilities. In particular, the observed variability of the North Atlantic SSTs is captured well by all models. These findings indicate that simulated variability and trends are primarily dictated by the atmospheric datasets which include the influence of ocean dynamics from nature superimposed onto anthropogenic effects. Despite these general agreements, there are many differences among the model solutions, particularly in the spatial structures of variability patterns. For example, the location of the maximum AMOC variability differs among the models between Northern and Southern Hemispheres.

*Keywords:*

Global ocean – sea-ice modelling, Ocean model comparisons, Atmospheric forcing, Inter-annual to decadal variability and mechanisms, Atlantic meridional overturning circulation variability, Variability in the North Atlantic

---

## 1. Introduction

This study presents an analysis of the simulated inter-annual to decadal variability and trends in the North Atlantic Ocean for the 1958–2007 period from a set of simulations participating in the second phase of the Coordinated Ocean-ice Reference Experiments (CORE-II). It is Part II of our companion paper, Danabasoglu et al. (2014) (hereafter DY14), where the mean states in the Atlantic basin from these

7 simulations are documented to provide a baseline for the present variability analysis.

8 Our primary focus is again on the Atlantic meridional overturning circulation  
9 (AMOC), but here we investigate representation of its inter-annual to decadal vari-  
10 ability and trends in the participating models. As stated in DY14, AMOC is pre-  
11 sumed to play a major role in decadal and longer time scale climate variability and in  
12 prediction of the earth's future climate on these time scales through its heat and salt  
13 transports and its impacts on sea surface temperatures (SSTs) and sea level. Due to  
14 lack of long and continuous AMOC observations, the main support for such an im-  
15 portant role for AMOC in influencing the earth's climate comes from coupled general  
16 circulation model (CGCM) simulations. In long control simulations with CGCMs,  
17 usually for pre-industrial conditions run without either changes in radiative forcings  
18 or inclusion of anthropogenic forcings, AMOC intrinsic variability is rather rich with  
19 a variety of time scales, e.g., inter-annual, decadal, centennial. Furthermore, such  
20 low frequency AMOC anomalies tend to precede the basin scale SST anomalies in the  
21 Atlantic Ocean, thus suggesting a driving role for AMOC in models (e.g., Delworth  
22 et al., 1993; Danabasoglu, 2008; Kwon and Frankignoul, 2012; Delworth and Zeng,  
23 2012; Danabasoglu et al., 2012). Hence, the basin scale, low frequency variability  
24 (40–70 year period) of the observed SSTs in the Atlantic Ocean is assumed to be  
25 linked to AMOC fluctuations. This basin scale SST variability is usually referred to  
26 as the Atlantic Multidecadal Variability (AMV) or Atlantic Multidecadal Oscillation.  
27 AMV represents an index of detrended, observed (North) Atlantic SST variability  
28 estimated from instrumental records and proxy data (Schlesinger and Ramankutty,  
29 1994; Kushnir, 1994; Delworth and Mann, 2000). We also note that some studies  
30 suggest that variability of AMOC and upper-ocean temperatures may be potentially  
31 predictable on decadal time scales (e.g., Griffies and Bryan, 1997; Pohlmann et al.,  
32 2004; Msadek et al., 2010; Branstator and Teng, 2010), thus making appropriate

33 initialization of the AMOC state for decadal prediction experiments an important  
34 endeavor.

35 For studies of AMOC variability and its mechanisms and prediction, CGCMs  
36 are an essential tool. However, their fidelity remains a serious concern, and a fun-  
37 damental understanding of the mechanisms of simulated AMOC variability remains  
38 elusive (see Liu (2012) and Srokosz et al. (2012) for recent reviews). For example,  
39 the magnitude and dominant time scales of AMOC variability and its mechanisms  
40 can differ substantially from one model to another (see above references), from one  
41 version of a model to another (Danabasoglu, 2008; Danabasoglu et al., 2012), and, in  
42 some cases, even from one time segment of a model simulation to another (Kwon and  
43 Frankignoul, 2012, 2014). Some oceanic subgrid scale parameterizations are shown  
44 to affect the variability of AMOC as well, e.g., magnitude of vertical diffusivity coef-  
45 ficients (Farneti and Vallis, 2011); representation of the Nordic Sea overflows (Yeager  
46 and Danabasoglu, 2012) and of meso- and submesoscale eddies (Danabasoglu et al.,  
47 2012). In addition, various aspects of AMOC variability are sensitive to both the  
48 atmosphere and ocean model resolutions (Bryan et al., 2006). Given these signif-  
49 icant model sensitivities and many unanswered questions, there is a critical need  
50 for improving our understanding of the mechanisms and assessing the fidelity and  
51 robustness of simulated AMOC variability against limited available observations.

52 The CORE-II hindcast experiments provide a common framework to address  
53 some of these issues. Specifically, they can be used to investigate AMOC variabil-  
54 ity and its mechanisms on seasonal, inter-annual, and decadal time scales and to  
55 understand and separate forced variability from natural variability – the latter in  
56 combination with (coupled) control experiments that exclude external and anthro-  
57 pogenic effects. Additionally, robustness of variability mechanisms across models can  
58 be evaluated. Continuous, observationally-based estimates of AMOC are available

59 only starting in early 2004 through the Rapid Climate Change transbasin observ-  
60 ing array installed along 26.5°N (RAPID; Cunningham et al., 2007). The CORE-II  
61 hindcasts – along with the reanalysis products – can provide complementary infor-  
62 mation on AMOC for the pre-RAPID era. Unfortunately, for our current work, the  
63 overlap period between the RAPID estimates and the model simulations is rather  
64 short, i.e., April 2004 through December 2007, making our annual-mean comparisons  
65 rather crude. Nevertheless, the solutions from the CORE-II hindcasts can be com-  
66 pared against other available observations in their representations of certain climate  
67 events, such as the mid-1990s warming of the subpolar North Atlantic. Identified  
68 variability mechanisms or their drivers associated with such events are expected to  
69 provide insight on AMOC variability in general, even though the CORE-II simu-  
70 lations cannot directly address intrinsic inter-annual to multi-decadal AMOC vari-  
71 ability because the forcing data sets include external and anthropogenic effects. We  
72 note that several individual model studies, using the CORE-II protocol, have already  
73 demonstrated many realistic features of mean and variability in the North Atlantic in  
74 CORE-II hindcasts, including an investigation of the AMOC variability mechanisms  
75 associated with the mid-1990s warming of the subpolar North Atlantic (e.g., Yeager  
76 et al., 2012; Yeager and Danabasoglu, 2014; Gusev and Diansky, 2014).

77 Use of such hindcast simulations to investigate variability in the North Atlantic,  
78 particularly of the AMOC, is not new (e.g., Häkkinen, 1999; Eden and Willebrand,  
79 2001; Bentsen et al., 2004; Beismann and Barnier, 2004; Böning et al., 2006; Biastoch  
80 et al., 2008; Deshayes and Frankignoul, 2008; Lohmann et al., 2009b; Brodeau et al.,  
81 2010; Robson et al., 2012). These studies employ various historical atmospheric  
82 datasets, e.g., National Centers for Environmental Prediction – National Center for  
83 Atmospheric Research (NCEP/NCAR) reanalysis (Kalnay et al., 1996), European  
84 Center for Medium-range Weather Forecasting (ECMWF) ERA-40 reanalysis (Up-

85 pala et al., 2005), or a combination of other datasets, to force regional Atlantic basin  
86 or global ocean models. They – along with the CORE-II hindcast studies men-  
87 tioned in the previous paragraph – show that AMOC variability on inter-annual to  
88 decadal time scales is connected to surface buoyancy fluxes and wind stress asso-  
89 ciated with the North Atlantic Oscillation (NAO). A particularly robust feature of  
90 these and other studies is the strengthening of AMOC during the last few decades  
91 of the twentieth century. Specifically, the persistent positive NAO (NAO+) that  
92 occurred between the early 1970s and the mid-1990s is credited with enhanced deep  
93 water formation (DWF) and associated deepening of mixed layers in the subpolar  
94 North Atlantic, particularly in the Labrador Sea (LS) region. This in turn results  
95 in increased AMOC and northward heat transports that have been identified as the  
96 major contributors to the mid-1990s subpolar North Atlantic warming (e.g., Robson  
97 et al., 2012; Yeager et al., 2012). We note that this AMOC variability mechanism  
98 suggesting a prominent role for the NAO is very similar to the AMOC intrinsic  
99 variability mechanisms found in many CGCM control simulations (e.g., Dong and  
100 Sutton, 2005; Teng et al., 2011; Danabasoglu et al., 2012).

101 In the present study, our primary goal is to provide an evaluation of how partici-  
102 pating models represent trends and variability in AMOC and in some other fields on  
103 inter-annual to decadal time scales under the common CORE-II forcing, with a focus  
104 on the North Atlantic. With the variability mechanism described above providing  
105 a background, other goals include i) an investigation of robust aspects of AMOC  
106 variability in these coarse resolution models in the presence of mean state differences  
107 discussed in DY14 and ii) an exploration of relationships between AMOC variability  
108 and those of some other fields such as NAO, mixed layer depths (MLDs), and the  
109 LS upper-ocean temperature, salinity, and density.

110 The paper is organized as follows. In section 2, we briefly summarize the CORE-



111 II framework, analysis methods, and participating models, including two additional  
112 contributions (labeled as FSU2 and GISS2) to those used in DY14. We document  
113 the variabilities in AMOC; North Atlantic SSTs; North Atlantic MLDs; upper-ocean  
114 central LS hydrographic properties; and subpolar gyre (SPG) circulation and SPG  
115 sea surface height (SSH) in sections 3 through 7. We then present the relationships  
116 between AMOC variability and i) those of meridional heat transport (MHT) in sec-  
117 tion 8 and ii) those of LS MLD, SPG circulation, SPG SSH, and NAO in section 9.  
118 The last section, i.e., section 10, has a summary and our conclusions. We provide  
119 short summaries of FSU2 and GISS2 along with a note on their vertical coordinate  
120 choices and a brief evaluation of their mean states in the North Atlantic in Appendix  
121 A. Appendix B details the departures from the CORE-II protocol that occurred in  
122 nearly half of the participating models. Finally, a list of major acronyms is included  
123 in Appendix C.

## 124 **2. CORE-II framework, models, and analysis methods**

125 The CORE-II experiments represent ocean – sea-ice hindcast simulations forced  
126 with the inter-annually varying atmospheric datasets over the 60-year period from  
127 1948 to 2007. These forcing datasets were developed by Large and Yeager (2004,  
128 2009). The CORE-II protocol requests that the simulations are integrated for no  
129 less than five repeat cycles of the 60-year forcing. There is no restoring term applied  
130 to SSTs. However, a form of surface salinity restoring may be used to prevent  
131 unbounded local salinity trends. Details of the CORE-II protocol are given in Griffies  
132 et al. (2012) and DY14.

133 Our present study includes two additional contributions to those used in DY14,  
134 thus bringing the total number of participating models to twenty. Both of the new  
135 participants, labeled as FSU2 and GISS2, are based on the HYbrid Coordinate Ocean

136 Model (HYCOM). The FSU simulation in DY14 uses an earlier HYCOM version  
137 which advects density and salinity, thus does not conserve heat. In contrast, FSU2  
138 employs a formulation that advects temperature and salinity, conserving heat. GISS2  
139 also uses this latter formulation and represents an updated version of the model de-  
140 scribed in Sun and Bleck (2006). Summaries of FSU2 and GISS2 model descriptions  
141 are provided in Appendix A.1 and Appendix A.2, respectively. For the descriptions  
142 of other models and their surface salinity restoring details, we refer to the Appen-  
143 dices in DY14. We use the same model naming convention in the present study as in  
144 DY14. For completeness and reference purposes, an updated list of the participating  
145 groups along with their model names and resolutions is reproduced in Table 1.

146 After the publications of DY14 and Griffies et al. (2014), it came to our attention  
147 that about half of the participating models did depart from the CORE-II protocol  
148 recommendations. These departures, detailed in Appendix B, include use of different  
149 bulk formulae, modifications of the Large and Yeager (2009) bulk formulae, and  
150 changes in the forcing datasets.

151 The 60-year repeat forcing cycle introduces an unphysical jump in the forcing  
152 from 2007 back to 1948 with the ocean state in 1948 identical to that of the end  
153 state of the forcing cycle. This approach impacts the solutions during the early years  
154 of the forcing period. Our analysis here uses only the 1958–2007 period from the  
155 fifth cycle of the simulations to partially avoid any adverse effects of this artificial  
156 jump in forcing. We employ standard correlation, regression, and empirical orthogo-  
157 nal function (EOF) analysis methods. The principal component (PC) time series are  
158 normalized to have unit variance. Thus, the EOF spatial pattern magnitudes cor-  
159 respond to one standard deviation changes in the PC time series. Unless otherwise  
160 noted, the time series are based on annual-mean data. In most of our analysis, we  
161 choose not to detrend the time series, because our interests include low-frequency,

162 e.g., decadal, variability and trends. As discussed in DY14, about half of the models  
163 reach a practical AMOC equilibrium state as measured by small root-mean-square  
164 differences and high correlations of their AMOC time series between the fourth and  
165 fifth forcing cycles. However, remaining models, i.e., AWI, FSU, GFDL-MOM, ICTP,  
166 INMOM, and KIEL, as well as the two new contributions FSU2 and GISS2, do not  
167 fully obtain such an equilibrium state and show ongoing drifts in their AMOCs (see  
168 Figs. 1 and 2 of DY14), likely impacting magnitudes of some of our calculated trends.  
169 The time series are decomposed into their high- and low-frequency contents, using  
170 a Butterworth filter with a somewhat arbitrary cutoff period of 7 years. In some  
171 of the figures with time series, we also include the time series for the multi model  
172 mean, denoted as MMM. The MMM time series do not include MRI-A – the only  
173 contribution with data assimilation. The solutions from this MRI-A simulation are  
174 also provided to the Karspeck et al. (2015) study where a comparison of AMOC  
175 mean, variability, and trends from six data assimilation products is presented.

176 The statistical significance of various lead–lag correlations is examined using a  
177 Monte Carlo approach called a *parametric bootstrap*. In this approach, we assume  
178 that the annual average statistical properties of the variables being considered (e.g.,  
179 AMOC and MLD) can be modeled as a first-order auto regressive process (AR1),  
180 with variance and damping coefficient estimated from the model time series (without  
181 low-pass filtering). Consistent with a standard t-test for evaluating the significance  
182 of correlation coefficients, we test the null-hypothesis that the two time series are in-  
183 dependent at all lags, but that sampling error may lead to a non-zero correlation. We  
184 build empirical distributions for each lag with which to evaluate this null-hypothesis  
185 using 2000 samples formed in the following way: two independent time series of  
186 length 50 years are generated from the AR1 process and the anomaly correlation  
187 coefficient is computed for each lag after low-pass filtering. This approach will nat-

188 urally account for changes in the degrees-of-freedom associated with the lag, the  
189 autocorrelation in the model, and the low-pass filtering. Obtained correlations that  
190 fall above (below) 97.5% (2.5%) of the samples from the empirical distribution at  
191 each lag are considered significant (i.e., statistically unlikely to have resulted from  
192 two uncorrelated time series) at the 95% confidence level.

193 As in DY14, we use the total AMOC transports in our analysis, i.e., the sum  
194 of the Eulerian-mean, mesoscale eddy, and submesoscale eddy contributions, if the  
195 latter two are available. Except INMOM, all models include a variant of the Gent  
196 and McWilliams (1990) parameterization (GM90) to represent the advective effects of  
197 the mesoscale eddies. Only four models (ACCESS, GFDL-GOLD, GFDL-MOM, and  
198 NCAR) employ a submesoscale eddy parameterization (Fox-Kemper et al., 2011) that  
199 contributes to the total transport. We note that in BERGEN the same submesoscale  
200 eddy parameterization is used only to modify the turbulent kinetic energy budget of  
201 the mixed layer model and it does not contribute to the total transport. Because we  
202 are primarily interested in large-scale sub-thermocline (below 500 m) characteristics  
203 of AMOC and the impacts of both the mesoscale and submesoscale eddies are largely  
204 confined to the upper few hundred meters in the North Atlantic, missing subgrid-  
205 scale contributions from some models is not expected to affect our findings. For  
206 convenience, we refer to total AMOC simply as AMOC in the rest of this paper.

207 Furthermore, we primarily use the representation of AMOC in depth–latitude  
208 space in our analysis. While this is the most common depiction and use of AMOC,  
209 an alternative is AMOC in density–latitude space – which we also consider, though  
210 briefly. As discussed in Kwon and Frankignoul (2014), the depth-space AMOC tends  
211 to stress sinking (deep water formation) across isopycnals. In contrast, the density-  
212 space AMOC is better at highlighting water mass transformations and, perhaps, at  
213 exposing the impacts of upper-ocean subpolar gyre in the North Atlantic. Zhang

214 (2010) also argues that the density-space AMOC better represents the meridional  
215 coherency of AMOC variability. Given that the information provided by either rep-  
216 resentation will likely be model dependent, both representations may be used to pro-  
217 vide complementary analysis for detailed variability mechanism studies (see Kwon  
218 and Frankignoul, 2014).

### 219 **3. AMOC variability**

220 We start with the AMOC maximum transport time series at  $26.5^\circ$  and  $45^\circ$ N shown  
221 in Figs. 1 and 2, respectively. The time series are based on AMOC obtained in depth  
222 – latitude space. They are anomalies from the respective 50-year (1958–2007) means  
223 for each model: these means are given in parentheses next to the model labels in each  
224 figure and they are also listed in Table 2. The MMM time series are included in the  
225 figures. These two latitudes are chosen to represent low- and mid-latitude AMOC  
226 variability, respectively. The  $26.5^\circ$ N time series additionally permit a comparison  
227 of models' AMOC variability to that of the RAPID based estimates during a short  
228 overlap period.

229 Focusing on decadal and longer time scales at both latitudes, AMOC variability,  
230 in general, can be characterized in three stages. During the first stage that lasts until  
231 the mid- to late-1970s, AMOC is relatively steady, usually remaining weaker than its  
232 long-term mean. Thereafter, AMOC intensifies with maximum transports achieved  
233 in the mid- to late-1990s. This intensification is then followed by a weakening trend  
234 that continues until the end of our integration period. Maximum transports appear  
235 to occur earlier and the weakening trend appears to be more pronounced at  $45^\circ$ N  
236 than at  $26.5^\circ$ N. Unfortunately, there are no long-term continuous observations to  
237 verify this *general* AMOC behavior in our CORE-II simulations. However, many  
238 modeling studies discussed in section 1 corroborate the AMOC variability depicted

239 in Figs. 1 and 2 (e.g., Häkkinen, 1999; Eden and Willebrand, 2001; Bentsen et al.,  
240 2004; Beismann and Barnier, 2004; Böning et al., 2006; Deshayes and Frankignoul,  
241 2008; Lohmann et al., 2009b; Brodeau et al., 2010; Robson et al., 2012). Similar  
242 trend behavior is also seen in some reanalysis products (Pohlmann et al., 2013).

243 There are, however, exceptions to the above generalizations. For example, CMCC,  
244 FSU, MIT, MRI-A, and NOCS show either very weak or no noticeable trends during  
245 the 1958–2007 period; KIEL does not show weakening during the last decade at  
246 26.5°N; and ICTP time series appear quite different than the other models at 45°N.  
247 There are also differences among the models in their ranges of anomaly magnitudes  
248 with AWI and GISS showing the largest peak-to-peak ranges with about 7 and 9 Sv  
249 ( $1 \text{ Sv} \equiv 10^6 \text{ m}^3 \text{ s}^{-1}$ ), respectively, at both latitudes. Nevertheless, the level of gen-  
250 eral agreement in the characteristics of the AMOC maximum transport time series,  
251 e.g., year-to-year variability and long-term trends, among the forward (non-data-  
252 assimilating) models participating in this study appears to be substantially greater  
253 than among various reanalysis products shown in Karspeck et al. (2015).

254 We provide a more quantitative assessment of the agreements and disagreements  
255 among the models in their representations of AMOC variability in Fig. 3, considering  
256 model – model correlations of the AMOC maximum transport time series discussed  
257 above. Specifically, the figure shows the high-pass filtered; low-pass filtered with  
258 trend; and low-pass filtered but linearly detrended time series correlations between  
259 the models. The majority of the models are in agreement in their representations  
260 of inter-annual variability at both latitudes (Figs. 3a and 3d). In general, model –  
261 model correlations are weaker at 45°N than at 26.5°N. MRI-A is the major outlier at  
262 26.5°N, with ACCESS, ICTP, and INMOM also showing less agreement. ICTP has  
263 the lowest correlations at 45°N. Figures 3b and 3e indicate that the model – model  
264 correlations are much weaker at decadal and longer time scales than at inter-annual

265 time scales. Again, the disagreement among the models is larger at 45°N than at  
266 26.5°N. At both latitudes, the primary outliers are MRI-A and NOCS with most  
267 of their correlation coefficients much less than 0.5. A comparison of the low-pass  
268 filtered correlations with trend and with the linear trend removed (Figs. 3b and 3e  
269 vs. Figs. 3c and 3f) shows that on decadal time scales the trend is the dominant  
270 signal over the 1958–2007 period at both latitudes – but more evident at 26.5°N.  
271 We note that although MRI-A emerges as an outlier when compared to the forward  
272 models in its representation of several AMOC variability characteristics considered  
273 in this paper, it is not an outlier among the reanalysis products analyzed in Karspeck  
274 et al. (2015).

275 The general characteristics of AMOC variability described above with reference  
276 to Fig. 3 appear to be consistent with findings of some previous studies (e.g., Bias-  
277 toch et al., 2008; Yeager and Danabasoglu, 2014). On inter-annual time scales and  
278 particularly at lower latitudes, variability is primarily wind-driven as suggested by  
279 the strong model – model correlations of Figs. 3a and 3d. Such high model – model  
280 correlations from the wind-driven component are expected because all the models  
281 are forced by the same wind dataset. On decadal and longer time scales, variabil-  
282 ity is dominated by buoyancy forcing, and there are larger discrepancies among the  
283 models. These latter differences are likely associated with differences in the models’  
284 DWF properties.

285 Figures 4 and 5 show the AMOC EOF1 spatial distributions and the correspond-  
286 ing PC1 time series, respectively, based on the depth–latitude space AMOC. Because  
287 we use undetrended time series, the patterns depicted in Fig. 4 are primarily asso-  
288 ciated with low frequency variability and trends, and PC1 time series are broadly  
289 similar to those of Figs. 1 and 2. Thus, most of the time series show strengthening  
290 of initially weak AMOC until about the mid- to late-1990s, followed by a weakening

291 trend. The exceptions to this generalization include FSU and, in particular, MRI-A.  
292 In general, the EOF1 distributions display a single cell pattern, covering the Atlantic  
293 basin south of 60°N. GISS, ICTP, and MRI-A have the largest amplitudes with more  
294 than 3.2 Sv per standard deviation. In this EOF measure, AWI does not stand out  
295 as one of the models with a large amplitude.

296 Based on their EOF1 spatial patterns, the models can be separated into three  
297 distinct groups. The first group, representing the majority with twelve models, has  
298 their maxima in the Northern Hemisphere, mostly between 30° and 50°N. The models  
299 in this group are AWI, BERGEN, CERFACS, CNRM, GFDL-GOLD, GFDL-MOM,  
300 GISS, GISS2, ICTP, INMOM, KIEL, and NCAR. Particularly for these models, the  
301 EOF1 pattern in its positive phase indicates strengthening and deeper penetration  
302 of the North Atlantic Deep Water (NADW) cell. The second group of models, i.e.,  
303 ACCESS, CMCC, FSU, FSU2, MIT, MRI-F, and NOCS, have their maxima in the  
304 Southern Hemisphere. With the exception of ACCESS, these are among the models  
305 with the weakest mean AMOC transports as shown in Fig. 3 of DY14, Fig. 17,  
306 and Table 2. MRI-A is the only member of the third group with its maximum  
307 located in the vicinity of the equator. Whether the AMOC EOF1 maximum is  
308 located in the Northern or Southern Hemisphere does not appear to be related to the  
309 characteristics / properties of the Southern Ocean meridional overturning circulations  
310 in these CORE-II simulations in any obvious way (see Farneti et al., 2015).

311 With its equatorially-enhanced EOF1 spatial structure and associated PC1 time  
312 series, MRI-A is one of the models with large differences from the general AMOC  
313 behavior described earlier. Similar, large amplitude AMOC variability at or near  
314 the equator is also present in other reanalysis products as shown in Karspeck et al.  
315 (2015). We think that such a prominent feature in reanalysis products, including  
316 MRI-A, may be associated with the mismatches in calendar time between the zonal



317 wind stress used to force the model and the potential temperature and salinity data  
318 used in data assimilation. This is because the equatorial circulation represents a  
319 balance between the zonal wind stress and zonal pressure gradients and any small  
320 discrepancies in this balance can produce anomalous circulation patterns. Thus, we  
321 believe that the MRI-A EOF1 likely represents spurious variability.

322 The EOF1s account for 40% to 70% of the total variances in AMOC. The highest  
323 variances occur in BERGEN (70%), ICTP (71%), GFDL-GOLD (74%), GISS (74%),  
324 and KIEL (77%). All of these models have their maxima in the Northern Hemisphere.  
325 In contrast, the models with the lowest variances, i.e., FSU2(40%), MRI-A (40%),  
326 FSU (46%), MIT (46%), and NOCS (47%), have their maxima in the Southern  
327 Hemisphere or near the equator.

328 For comparison purposes, we note that the second EOFs of AMOC (not shown)  
329 account for only 7% to 22% of the total variance with fourteen models having vari-  
330 ances of  $< 15\%$ . Not surprisingly, the models with the larger EOF2 variances corre-  
331 spond to the ones with the smallest variances in their EOF1s. With the exception  
332 of a few models, the EOF2 spatial patterns can be described as two north – south  
333 counter-rotating (dipole) cells, extending from the surface to the ocean bottom (not  
334 shown, but see Fig. 2 of Danabasoglu et al. (2012) for an example). The crossover  
335 latitude between these two cells varies between  $0^\circ$  and  $30^\circ\text{N}$  among the models, but  
336 it is near  $0^\circ$  in the models with the largest EOF2 variances. These models are also  
337 the ones with their EOF1 maxima in the Southern Hemisphere.

338 We find qualitatively very similar results when AMOC variability is analyzed in  
339 density ( $\sigma_2$ ) – latitude space as presented in Fig. 6. For example, relative model dif-  
340 ferences are largely preserved, with models which have weaker (stronger) AMOC am-  
341 plitudes in depth space still showing weaker (stronger) amplitudes in density space.  
342 In addition to GISS and ICTP, AWI, GFDL-MOM, and KIEL also show variability

343 > 3.2 Sv per standard deviation. All of the models with their maximum variability  
344 in the Northern Hemisphere in depth-space AMOC also retain their maxima in the  
345 same hemisphere, but the latitudes of the maxima are shifted northwards in density  
346 space. MRI-A has its maximum still near the equator. The models with their max-  
347 ima in the Southern Hemisphere in depth space display less consistency in density  
348 space. For example, while CMCC and FSU have their maxima still in the Southern  
349 Hemisphere, the location of maxima is shifted to the Northern Hemisphere in FSU2  
350 and MRI-F. In ACCESS, there is an additional maximum location in the Northern  
351 Hemisphere.

352 The density-space EOF1s account for 35% to 72% of the total variance in AMOC  
353 – a very similar spread as in the depth-space analysis. However, the individual  
354 model variances are reduced in density space in all models with the exception of  
355 AWI, FSU, and GFDL-MOM. The lowest variances occur in MIT (35%), NOCS  
356 (36%), FSU2 (37%), and MRI-A (37%) – all among the lowest-variance-models in  
357 depth space as well. GFDL-MOM and KIEL have the highest variances with > 70%.  
358 The corresponding PC1 time series (not shown) are very similar to those of Fig. 5  
359 for AMOC in depth space, broadly duplicating the low frequency AMOC variability.

360 As illustrated above, the most prominent features of the AMOC maximum trans-  
361 port and PC1 time series are the strengthening of transports between about the  
362 mid-1970s and the mid- to late-1990s, followed by a weakening trend that continues  
363 until the end of the integration period. To provide a quantitative assessment of these  
364 tendencies, we present the AMOC linear trends in Table 2, calculated using the time  
365 series of Figs. 1 and 2 for 26.5° and 45°N, respectively. The trends are calculated  
366 for the 1978–1998 and 1998–2007 periods at 26.5°N and for the 1975–1995 and  
367 1995–2007 periods at 45°N to roughly represent the time frames with increases and  
368 decreases in AMOC, respectively. The shifts in the time periods between the two

369 latitudes are intended to account for the apparent lag of AMOC changes at 26.5°N  
370 in comparison to those at 45°N as alluded to earlier in this section (also see section  
371 9). The trends that meet the 95% confidence level based on a two-sided Student's  
372 t-test are shown in bold.

373 We compute the MMM trends as 0.70 and  $-1.73 \text{ Sv decade}^{-1}$  at 26.5°N and  
374 0.82 and  $-1.54 \text{ Sv decade}^{-1}$  at 45°N. Particularly for the later period, these trends  
375 are impacted by the large negative trends in GISS, the exclusion of which reduces  
376 the MMM trends at both latitudes to about  $-1.37 \text{ Sv decade}^{-1}$ . A notable feature  
377 of the MMM trends is that the weakening rate is nearly double that of strength-  
378 ening. We note that the models that have their AMOC maximum variability in  
379 the Southern Hemisphere or in the vicinity of the equator tend to show weaker and  
380 statistically less significant trends. For the 1978–1998 period at 26.5°N, all mod-  
381 els show positive trends, ranging from 0.09 (MRI-A) to 1.62 (GISS)  $\text{Sv decade}^{-1}$ .  
382 For the 1998–2007 period at the same latitude, while sixteen models have negative  
383 trends – from  $-0.02$  (FSU) to  $-8.13$  (GISS)  $\text{Sv decade}^{-1}$  – four models, i.e., KIEL,  
384 MIT, MRI-A, and NOCS, show positive trends. Except KIEL, these models have  
385 their maximum AMOC anomalies in the Southern Hemisphere. At 45°N, the models  
386 are again unanimous in their trend signs, all showing AMOC intensification for the  
387 1975–1995 period, ranging from 0.03 (NOCS) to 2.06 (GISS)  $\text{Sv decade}^{-1}$ . For the  
388 1995–2007 period at 45°N, all but three models show weakening of AMOC with  
389 trends ranging from  $-0.02$  (FSU) to  $-4.81$  (GISS)  $\text{Sv decade}^{-1}$ . The exceptions are  
390 MIT, MRI-F, and NOCS, again all with maximum AMOC anomalies in the Southern  
391 Hemisphere.

392 We make the following additional observations based on Table 2: GISS emerges  
393 as the model with the largest trends of both signs at both latitudes; the trends in  
394 NOCS are positive at both latitudes regardless of the time period; and MRI-A and

395 NOCS are the only models in which all trends remain below our confidence level.

396 The spatial patterns of AMOC linear trends are very similar to those of the  
397 EOF1s depicted in Fig. 4 and, therefore, not shown. The intensification of AMOC  
398 during the earlier period is associated with strengthening and deeper penetration  
399 of the NADW cell. We finally note that MRI-A appears to be an outlier in its  
400 trend spatial patterns (not shown), revealing strong negative trends in the Southern  
401 Hemisphere. Regarding such reanalysis products, Karspeck et al. (2015) show quite  
402 diverse representations of AMOC trends over a similar time period among several  
403 reanalysis datasets – perhaps even more diverse than those depicted in Table 2 for  
404 the present CORE-II simulations.

#### 405 4. SST variability

406 An important test for evaluation of the CORE-II hindcast simulations is their  
407 ability to reproduce observed spatial patterns and temporal characteristics of SST  
408 variability. This is not assured in these simulations as discussed in Doney et al. (2007)  
409 where it is shown that ocean processes considerably affect SST and upper-ocean heat  
410 content variability. Thus, disagreements with observations can be expected in the  
411 North Atlantic where ocean, particularly advective, heat transports are significant.

412 We show the model SST EOF1 spatial distributions and the associated PC1 time  
413 series in Figs. 7 and 8, respectively, including the corresponding distributions from  
414 the HadISST observational dataset (Hurrell et al., 2008). The EOFs are obtained  
415 for the North Atlantic region bounded by  $80^{\circ}\text{W}$ – $10^{\circ}\text{E}$  and  $10^{\circ}$  –  $70^{\circ}\text{N}$ . Although  
416 we do not use any detrending or low-pass filtering of the SST time series, the EOF1  
417 patterns still produce the familiar AMV pattern (e.g., Sutton and Hodson, 2005)  
418 with a basin scale, single-sign SST anomaly (positive in Fig. 7). In HadISST, the  
419 maximum variability occurs east of Newfoundland with an amplitude of  $> 0.7^{\circ}\text{C}$  per

420 standard deviation. Additional maxima are present in LS and western Irminger Sea.  
421 There is an opposite-signed anomaly (negative in Fig. 7) just off the east coast of  
422 North America with a small amplitude of about  $0.1^{\circ}\text{C}$  per standard deviation. The  
423 CORE-II simulations broadly reproduce observed SST characteristics, but there are  
424 many differences from observations in details. Perhaps the most visible of these is the  
425 amplitude, location, and spatial extent of the largest SST anomaly. This discrepancy  
426 is particularly evident in AWI, CERFACS, CNRM, GFDL-GOLD, GFDL-MOM,  
427 GISS, GISS2, and KIEL with maximum anomalies of  $> 1^{\circ}\text{C}$  per standard deviation  
428 and with substantially broader spatial extent of this maximum in comparison with  
429 observations. We think that these discrepancies together with somewhat smaller  
430 differences in the details of the negative SST anomalies off the east coast of North  
431 America are partly due to the incorrect separation of the models' Gulf Stream and  
432 the failure of the subsequent North Atlantic Current (NAC) to reconnect with the  
433 topography off the Grand Banks, resulting in a too-zonal path. As discussed in  
434 Danabasoglu (2008), these persistent biases can impact model variability in the North  
435 Atlantic. The model SST EOF1s account for 29% to 40% of the total SST variance  
436 in good agreement with the observational variance of 40%.

437 The correspondence between the model – model and model – observational SST  
438 PC1 time series is remarkably good, both at inter-annual and decadal time scales  
439 (Fig. 8). The PC1 time series show large amplitude low frequency variability su-  
440 perimposed onto inter-annual changes. There is an evident warming trend, roughly  
441 between the late 1980s and the late 1990s, producing peak SSTs around 1998. An-  
442 other peak occurs during 2005–2006 after a short-lived cooling in between. Good  
443 agreements among all of these PC1 time series, particularly with modeled and ob-  
444 served variability, indicate that the temporal character of the basin-scale SST is  
445 primarily dictated by the variability and trends in the atmospheric datasets which

446 already include the impacts of ocean dynamics from nature superimposed onto an-  
447 thropogenic effects. The role of the *simulated* ocean dynamics in the models, e.g.,  
448 NAC and AMOC, in influencing smaller-scale SSTs and upper-ocean heat contents  
449 is demonstrated by the differences among the models in their SST EOF1 spatial  
450 structures (Fig. 7). Indeed, the role of enhanced AMOC transports in the con-  
451 text of the mid-1990s subpolar North Atlantic warming in the upper-ocean has been  
452 unequivocally shown in Robson et al. (2012) and Yeager et al. (2012).

## 453 5. MLD variability

454 We assess the variability of the models' DWF regions in the northern North At-  
455 lantic, considering the March-mean MLD time series. Following the same procedure  
456 as in DY14, we adopt a density-based approach to determine MLDs where they  
457 are calculated as the depths at which the potential density (referenced to surface)  
458 changes by  $0.125 \text{ kg m}^{-3}$  from its surface value. MLD is calculated offline using the  
459 March-mean potential density obtained from March-mean potential temperature and  
460 salinity by each participating group.

461 Figures 9 and 10 present the March-mean MLD EOF1 spatial distributions and  
462 the PC1 time series, respectively. Despite differences in their mean MLDs (see Fig. 13  
463 of DY14), the majority of the models show the area extending from the southeast  
464 LS into the Irminger Sea as the region with the largest MLD variability. Such broad  
465 regions with deep MLDs appear to be rather extensive in comparison with some  
466 observations (e.g., Lavender et al., 2002) which show only relatively small areas of  
467 deep mixing, mostly confined to just north of Labrador. In more than half of the  
468 models, the maximum amplitude is  $> 800 \text{ m}$  per standard deviation. However, the  
469 amplitude and spatial extent of the maximum MLD variability in the LS – Irminger  
470 Sea region differ considerably among the models. There are three exceptions to the

471 dominance of this region: in KIEL and MRI-F, the MLD variability is as strong in  
472 the Nordic Seas; and NOCS has its largest variability in the Nordic Seas with rather  
473 weak variability in the LS region. In some of the models, e.g., BERGEN, CERFACS,  
474 CMCC, GFDL-MOM, and, NCAR, the deeper MLDs in the LS region – as depicted  
475 in Fig. 9 – are accompanied by shallower MLDs in the northern LS. Small amplitude  
476 negative MLD anomalies are also evident in the Nordic Seas in CERFACS, CMCC,  
477 CNRM, GFDL-MOM, GISS, GISS2, and MRI-A. The interior white areas in Fig. 9  
478 indicate regions of no variability as the time-mean MLDs reach the ocean bottom  
479 in some models. A prominent example is ICTP where the time-mean MLDs are  
480 always as deep as the ocean bottom. The MLD EOF1s account for 19% to 49%  
481 of the total variance in MLD. While BERGEN (40%) GFDL-GOLD (41%), GFDL-  
482 MOM (41%), and NCAR (49%) have the highest variances, INMOM and NOCS  
483 have the smallest variances with 19% each. We note that larger (smaller) MLD  
484 EOF1 variances do not imply similarly larger (smaller) AMOC EOF1 variances. For  
485 example, MLD EOF1 variances are very similar for CMCC, ICTP, AWI, ACCESS,  
486 and MIT (35 – 37%), but their AMOC EOF1 variances range from 46% in MIT to  
487 71% in ICTP. Likewise, we do not find any obvious connections between the MLD  
488 EOF1 spatial pattern characteristics and where the AMOC EOF1 maxima occur,  
489 i.e., Southern vs. Northern Hemisphere.

490 The PC1 time series (Fig. 10) show general agreement among most of the mod-  
491 els, particularly in their representations of low frequency variability. With the sign  
492 convention depicted in Fig. 9 and primarily referring to the LS MLDs, MLDs get  
493 shallower and stay shallower during the first decade. This is followed by a tendency  
494 towards deeper MLDs until the early- to mid-1990s. Finally, we identify a tendency  
495 towards shallower MLDs till the end of the integration period. This characterization  
496 of the time series is consistent with changes in AMOC and is discussed further in

497 section 9. The exceptions to the generalization include: NOCS with its near-neutral  
498 MLDs between the late 1980s and the late 1990s; KIEL, MIT, MRI-F, and NOCS  
499 with their mostly positive MLD anomalies after 1998; and GISS with a sharper  
500 increase and a sharper decrease of MLDs in the early 1970s and the early 1990s,  
501 respectively. We note that KIEL and NOCS deviate significantly from CERFACS,  
502 CMCC, and CNRM – the other NEMO-based models – in their PC1 time series,  
503 particularly after the mid-1980s.

## 504 **6. Comparisons with hydrographic data in central LS**

505 Unfortunately, it is rather difficult to verify the fidelity of the simulated MLD  
506 variability in the northern North Atlantic discussed above due to very limited ob-  
507 servations. Instead, following Yeager and Danabasoglu (2014), we focus on a small  
508 central LS region, taking advantage of a compilation of hydrographic observations  
509 from Yashayaev (2007) which includes data from research vessels and profiling Argo  
510 floats. Specifically, we generate time series of potential temperature ( $\theta$ ), salinity ( $S$ ),  
511 and density ( $\sigma_0$ ) by averaging over a region bounded by  $49^\circ - 56^\circ\text{W}$  and  $56^\circ - 61^\circ\text{N}$ .  
512 We compute vertical averages in depth coordinates, rather than in density coordi-  
513 nates, for the 150–1000 m depth range because the observations are available at  
514 depths greater than 150 m. We use May-mean  $\theta$  and  $S$  from the models to roughly  
515 match the mostly Spring-time observations. Density is calculated using a common  
516 equation of state for all models, based on these May-mean  $\theta$  and  $S$ .

517 We present the resulting model and observational time series for  $\theta$ ,  $S$ , and density  
518 in Figs. 11, 12, and 13, respectively, as anomalies from the 1958–2007 period. For  
519 this comparison, the data from fourteen of the participating models are available.  
520 Also, the observational data are missing for some years roughly between 1975 and  
521 1990. The figures also include the root-mean-square (rms) model – observations time



522 series differences as well as the correlation coefficients between the model and obser-  
523 vational time series for each model. These two metrics are evaluated only for years  
524 with available observations and, as such, they are less focused on the 1975–1990  
525 period where missing data occurs. We note that low rms differences and high corre-  
526 lation coefficients indicate good agreements with observations.

527 The observations show decadal-scale variability in  $\theta$  and  $S$  from warm and salty  
528 anomalies in the 1960s and the early 1970s to mostly cold and fresh anomalies un-  
529 til about the early 2000s and then back to warm and salty anomalies. There are  
530 substantial compensations of  $\theta$  and  $S$  anomalies in their contributions to density,  
531 but the density anomalies between about 1985 and 2000 are set primarily by the  $\theta$   
532 anomalies. The largest positive density anomalies occur in the mid-1990s, roughly  
533 coinciding with the deepest MLDs. There is modest agreement between the observa-  
534 tional and simulated decadal-scale variability, particularly evident in  $\theta$  and density  
535 time series. We compute the MMM correlation coefficients, i.e., the mean of the  
536 correlation coefficients and excluding MRI-A, for  $\theta$  and density as 0.58 and 0.61,  
537 respectively. The corresponding value for  $S$  is much lower at 0.26. We note that  
538 MRI-A, which assimilates data, usually has the lowest rms and the highest cor-  
539 relation coefficients, producing one of the better agreements with observations by  
540 construction. Therefore, in the following discussion, we focus our attention to the  
541 performance of the forward models.

542 In  $\theta$  (Fig. 11), while the smallest rms differences are in INMOM (0.19°C) and  
543 CERFACS (0.20°C), the largest departures from observations are in NCAR (0.31°C),  
544 BERGEN (0.32°C), and AWI (0.36°C). ICTP has the lowest correlation coefficient  
545 with 0.28. INMOM and CERFACS show the highest correlations with 0.76 and  
546 0.79, respectively. Thus, in these measures, CERFACS and INMOM have the best  
547 agreements with observations. We note that, with the exception of NOCS, all models

548 exhibit a prominent cold bias that leads to a positive density bias roughly during the  
549 1983–1985 period. Because such a cold bias also exists in all the reanalysis products  
550 analyzed in Karspeck et al. (2015), we speculate that it may indicate a deficiency  
551 with the observational data.

552 In  $S$  (Fig. 12), the models with the lowest and highest rms differences are CNRM  
553 (0.025 psu), CERFACS (0.026 psu), CMCC (0.028 psu) and GFDL-MOM (0.040  
554 psu), INMOM (0.042 psu), ICTP (0.047 psu), respectively. As indicated above, the  
555 correlation coefficients for  $S$  are much lower than those of  $\theta$  and density. Indeed, the  
556 correlation is even negative in ICTP ( $-0.06$ ) and near-zero in three of the models,  
557 i.e., GFDL-GOLD (0.00), INMOM (0.04), and GFDL-MOM (0.06). The highest  
558 correlation occurs in AWI with only 0.50. Although these metrics do not favor  
559 a particular model as better than the others, ICTP, INMOM, and GFDL-MOM  
560 produce the largest departures from observations.

561 While the largest rms density differences (Fig. 13) occur in AWI ( $0.026 \text{ kg m}^{-3}$ ),  
562 GFDL-MOM ( $0.026 \text{ kg m}^{-3}$ ), and INMOM ( $0.033 \text{ kg m}^{-3}$ ), the lowest rms differences  
563 are in FSU2 ( $0.014 \text{ kg m}^{-3}$ ) and CMCC ( $0.016 \text{ kg m}^{-3}$ ). The smallest correlations are  
564 in MRI-F, NOCS, and AWI with 0.48, 0.53, and 0.54, respectively. CMCC and FSU2  
565 reveal the highest correlation coefficients with 0.69 and 0.73, respectively. Thus, these  
566 two models emerge as the models with the best agreements with the observations in  
567 density – even better than in MRI-A. In contrast, AWI appears to show the least  
568 agreement. As indicated earlier, the density time series include compensating biases  
569 in  $\theta$  and  $S$  in their contributions to density. A notable example of this compensation  
570 occurs after 1998 where most models show warm and salty biases.

571 Finally, we compute the linear trends in density for the 1970–1995 period for  
572 each model and for the observations as another evaluation metric. The MMM trend  
573 of  $0.025 \text{ kg m}^{-3} \text{ decade}^{-1}$  compares rather favorably with the observationally-based

574 trend of  $0.024 \text{ kg m}^{-3} \text{ decade}^{-1}$ . The range for individual model trends is between  
575  $0.009$  and  $0.049 \text{ kg m}^{-3} \text{ decade}^{-1}$  with NOCS and INMOM at the low and high  
576 end of this range, respectively. The simulated trends are within 20% of the obser-  
577 vational value in six of the models. These models are (with their trends in  $\text{kg m}^{-3}$   
578  $\text{decade}^{-1}$ ) AWI (0.020), NCAR (0.022), BERGEN (0.022), CMCC (0.023), GFDL-  
579 MOM (0.027), and FSU2 (0.028). We note that the trend in MRI-A is  $0.017 \text{ kg m}^{-3}$   
580  $\text{decade}^{-1}$ .

## 581 **7. Gyre and sea surface height variability in the subpolar North Atlantic**

582 Several recent observational and modeling studies highlight the importance and  
583 impacts of the North Atlantic SPG circulation variability on the climate of the North  
584 Atlantic (e.g., Häkkinen and Rhines, 2004; Böning et al., 2006; Lohmann et al., 2009a;  
585 Yeager and Danabasoglu, 2014). Because the SPG transport itself is not easily ob-  
586 served, the satellite-based SSH data (available since 1993) is used instead to deter-  
587 mine observed changes in the SPG as well as to evaluate model-based findings (since  
588 the strength of the SPG is directly connected to the SSH gradients via geostrophy).  
589 As discussed earlier, the previous studies also show that there is a close connection  
590 between the SPG / SSH variability and that of AMOC via the NAO-related surface  
591 fluxes and associated changes in DWF, i.e., convective events. Indeed, Yeager and  
592 Danabasoglu (2014) suggest monitoring of the variations in the LS SSH as a proxy  
593 for AMOC changes.

594 A detailed evaluation of the simulated, global sea level mean and variability  
595 for the 1993–2007 period for most of the models participating in CORE-II is pre-  
596 sented in Griffies et al. (2014). In the present study, we specifically focus on the  
597 SSH – strictly speaking, dynamic sea level – changes in the SPG region defined as  
598 the area between  $15^\circ - 60^\circ\text{W}$  and  $48^\circ - 65^\circ\text{N}$  to provide an assessment of fidelity

599 of model simulations in this important metric in comparison with the data from  
 600 the AVISO project (Archiving, Validation, and Interpolation of Satellite Oceanographic  
 601 Data; Le Traon et al., 1998; Ducet et al., 2000). Here, we use a product  
 602 available from a NASA Jet Propulsion Laboratory web site located at [podaac.jpl.nasa.gov/dataset/AVISO\\_L4\\_DYN\\_TOPO\\_1DEG\\_1MO](http://podaac.jpl.nasa.gov/dataset/AVISO_L4_DYN_TOPO_1DEG_1MO). The SSH time series  
 603 anomalies calculated as the average SSHs for the SPG region with respect to the  
 604 1993–2007 mean are given in Fig. 14. The AVISO time series are included in each  
 605 panel as the black lines. The figure also shows the correlation coefficients between  
 606 the AVISO and models' time series as well as the linear trends for the 1993–2007  
 607 period for each model and from the AVISO data. NOCS clearly emerges as the major  
 608 outlier in comparison with the AVISO data as the only model with a negative  
 609 correlation coefficient ( $-0.19$ ) and as the only model with a negative trend ( $-0.15$   
 610  $\text{cm yr}^{-1}$ ). Half of the models have quite high correlations with the AVISO data  
 611 with correlation coefficients of 0.96 or higher. The lowest correlations are in MIT  
 612 and KIEL with 0.69 and 0.75, respectively. The trend in AVISO data is  $0.45 \text{ cm yr}^{-1}$ .  
 613 The simulated trends are within 20% of this value in six of the models. These  
 614 models are (with their trends in  $\text{cm yr}^{-1}$ ) GISS2 (0.38), FSU2 (0.39), GFDL-GOLD  
 615 (0.39), NCAR (0.39), GFDL-MOM (0.48), and BERGEN (0.51). The largest trend  
 616 is in ICTP with  $0.62 \text{ cm yr}^{-1}$  which is within 30% of the AVISO-based value. The  
 617 smallest positive trends occur in MIT, MRI-F, CNRM, and KIEL with 0.05, 0.07,  
 618 0.08, and  $0.08 \text{ cm yr}^{-1}$ , respectively.

## 620 8. AMOC and meridional heat transport variability

621 AMOC is the principal contributor to the Atlantic Ocean MHT in both obser-  
 622 vations and model simulations (see, e.g., Böning et al., 2001; Biastoch et al., 2008;  
 623 Johns et al., 2011; Msadek et al., 2013). Here, we assess the relationships between

624 the AMOC variability and that of the MHT by considering their simultaneous cor-  
625 relations and by performing simultaneous regressions of MHT onto AMOC. For this  
626 purpose, we use the AMOC maximum transports and MHT values obtained at  $26.5^{\circ}\text{N}$   
627 for two reasons: i) this latitude is within the range of latitudes for maximum MHTs,  
628 and ii) there are observationally-based estimates from the RAPID data (Johns et al.,  
629 2011). We note again that the overlap period between the model simulations and the  
630 observations is very short: while we analyze the annual-mean data for the 1958–2007  
631 period from the simulations, the observational data are available starting in April  
632 2004 and their analyses usually use 10-day and 30-day means. The implications of  
633 such differences are discussed below.

634 Table 3 summarizes our results. We find that AMOC and MHT variability are  
635 very highly correlated with correlation coefficients of  $\geq 0.9$  in all, but two, of the  
636 models. The lowest correlations occur in INMOM and MRI-A with 0.86. These  
637 high correlations are consistent with the RAPID-based estimate of 0.97. Such good  
638 agreements between the model and RAPID-based AMOC and MHT correlations ap-  
639 pear to be independent of the range of time averaging applied in the calculations.  
640 For example, we obtain similarly high correlations of 0.93 (for 1958–2007) and 0.96  
641 (for 2004–2007) for NCAR when monthly-mean data are used. The regression co-  
642 efficients vary between 0.042 and 0.068  $\text{PW Sv}^{-1}$  with INMOM at the low end and  
643 CMCC, FSU, MRI-F, and NOCS at the high end of this range. We note that the lat-  
644 ter four are among the models where the maximum anomalies in AMOC occur in the  
645 Southern Hemisphere. The model regression coefficients are all smaller than those  
646 of the RAPID-based estimates which are 0.079  $\text{PW Sv}^{-1}$  (Johns et al., 2011) and  
647 0.083  $\text{PW Sv}^{-1}$  (Msadek et al., 2013) obtained using 10-day and 30-day means, re-  
648 spectively, for the April 2004 – October 2007 period and 0.077  $\text{PW Sv}^{-1}$  (W. Johns,  
649 personal communication) obtained using 10-day means from April 2004 to mid-2014.

650 We think that this discrepancy between the model and observationally-based regres-  
651 sions is due to the use of annual-mean vs. 10-day or 30-day mean data in model  
652 vs. observational analysis. Specifically, we get  $0.074 \text{ PW Sv}^{-1}$  (for 1958–2007) and  
653  $0.078 \text{ PW Sv}^{-1}$  (for 2004–2007) for NCAR when monthly-mean data are employed,  
654 both in rather good agreement with the RAPID-based estimates – in contrast with  
655 the annual-mean-based regression coefficient of  $0.062 \text{ PW Sv}^{-1}$ . Similarly, we find  
656 that the RAPID-based regression coefficient reduces to  $0.067 \text{ PW Sv}^{-1}$  when calcu-  
657 lated with annual-mean data for the April 2004 – March 2014 period. The models  
658 are evenly divided in their intercept values with half above zero and half below zero  
659 values. While GISS2 has the highest intercept with  $+0.177 \text{ PW}$ , MRI-F has the  
660 lowest value with  $-0.117 \text{ PW}$ . As discussed in Msadek et al. (2013), the differences  
661 in regression coefficients and in intercept values among the models can be due to  
662 many reasons, and it is beyond the scope of the present study to investigate causes  
663 of these differences in each model. However, following Msadek et al. (2013), we offer  
664 differences in mean AMOC magnitudes; in correlations between AMOC and temper-  
665 ature fluctuations; and in the gyre component contributions and their variability as  
666 possible causes.

## 667 9. Variability relationships between AMOC and other fields

668 In this section, we investigate relationships between the simulated AMOC vari-  
669 ability and those of MLD, SPG circulation, SPG SSH, and NAO. We use the AMOC  
670 maximum transport at  $45^\circ\text{N}$  time series as our primary AMOC index.

671 We first present in Fig. 15 the low-pass filtered, MMM time series of the AMOC  
672 index, March-mean MLD, and SPG barotropic streamfunction (BSF) (top panel),  
673 and the AMOC maximum transport time series at  $26.5^\circ\text{N}$  and SPG SSH (bottom  
674 panel). The top panel also includes a low-pass filtered NAO index, and our primary

675 AMOC index is repeated in the bottom panel. Here, MLD is calculated as an average  
676 for the LS – Irminger Sea region defined as the area between  $15^{\circ}$  –  $60^{\circ}$ W and  
677  $48^{\circ}$  –  $60^{\circ}$ N, thus including the region extending from the southeast LS to the Irminger  
678 Sea which contains the largest MLD variability in the majority of the models (see Fig.  
679 9). The SPG BSF and SSH represent average transport and surface height for the  
680 SPG domain defined in section 7. For NAO, we adopt the winter (December–March)  
681 sea level pressure PC1 time series from the CORE-II data sets as our index. The  
682 NAO index shows a stronger-than-normal subtropical high and a deeper-than-normal  
683 Icelandic low in its positive phase (NAO+). We note that all models are subject to  
684 the same NAO index because it is part of the forcing datasets. All time series are  
685 anomalies with respect to the 1958–2007 period, and shadings denote one standard  
686 deviation spreads of the models’ time series from those of the respective MMM.

687 The figure shows several noteworthy features. First, changes in MLD tend to lead  
688 changes in AMOC. This is particularly evident after 1980: deepening in MLD leads  
689 AMOC intensification by a few years with the deepest MLDs and the largest AMOC  
690 transports occurring in 1992–1993 and 1995, respectively. Second, the NAO time  
691 series similarly lead those of AMOC, with changes in NAO and MLD tending to co-  
692 vary. There is a suggestion that NAO slightly leads MLD after about 1990. Third,  
693 AMOC and SPG BSF and SSH anomalies appear to be largely in-phase, noting  
694 that the negative BSF and SSH anomalies indicate strengthening of the cyclonic  
695 SPG circulation. However, the SPG SSH time series suggest that they tend to lead  
696 those of AMOC by a few years. In Yeager (2015), these co-variations of AMOC  
697 and SPG anomalies are shown to be associated with the bottom pressure torque  
698 which emerges as the primary driver in the barotropic vorticity equation responsible  
699 for decadal, buoyancy-forced changes in the gyre circulation, thus providing AMOC  
700 and SPG coupling. Finally, we note that the two AMOC time series do not show

701 an appreciable lead–lag relationship until about 1985. Thereafter, anomalies at  
702  $45^{\circ}\text{N}$  lead those at  $26.5^{\circ}\text{N}$  by about 5 years. A prominent example is the emergence  
703 and strengthening of positive AMOC anomalies at  $26.5^{\circ}\text{N}$  during the 1989–2000  
704 period which follow a similar AMOC intensification at  $45^{\circ}\text{N}$  that occurs during the  
705 1984–1995 period.

706 To establish the lead–lag relationships between the AMOC index time series and  
707 those of the MLD, SPG BSF, SPG SSH, and NAO, we next calculate the correlation  
708 functions among these time series. The resulting lead–lag correlations for each model  
709 are shown in Fig. 16 where the AMOC index leads for positive lags. The correlations  
710 are obtained using the low-pass filtered anomalies with respect to the 1958–2007  
711 period. The figure also includes the MMM correlation function evaluated as the  
712 mean of the individual model correlations as well as 95% confidence levels calculated  
713 using a parametric bootstrap method (see section 2 for details). As above, MLD and  
714 BSF time series are evaluated as spatial averages for their respective regions, and  
715 SSH spatial averages use the same domain as in BSF.

716 We first summarize our analysis considering the MMM correlations shown as the  
717 black lines in Fig. 16. The maximum correlations ( $\approx 0.75$ ) occur when positive MLD  
718 anomalies, i.e., MLD deepening, lead AMOC intensification by 2–3 years. As also  
719 suggested by Fig. 15, the correlation coefficient between the AMOC index and the  
720 SPG BSF time series is a maximum ( $\approx |0.7|$ ) at lag of  $-1$  to  $-2$ , again noting that the  
721 negative correlations indicate in-phase strengthening and weakening of AMOC and  
722 SPG. We see a similar relationship between the AMOC index and the SPG SSH time  
723 series with the largest negative correlations of about 0.6 occurring when SSH leads  
724 by 2–3 years. These lead–lag relationships between the AMOC index time series  
725 and those of SPG BSF and SSH along with the time series plots of Fig. 15 support  
726 the idea of monitoring the variations in the LS SSH as a proxy for AMOC changes as



727 suggested by Yeager and Danabasoglu (2014). Lastly, we note that the NAO index  
728 leads the AMOC index by 2–4 years with a maximum correlation coefficient of about  
729 0.6.

730 There are many differences among the individual correlation functions, for ex-  
731 ample, in their correlation coefficient magnitudes as well as in their lead–lag times  
732 for maximum correlations. We discuss only a few of these differences here both to  
733 provide some examples of such differences and to identify some models that depart  
734 from our MMM characterization. Starting with the AMOC and MLD correlation  
735 functions, we note that although INMOM also shows relatively strong correlations  
736 when MLD leads AMOC, it is the only model which has its maximum correlation  
737 when AMOC leads, indicating that MLDs continue to get deeper while AMOC be-  
738 gins to weaken. The maximum correlations vary between about 0.45 and 0.9 among  
739 the models, with ICTP at the low end and AWI, BERGEN, CNRM, INMOM, KIEL,  
740 MRI-F, and NCAR at the high end of this range. The low correlations in ICTP that  
741 are not statistically significant are likely due to low MLD variability in the LS –  
742 Irminger Sea region (Fig. 9) where the time-mean MLDs always remain very deep  
743 and the largest variabilities occur in the southern portion. In contrast with the rest  
744 of the models, GFDL-GOLD, GISS, MRI-A, and NOCS show earlier transitions to  
745 negative correlations starting at lag of 0. Consequently, these models have the largest  
746 negative correlation coefficients among the models. Although there does not seem to  
747 exist any clear relationships between the AMOC – MLD correlations and where the  
748 deepest MLDs occur in the models, we note that in MRI-A and NOCS – two of the  
749 models with earlier transitions to negative correlations – AMOC EOF1 anomalies  
750 are very weak at 45°N, indeed negative as shown in Fig. 4. Continuing with the  
751 AMOC and SPG BSF correlation functions, we find GISS2 and, to some degree,  
752 FSU distributions – both below the confidence levels – difficult to interpret due to

753 their pronounced oscillatory behavior with relatively small correlation coefficients.  
754 In BERGEN, INMOM, and NCAR, the extrema in SPG transports are attained  
755 more than 2 years after the extrema in AMOC. Not surprisingly, there are general  
756 similarities in many individual model correlations between the AMOC vs. BSF and  
757 AMOC vs. SSH relationships. Only GFDL-GOLD and CNRM appear to have the  
758 longest lead times for SSH with 9 to 10+ years. Finally considering AMOC and NAO  
759 relationships, we identify MRI-A and NOCS as the major outliers, noting that while  
760 MRI-A is below our confidence limit, the minimum in NOCS is very near the 95%  
761 limit. They have small or even negative correlations prior to an AMOC maximum,  
762 and negative correlations persist through positive lags. As discussed above regarding  
763 AMOC – MLD relationships, this behavior in MRI-A and NOCS is likely related  
764 to the negative AMOC EOF1 anomalies present at the latitude of our AMOC index  
765 (Fig. 4), in contrast with the other models which show positive anomalies. Further,  
766 in MRI-A, data assimilation presumably impacts the relationship between AMOC  
767 and NOA. To the extent that NAO+ plays an important role in driving AMOC  
768 variability through its associated surface fluxes, as discussed previously, the NAO  
769 appears to be not a major factor in influencing AMOC variability in these two mod-  
770 els. We also note that FSU has its largest positive correlations between AMOC and  
771 NAO following an AMOC intensification.

## 772 10. Summary and conclusions

773 We have presented an analysis of the simulated inter-annual to decadal variabil-  
774 ity and trends in the North Atlantic Ocean for the 1958–2007 period from twenty  
775 simulations participating in the CORE-II effort. A major focus has been the represen-  
776 tation of AMOC variability. In addition, we have investigated connections between  
777 AMOC variability and those of some other fields such as NAO, subpolar MLDs, and

778 LS hydrographic properties to elucidate some variability mechanisms. This study  
779 is Part II of our companion paper, DY14, which documents the mean states in the  
780 North Atlantic from the same models, providing a baseline for the present variability  
781 analysis.

782 In general, AMOC variability shows three distinct stages on decadal time scales.  
783 During the first phase that lasts from 1958 until the mid- to late-1970s, AMOC re-  
784 mains weaker than its long-term (1958–2007) mean. Thereafter, AMOC intensifies  
785 with maximum transports achieved in the mid- to late-1990s. This enhancement  
786 is then followed by a weakening trend that continues until the end of our integra-  
787 tion period. This sequence of low frequency AMOC variability cannot be directly  
788 confirmed by observations. However, it is consistent with the results of many other  
789 ocean hindcast simulations (see section 1 for a sampling of references) forced with var-  
790 ious historical atmospheric datasets, including NCEP/NCAR and ECMWF ERA-40  
791 reanalysis products.

792 A prominent and robust feature of the above characterization of the low frequency  
793 variability is the strengthening of AMOC between about the mid-1970s and the mid-  
794 to late-1990s, distinguished by an intensified and deeper-penetrating NADW cell.  
795 Previous studies show that this AMOC intensification is connected to enhanced DWF  
796 and associated mixed layer deepening in the subpolar North Atlantic, particularly in  
797 the LS region, driven by surface buoyancy fluxes and wind stress resulting from the  
798 persistent positive phase of the NAO. Increase in AMOC is then accompanied by  
799 more heat transport into the subpolar North Atlantic, contributing to the warming  
800 observed in the mid-1990s there. Although an in-depth analysis of AMOC variability  
801 mechanisms in the participating models is beyond the scope of the present study, our  
802 results support this variability mechanism. In particular, positive density and MLD  
803 anomalies precede AMOC intensification, and lead–lag relationships show that both

804 MLD and NAO indices lead AMOC enhancement by 2–4 years. Such a variability  
805 mechanism that suggests an important role for the NAO appears to be very similar  
806 to AMOC intrinsic variability mechanisms found in some CGCM control simulations  
807 (e.g., Danabasoglu et al., 2012).

808 The analysis of the mean states presented in DY14 shows that the larger AMOC  
809 mean transports are associated with deeper MLDs, resulting from increased salt con-  
810 tent in the LS region. In sharp contrast, the increase in AMOC, i.e., the positive  
811 AMOC anomaly, discussed above is primarily associated with negative temperature  
812 anomalies in the LS region in both model simulations and in observations (see also  
813 Yeager and Danabasoglu, 2014). Concerning any links between the Nordic Seas over-  
814 flow transports and AMOC, DY14 finds no clear links between the mean AMOC and  
815 overflow transports. Unfortunately, an investigation of this relationship for variabil-  
816 ity purposes remains beyond the scope of the present study, requiring a dedicated  
817 effort of its own with additional model outputs that are not currently available.

818 Arguably, the level of general agreement in the representation of AMOC variabil-  
819 ity, including year-to-year changes and long-term trends, among the forward models  
820 participating in CORE-II appears to be substantially greater than among various  
821 reanalysis products (Karspeck et al., 2015). Such a general agreement among the  
822 models also extends to characterization of MLD and SSH variability in the subpolar  
823 North Atlantic. Furthermore, the observed variability of the North Atlantic SSTs is  
824 reproduced remarkably well by all the models. These findings suggest that simulated  
825 temporal characteristics of the variables considered here are primarily dictated by the  
826 variability and trends in the CORE-II atmospheric datasets which include the im-  
827 pacts of ocean dynamics from nature superimposed onto external and anthropogenic  
828 effects. The general agreements among the models in their depictions of AMOC,  
829 MLD, and SSH variability and trends in the North Atlantic do not necessarily indi-

830 cate that the models accurately capture variability and trends seen in nature because  
831 there are undoubtedly errors in the forcing datasets and the models have errors and  
832 common, systematic biases. Indeed, agreements in variability and trends occur in  
833 the presence of large mean-state differences among the models – as well as large  
834 mean biases from observations – as documented in DY14. In that study, the over-  
835 arching hypothesis, namely that global ocean – sea-ice models integrated using the  
836 same inter-annually varying atmospheric forcing datasets will produce qualitatively  
837 similar mean and variability in their simulations, is found to be not satisfied for the  
838 mean states in the North Atlantic. In contrast, based on the present results, there  
839 appears to be more support for this hypothesis for variability in the North Atlantic.  
840 A similar conclusion is also reported in Wang et al. (2015) where the variability  
841 in the freshwater content and transports and sea-ice in the Arctic Ocean is found  
842 to be represented rather consistently among the models participating in CORE-II  
843 in spite of substantial differences in their mean states and mean state biases from  
844 observations.

845 Despite these general agreements, there are many differences – some significant  
846 – among the models, particularly in the spatial structures of variability patterns.  
847 For example, amplitudes and spatial extents of the largest SST and MLD anomalies  
848 differ among the models, reflecting the role of simulated ocean dynamics. Another  
849 notable difference occurs in the location of the largest AMOC anomalies (positive  
850 as depicted in Fig. 4). While the majority of the models have their maximum  
851 variability in the Northern Hemisphere, other models show enhanced variability in  
852 the Southern Hemisphere. Whether the maximum anomalies are located in the  
853 Northern or Southern Hemispheres does not appear to be related to the properties  
854 of the Southern Ocean meridional overturning circulations in these simulations (see  
855 Farneti et al., 2015). Similarly, there are no obvious connections between the subpolar

856 North Atlantic MLDs and where the maximum AMOC variability occurs. We do  
857 find, however, that the models that have their maximum variability in the Southern  
858 Hemisphere or in the vicinity of the equator tend to show weaker and statistically less  
859 significant AMOC trends, and their AMOC EOF1s account for a smaller fraction of  
860 their total variance in AMOC in comparison to those models with AMOC maximum  
861 variability in the Northern Hemisphere.

862 As in DY14, the differences among the model solutions do not suggest an obvious  
863 grouping of the models based on either their lineage, vertical coordinate represen-  
864 tations, or surface salinity restoring strengths. Again, we attribute these differences  
865 primarily to use of different subgrid scale parameterizations and their parameter  
866 values; differences in horizontal and vertical grid resolutions; and use of different  
867 sea-ice models along with diverse snow and sea-ice albedo treatments. Among the  
868 forward models, NOCS appears to deviate substantially in some of its low-frequency  
869 and trend characteristics from the other models. For example, it is the only model  
870 with a negative SSH trend in the subpolar North Atlantic for the 1993–2007 pe-  
871 riod; it is the only model with positive AMOC trends at both  $26.5^\circ$  and  $45^\circ\text{N}$  for  
872 the 1975–2007 period; and it shows the lowest trend in its LS upper-ocean density  
873 time series for the 1970–1995 period. These NOCS features are certainly in contrast  
874 with the solutions from the other NEMO-based models and the reasons for these  
875 differences remain unclear. However, several preliminary NOCS simulations that are  
876 underway in which the skew-flux form of GM90 is replaced with its advective form  
877 and / or associated tapering of both the thickness and isopycnal diffusivities within  
878 the surface mixed layer has been modified appear to show low frequency variability  
879 and trends that are in much better agreement with the other NEMO-based models.  
880 Based on both our present study and other work (e.g., Yeager et al., 2012; Yea-  
881 ger and Danabasoglu, 2014), we think that the CORE-II experimental protocol and

882 resulting simulations can be confidently used for studies concerning variability and  
883 its mechanisms on inter-annual and decadal times scales in the North Atlantic and  
884 elsewhere (e.g., Griffies et al., 2014; Farneti et al., 2015). The CORE-II effort has  
885 gained unprecedented momentum and exposure over the past few years, attract-  
886 ing participation of many ocean and climate modeling groups worldwide. As such,  
887 we think that it has now reached a mature state as the community standard for  
888 global ocean – sea-ice simulations. Encouraged by these developments, the CORE-II  
889 framework is recently proposed and endorsed as an Ocean Model Inter-comparison  
890 Project (OMIP) for inclusion in the Coupled Model Inter-comparison Project phase  
891 6 (CMIP6), again coordinated by the CLIVAR Ocean Model Development Panel  
892 (OMDP).

### 893 **Acknowledgments**

894 NCAR is sponsored by the U. S. National Science Foundation (NSF). The CESM  
895 is supported by the NSF and the U. S. Department of Energy. S. G. Yeager was  
896 supported by the NOAA Climate Program Office under Climate Variability and  
897 Predictability Program grants NA09OAR4310163 and NA13OAR4310138 and by  
898 the NSF Collaborative Research EaSM2 grant OCE-1243015 to NCAR. W. M. Kim  
899 was supported by the NOAA Climate Program Office under Climate Variability and  
900 Predictability Program grant NA13OAR4310136 to Texas A&M University. AC-  
901 CESS modeling work has been undertaken as part of the Australian Climate Change  
902 Science Program, funded jointly by the Department of Climate Change and En-  
903 ergy Efficiency, the Bureau of Meteorology and CSIRO, and was supported by the  
904 National Computational Infrastructure facility at the Australian National Univer-  
905 sity. AWI is a member of the Helmholtz Association of German Research Centers.  
906 Q. Wang and D. Sidorenko were funded by the Helmholtz Climate Initiative REK-

907 LIM (Regional Climate Change) project. The BERGEN contribution was supported  
908 by the Research Council of Norway through the EarthClim (207711/E10) and NO-  
909 TUR/NorStore projects, as well as the Centre for Climate Dynamics at the Bjerknes  
910 Centre for Climate Research. The CMCC contribution received funding from the  
911 Italian Ministry of Education, University, and Research and the Italian Ministry of  
912 Environment, Land, and Sea under the GEMINA project. INMOM was sponsored  
913 by the Russian Science Foundation (project number 14-27-00126). The KIEL con-  
914 tribution acknowledges support within the Co-Operative Project RACE - Regional  
915 Atlantic Circulation and Global Change funded by the German Federal Ministry  
916 for Education and Research (BMBF) under grant number 03F0651B and comput-  
917 ing resources from the North-German Supercomputing Alliance (HLRN). P. G. Fogli  
918 thanks W. G. Large, J. Tribbia, M. Vertenstein, G. Danabasoglu, and D. Bailey for  
919 their support and help in bringing NEMO into the CESM framework while visiting  
920 NCAR. E. Fernandez was supported by the BNP-Paribas foundation via the PRE-  
921 CLIDE project under the CNRS research convention agreement 30023488. We thank  
922 M. Harrison and R. Hallberg of GFDL for assistance with defining the GFDL-GOLD  
923 configuration, and R. Msadek and Y. M. Ruprich-Robert of GFDL for comments on  
924 an earlier version of the manuscript. Finally, we thank both the international CLI-  
925 VAR and U. S. CLIVAR projects for patiently sponsoring the Working Group on  
926 Ocean Model Development (now, Ocean Model Development Panel) over the years  
927 as COREs were developed.

## 928 **Appendix A. Two new HYCOM simulations**

929 The FSU HYCOM used in DY14 was based on an earlier version of HYCOM  
930 which advects density and  $S$  (instead of  $\theta$  and  $S$ ) and therefore does not conserve  
931 heat – see Griffies et al. (2014) for a discussion of impacts of this choice on sea



932 level. For the present study, a new HYCOM simulation, denoted as FSU2, has  
933 been performed with the formulation that advects  $\theta$  and  $S$ , thus conserving heat.  
934 Another new contribution that also uses the heat conserving formulation of HYCOM  
935 is GISS2. Here, we give brief summaries of these two new contributions in Appendix  
936 A.1 and Appendix A.2 for FSU2 and GISS2, respectively. Appendix A.3 includes  
937 a note on the use of  $\sigma_1$  vs.  $\sigma_2$  vertical coordinates in HYCOM. A short description  
938 of FSU2 and GISS2 time-mean solutions is presented in Appendix A.4, considering  
939 only AMOC and MHT distributions.

#### 940 *Appendix A.1. FSU2*

941 FSU2 is a global configuration of HYCOM (Bleck, 2002; Chassignet et al., 2003;  
942 Halliwell, 2004). The grid is a tripolar (Mercator grid smoothly connecting to a  
943 bipolar grid patch at about 47°N) Arakawa C-grid of 0.72° horizontal resolution  
944 with refinement at the equator. There are 500 and 382 grid cells in the zonal and  
945 meridional directions, respectively. The bottom topography is derived from the 2-  
946 minute NAVO / Naval Research Laboratory DBDB2 global dataset. The vertical  
947 discretization combines pressure coordinates at the surface, isopycnic coordinates in  
948 the stratified open ocean, and sigma coordinates over shallow coastal regions (Chas-  
949 signet et al., 2003, 2006). Thirty-two hybrid layers whose  $\sigma_2$  target densities range  
950 from 28.10 to 37.25 kg m<sup>-3</sup> are used. The initial conditions in  $\theta$  and  $S$  are given  
951 by the Polar Science Center Hydrographic Climatology version 2 dataset (PHC2; a  
952 blending of the Levitus et al. (1998) dataset with modifications in the Arctic Ocean  
953 based on Steele et al. (2001)). The ocean model is coupled to the sea-ice model CICE  
954 (Hunke and Lipscomb, 2010) that provides the ocean-ice fluxes. Turbulent air-sea  
955 fluxes are computed using the Large and Yeager (2009) bulk formulae. Surface fresh-  
956 water fluxes are applied as virtual salt fluxes as in FSU. Surface salinity is restored

957 over the entire domain with a piston velocity of 50 m over 4 years everywhere, except  
958 for the Antarctic region where the piston velocity is 50 m over 6 months. In addi-  
959 tion, a global normalization is applied to the restoring salinity flux at each time step.  
960 Vertical mixing is provided by the K-Profile Parameterization (KPP; Large et al.,  
961 1994) with a background diffusivity of  $10^{-5} \text{ m}^2 \text{ s}^{-1}$  and tracers are advected using a  
962 second-order flux corrected transport scheme. Lateral Laplacian diffusion of  $0.03\Delta x$   
963 is applied on  $\theta$  and  $S$  and a combination of Laplacian ( $0.03\Delta x$ ) and biharmonic  
964 ( $0.05\Delta x^3$ ) dissipation is applied on the velocities. Here,  $\Delta x$  represents grid spacing.  
965 Interface pressure smoothing, corresponding to GM90 as discussed in Gent (2011),  
966 is applied through a biharmonic operator, with a mixing coefficient determined by  
967 the grid spacing (in m) times a velocity scale of  $0.02 \text{ m s}^{-1}$  everywhere except in  
968 the Pacific and Atlantic north of  $40^\circ\text{N}$  where a Laplacian operator with a velocity  
969 scale of  $0.01 \text{ m s}^{-1}$  is used. The use of a biharmonic operator differs from GM90,  
970 but still ensures conversion from mean available potential energy to eddy potential  
971 energy. The interface pressure smoothing tapers off when the generalized vertical  
972 coordinate of HYCOM switches from isopycnal to pressure, mostly in the mixed  
973 layer and in unstratified regions. In such regions, lateral diffusion is oriented along  
974 pressure surfaces rather than rotated to neutral directions. No parameterization has  
975 been implemented for abyssal overflows.

976 We summarize the main differences between FSU2 and the version introduced in  
977 DY14 – labeled as FSU – as follows (FSU2 vs. FSU): (i) turbulent air-sea fluxes use  
978 Large and Yeager (2009) bulk formulae vs. Kara et al. (2005) bulk formulae; (ii)  
979 version 2.2.74 vs. version 2.2.21; (iii)  $\theta$  and  $S$  advection vs. density and  $S$  advection;  
980 (iv) tripolar grid of finer resolution ( $0.72^\circ$  vs.  $1^\circ$ ); (v) sea-ice model CICE v4.0 vs.  
981 CSIM (Community Sea-Ice Model; Briegleb et al., 2004; Holland et al., 2006); and  
982 (vi) surface salinity restoring time scale of 6 months vs. 4 years over 50 m in the

983 Antarctic region.

984 *Appendix A.2. GISS2*

985 The HYCOM version used at the National Aeronautics and Space Administration  
986 (NASA) Goddard Institute for Space Studies (GISS), denoted as GISS2, represents  
987 an updated version of the ocean component of the climate model described in Sun  
988 and Bleck (2006). It uses a Mercator grid, which smoothly connects to a bipolar  
989 grid patch at about 57°N. The horizontal mesh in the Mercator domain is  $1^\circ \times$   
990  $1^\circ \cos(\text{latitude})$ , but meridional resolution is enhanced near the equator, resulting in  
991 a  $1/3^\circ$  meridional mesh size at the equator. There are 360 and 387 grid points in  
992 the zonal and meridional directions, respectively (with the Bering Strait being the  
993 northernmost grid point in the extended Atlantic). The model is configured with 26  
994 hybrid  $\sigma_1$  coordinate levels. The adoption of this  $\sigma_1$  coordinate differs from Sun and  
995 Bleck (2006) where a  $\sigma_2$  coordinate was used. The bottom topography is obtained by  
996 spatially integrating ETOPO5 data of 5 minute spatial resolution over each model  
997 grid cell, without further smoothing. The initial  $\theta$  and  $S$  are given by the PHC3  
998 climatology. A non-slab KPP mixed layer sub-model (Halliwell, 2004) is employed.  
999 GISS2 uses the same prescriptions to specify lateral diffusivity and viscosity as in  
1000 FSU2 with the exception that the velocity scale used in the biharmonic operator is  
1001 a global constant set at  $0.05 \text{ m s}^{-1}$ .

1002 As in the original FSU contribution, GISS2 deviates from the suggested CORE-  
1003 II protocol in one important aspect. Namely, turbulent air-sea fluxes are computed  
1004 using the Kara et al. (2005) bulk formulae, instead of the Large and Yeager (2009)  
1005 bulk formulae. However, the other details of the forcing follow the protocol. Thus,  
1006 no restoring is applied to SSTs and no additional adjustment of surface heat flux  
1007 components, e.g., shortwave heat flux, are made. As a consequence, the global-mean

1008  $\theta$  in GISS2 increases by  $1/3^\circ\text{C}$  over the course of the 300-year simulation. Surface  
1009 freshwater fluxes are applied as virtual salt fluxes. Surface salinity is restored over the  
1010 entire domain with a piston velocity of 50 m over 4 years. Precipitation is multiplied  
1011 by a factor which aims to prevent long-term salinity trends. This factor is updated  
1012 monthly based on the departure of global-mean salt content from its initial value,  
1013 using a 1-year time scale. The adjustment factor stabilizes around 0.97, implying a  
1014 roughly 3% reduction of the imposed precipitation.

1015 The sea-ice model employed in GISS2 is a single-layer thermodynamic model  
1016 with ice advection by surface currents and a *shaving* device that laterally spreads ice  
1017 exceeding a prescribed thickness. Thus, it differs from Sun and Bleck (2006), where  
1018 the coupled ocean-atmosphere climate simulations at GISS use a more realistic sea-ice  
1019 model. One shortcoming of this highly simplified model is that melting and freezing  
1020 processes do not involve any exchange of water mass between ice and water; instead,  
1021 they spawn virtual salt fluxes. Since melting (freezing) reduces (increases) ocean  
1022 salinity, sea ice in this scheme contributes with a minus sign to the salt budget.  
1023 When attempting to reconcile surface freshwater fluxes with trends in the overall  
1024 oceanic salt content, one must be aware of this somewhat counter-intuitive aspect of  
1025 the sea-ice model.

1026 We identify five major differences between FSU2 and GISS2 configurations. They  
1027 are (FSU2 vs. GISS2): (i) nominal horizontal resolution of  $0.72^\circ$  vs.  $1^\circ$ ; (ii)  $\sigma_2$   
1028 vertical coordinate with 32 layers vs.  $\sigma_1$  vertical coordinate with 26 layers; (iii)  
1029 tripolar grid matching at  $47^\circ\text{N}$  vs. at  $57^\circ\text{N}$ ; (iv) CICE4.0 sea-ice model vs. one-layer  
1030 thermodynamic sea-ice model; and (v) use of Large and Yeager (2009) vs. Kara et al.  
1031 (2005) bulk formulae.

1032 *Appendix A.3. A note on use of  $\sigma_1$  vs.  $\sigma_2$  vertical coordinates in HYCOM*

1033 A few remarks are in order to explain the choice of  $\sigma_1$  as vertical coordinate  
1034 in GISS2 in contrast with the use of  $\sigma_2$  coordinate in FSU and FSU2. A major  
1035 problem in models featuring sloping coordinate surfaces is the two-term expression  
1036 for the horizontal pressure gradient force. In HYCOM, the numerically challenging  
1037 two-term pressure gradient force is transformed into a more benign, single-term ex-  
1038 pression by treating sea water as incompressible and, for dynamic consistency with  
1039 this approximation, by replacing density with a globally referenced potential density  
1040 ( $\rho_{pot}$ ) in the equation of state (Spiegel and Veronis, 1960).

1041 One shortcoming of the above approximation is that a water column which is  
1042 stably stratified in the real ocean may not be stably stratified in  $\rho_{pot}$  space. The choice  
1043 of  $\sigma_2$  in HYCOM, traditionally regarded as the best compromise, is particularly  
1044 problematic in the upper Southern Ocean where convection triggered by a reversal  
1045 of the vertical  $\rho_{pot}$  gradient can weaken the seasonal summertime halocline to the  
1046 point where it becomes hard to form new ice in the fall. Without ice cover, the  
1047 Southern Ocean acts as a heat source in austral winter, with grave consequences in  
1048 a coupled climate model.

1049 It is for this reason that in the GISS version of HYCOM, i.e., GISS2,  $\sigma_2$  has  
1050 been replaced by  $\sigma_1$ , both in the equation of state and as vertical coordinate. Static  
1051 stability problems in the abyssal Atlantic due the use of  $\sigma_1$  as vertical coordinate  
1052 have been found to be less serious than expected – in the sense that they do not  
1053 appear to preclude the existence of an abyssal, Southern-Ocean driven overturning  
1054 cell.

1055 The HYCOM versions in FSU and FSU2 add the thermobaricity treatment of Sun  
1056 et al. (1999) to the basic Boussinesq-related approximations listed above. Accounting  
1057 for thermobaric effects has been found to reduce Southern Ocean sea-ice biases in

1058 the  $\sigma_2$ -based FSU and FSU2 models. GISS2 does not account for thermobaricity,  
1059 relying instead on the use of  $\sigma_1$  to alleviate this problem.

1060 *Appendix A.4. Time-mean AMOC and MHT in FSU2 and GISS2*

1061 A detailed analysis of the time-mean solutions from FSU2 and GISS2, as was  
1062 done in DY14 for the other participating models, is beyond the scope of the present  
1063 study. Instead, we only provide a brief assessment of their time-mean AMOC and  
1064 MHT distributions, considering the solutions from the fifth cycle of their CORE-II  
1065 simulations.

1066 Figure 17 shows the time-mean (years 1988–2007 mean) AMOC distributions  
1067 in depth–latitude space from FSU, FSU2, and GISS2, corresponding to Fig. 3 of  
1068 DY14. With  $< 8$  Sv, FSU has the weakest NADW maximum transport among all  
1069 the participating models. This maximum transport is  $> 14$  Sv and  $> 22$  Sv in FSU2  
1070 and GISS2, respectively. The NADW penetration depth as measured by the depth of  
1071 the zero contour line is deeper in FSU2 and GISS2 than in FSU. Indeed, the NADW  
1072 penetration depth exceeds 5 km in GISS2. In both FSU2 and GISS2, the transports  
1073 associated with the Antarctic Bottom Water (AABW) are quite weak.

1074 We provide a quantitative comparison of the AMOC profiles from FSU2 and  
1075 GISS2 to the profile based on the RAPID data (Cunningham et al., 2007) at  $26.5^\circ\text{N}$   
1076 in Fig. 18a. The figure corresponds to Fig. 5 of DY14 and uses the 4-year mean  
1077 for years 2004–2007 for the model data while the RAPID data represents the 4-year  
1078 mean for April 2004 – March 2008. The profile for FSU is also included for reference  
1079 purposes. We note that the profiles show the total integrated transport between the  
1080 surface and a given depth, with negative and positive slopes indicating northward  
1081 and southward flow, respectively. The RAPID estimate for the NADW maximum  
1082 transport at this latitude is 18.6 Sv, occurring at about 1000-m depth, with about

1083  $\pm 1$  Sv as its annual-mean range over this short period.

1084 As indicated above, FSU has the lowest NADW maximum transport among all  
1085 the models with only 5.3 Sv, and its profile deviates quite substantially from the  
1086 RAPID profile. FSU2 shows major improvements from FSU in both the NADW  
1087 maximum transport magnitude with 11.5 Sv and the vertical structure of the trans-  
1088 port profile. Nevertheless, the NADW maximum transport in FSU2 still remains  
1089 considerably lower than in RAPID. In GISS2, the NADW maximum transport of  
1090 about 19.2 Sv is only slightly stronger than in RAPID and its profile captures that  
1091 of RAPID well, including the NADW penetration depth. As in all the other partic-  
1092 ipating models (see Fig. 5 of DY14), both FSU2 and GISS2 show significant depar-  
1093 tures from the RAPID profile in their representations of the AABW with near-zero  
1094 transports at this latitude. We note that the GISS2 profile arguably shows one of  
1095 the best comparisons with that of RAPID among all the participating models.

1096 We present the time-mean (years 1988–2007 mean) Atlantic Ocean MHT distri-  
1097 butions from FSU, FSU2, and GISS2 in Fig. 18b, as in Fig. 6 of DY14. The figure also  
1098 includes the implied transport estimates from Large and Yeager (2009) calculated us-  
1099 ing the CORE-II datasets with observed SSTs and sea-ice for the 1984–2006 period,  
1100 and the direct estimates with their uncertainty ranges from Bryden and Imawaki  
1101 (2001) and the estimate from the RAPID data (Johns et al., 2011). As a result of its  
1102 weakest NADW transport, FSU has the lowest MHT among the participating mod-  
1103 els with about 0.40 PW. In addition, FSU is the only model with southward heat  
1104 transport in the Atlantic basin. Again, FSU2 represents an improved solution over  
1105 FSU, with a maximum MHT of 0.86 PW. Still, however, FSU2 MHT distribution  
1106 remains below the range of the estimates, except south of 10°S. With the exception  
1107 of north of 60°N, GISS2 distribution is within the bounds of the estimates, with  
1108 maximum heat transports of about 1.1 PW, occurring at 10°N and 30°N. Including

1109 FSU2 and GISS2, none of the models participating in CORE-II is able to obtain the  
1110 RAPID based estimate of 1.33 PW at this latitude for this time period – see Msadek  
1111 et al. (2013) and DY14 for a discussion of lower MHTs in the model simulations.

## 1112 **Appendix B. Departures from the CORE-II protocol**

1113 Despite our best efforts, about half of the participating models did not follow the  
1114 recommendations of the CORE-II protocol exactly. The departures include use of  
1115 different bulk formulae, modifications of the Large and Yeager (2009) bulk formulae,  
1116 and changes in the forcing datasets.

1117 For historical reasons, INMOM uses the bulk formulae adopted from the Arctic  
1118 Ocean Model Inter-comparison Project (AOMIP), while FSU and GISS2 use the Kara  
1119 et al. (2005) formulae. In MRI-F and MRI-A (data assimilated version of MRI-F),  
1120 the air-ice neutral bulk transfer coefficients are modified to follow the values in Mellor  
1121 and Kantha (1989), because the thermodynamic part of their sea-ice model is based  
1122 on Mellor and Kantha (1989). Specifically, the momentum transfer coefficient is set  
1123 to  $3 \times 10^{-3}$  and the transfer coefficients for sensible heat and evaporation are set to  
1124  $1.5 \times 10^{-3}$ , in contrast with a value of  $1.63 \times 10^{-3}$  used in Large and Yeager (2009).

1125 Regarding the modifications of the forcing datasets, CERFACS, CNRM, and  
1126 NOCS impose a seasonal cycle to the Antarctic runoff whereby four times the annual-  
1127 mean value is applied over the summer months, i.e., January, February, and March,  
1128 and zero runoff is used for the rest of the year. In the CORE-II protocol, the Antarctic  
1129 runoff is time-invariant.

1130 In addition to using different bulk formulae, INMOM adds  $1 \text{ m s}^{-1}$  to the CORE-  
1131 II wind data uniformly, prior to the calculation of the wind stress to improve their  
1132 sea-ice simulations, particularly in the Arctic basin. As a result, the wind stress for  
1133 INMOM is larger than in any other model – see Fig. 3 of Farneti et al. (2015).



1134 Finally, KIEL has three differences from the protocol: i) the wind stress near  
1135 Antarctica is modified to include a parameterization of katabatic winds; ii) a differ-  
1136 ent runoff dataset – though still based on Dai and Trenberth (2002) – is adopted; and  
1137 iii) model potential temperature and salinity are restored to observed monthly-mean  
1138 climatology in the Gulf of Cadiz region to improve the representation of the Mediter-  
1139 ranean outflow. This restoring is applied within the 627–1297 m depth range and  
1140 its strength varies with depth and distance from the coast.

1141 We do not know the impacts of these departures from the CORE-II protocol on  
1142 model solutions. While some, e.g., transfer coefficient changes, are expected to have  
1143 minor impacts, the use of different bulk formulae can result in larger changes in  
1144 model solutions. It is, nevertheless, clear that, despite our best efforts, we are still  
1145 short of achieving our ultimate goal of having all groups follow the protocol fully.  
1146 The protocol does not specify a particular recipe for surface salinity restoring; it  
1147 is left to the modeling groups to choose their optimal salinity restoring procedure.  
1148 Thus, given the diversity among the models in their use of quite different restoring  
1149 time scales – see Appendix C of DY14 – it is possible that the differences in model  
1150 solutions due to their departures from the CORE-II protocol could be substantially  
1151 masked.

## 1152 **Appendix C. List of Major Acronyms**

- 1153 – ACCESS: Australian Community Climate and Earth System Simulator
- 1154 – AMOC: Atlantic meridional overturning circulation
- 1155 – AMV: Atlantic multi-decadal variability
- 1156 – AVISO: Archiving, Validation, and Interpolation of Satellite Oceanographic

- 1157 Data
- 1158 – AWI: Alfred Wegener Institute
- 1159 – BSF: Barotropic streamfunction
- 1160 – CERFACS: Centre Européen de Recherche et de Formation Avancée en Calcul  
1161 Scientifique
- 1162 – CESM: Community Earth System Model
- 1163 – CGCM: Coupled general circulation model
- 1164 – CICE: Sea ice model
- 1165 – CLIVAR: Climate Variability and Predictability
- 1166 – CMCC: Centro Euro-Mediterraneo sui Cambiamenti Climatici
- 1167 – CNRM: Centre National de Recherches Météorologiques
- 1168 – CORE-II: Coordinated Ocean-ice Reference Experiments phase II
- 1169 – CSIM: Community Sea Ice Model
- 1170 – DWF: Deep water formation
- 1171 – DY14: Danabasoglu et al. (2014)
- 1172 – ECMWF: European Center for Medium-range Weather Forecasting
- 1173 – EOF: Empirical orthogonal function
- 1174 – FESIM: Finite Element Sea-ice Model

- 1175 – FESOM: Finite Element Sea-ice Ocean Model
- 1176 – FSU: Florida State University
- 1177 – FSU2: Version 2 of the FSU contribution
- 1178 – GFDL: Geophysical Fluid Dynamics Laboratory
- 1179 – GISS: Goddard Institute for Space Studies
- 1180 – GISS2: HYCOM contribution from GISS
- 1181 – GM90: Gent and McWilliams (1990) parameterization
- 1182 – GOLD: Generalized Ocean Layer Dynamics
- 1183 – HYCOM: HYbrid Coordinate Ocean Model
- 1184 – ICTP: International Centre for Theoretical Physics
- 1185 – INMOM: Institute of Numerical Mathematics Ocean Model
- 1186 – KIEL: Refers to the contribution from the Helmholtz Center for Ocean Re-  
1187 search from Kiel
- 1188 – KPP: K-Profile Parameterization (Large et al., 1994)
- 1189 – LIM: Louvain-la-Neuve Sea Ice Model
- 1190 – LS: Labrador Sea
- 1191 – MHT: Meridional heat transport
- 1192 – MICOM: Miami Isopycnal Coordinate Ocean Model

- 1193 – MIT: Massachusetts Institute of Technology
- 1194 – MITgcm: Massachusetts Institute of Technology general circulation model
- 1195 – MLD: mixed layer depth
- 1196 – MMM: Multi-model mean
- 1197 – MOM: Modular Ocean Model
- 1198 – MOVE: Multivariate Ocean Variational Estimation
- 1199 – MRI: Meteorological Research Institute
- 1200 – MRI.COM: Meteorological Research Institute Community Ocean Model
- 1201 – MRI-A: Data assimilated version of MRI-F
- 1202 – MRI-F: MRI contribution
- 1203 – NAC: North Atlantic Current
- 1204 – NADW: North Atlantic Deep Water
- 1205 – NAO: North Atlantic Oscillation
- 1206 – NASA: National Aeronautics and Space Administration
- 1207 – NCAR: National Center for Atmospheric Research
- 1208 – NCEP: National Centers for Environmental Prediction
- 1209 – NEMO: Nucleus for European Modelling of the Ocean
- 1210 – NOAA: National Oceanic and Atmospheric Administration

- 1211 – NOCS: National Oceanography Centre Southampton
- 1212 – NorESM-O: Norwegian Earth System Model ocean component
- 1213 – OMDP: Ocean Model Development Panel
- 1214 – ORCA: Ocean model configuration of the NEMO model
- 1215 – PC: Principal component
- 1216 – PHC: Polar Science Center Hydrographic Climatology
- 1217 – POP2: Parallel Ocean Program version 2
- 1218 – RAPID: Rapid Climate Change mooring data
- 1219 – SIS: GFDL Sea Ice Simulator
- 1220 – SPG: Subpolar gyre
- 1221 – SSH: Sea surface height
- 1222 – SST: Sea surface temperature

1223 **References**

- 1224 Beismann, J.-O., Barnier, B., 2004. Variability of the meridional overturning circu-  
1225 lation of the North Atlantic: sensitivity to overflows of dense water masses. *Ocean*  
1226 *Dynamics* 54, 92–106.
- 1227 Bentsen, M., Drange, H., Furevik, T., Zhou, T., 2004. Simulated variability of the  
1228 Atlantic meridional overturning circulation. *Clim. Dyn.* 22, 701–720.
- 1229 Biastoch, A., Böning, C. W., Getzlaff, J., Molines, J.-M., Madec, G., 2008. Causes  
1230 of interannual - decadal variability in the meridional overturning circulation of the  
1231 mid-latitude North Atlantic Ocean. *J. Climate* 21, 6599–6615.
- 1232 Bleck, R., 2002. An oceanic general circulation model framed in hybrid isopycnic-  
1233 Cartesian coordinates. *Ocean Modelling* 4, 55–88.
- 1234 Böning, C. W., Dieterich, C., Barnier, B., Jia, Y. L., 2001. Seasonal cycle of the  
1235 meridional heat transport in the subtropical North Atlantic: a model intercom-  
1236 parison in relation to observations near 25°N. *Prog. Oceanogr.* 48, 231–253.
- 1237 Böning, C. W., Scheinert, M., Dengg, J., Biastoch, A., Funk, A., 2006. Decadal  
1238 variability of subpolar gyre transport and its reverberation in the North Atlantic  
1239 overturning. *Geophys. Res. Lett.* 33, L21S01.
- 1240 Branstator, G., Teng, H., 2010. Two limits of initial-value decadal predictability in  
1241 a CGCM. *J. Climate* 23, 6292–6311.
- 1242 Briegleb, B. P., Bitz, C. M., Hunke, E. C., Lipscomb, W. H., Holland, M. M.,  
1243 Schramm, J. L., Moritz, R. E., 2004. Scientific description of the sea-ice compo-  
1244 nent in the Community Climate System Model, version three. NCAR Tech. Note

- 1245 NCAR/TN-463+STR, National Center for Atmospheric Research, Boulder, Col-  
1246 orado.
- 1247 Brodeau, L., Barnier, B., Treguier, A. M., Penduff, T., Gulev, S., 2010. An ERA40-  
1248 based atmospheric forcing for global ocean circulation models. *Ocean Modelling*  
1249 31, 88–104.
- 1250 Bryan, F. O., Danabasoglu, G., Nakashiki, N., Yoshida, Y., Kim, D. H., Tsutsui,  
1251 J., Doney, S. C., 2006. Response of North Atlantic thermohaline circulation and  
1252 ventilation to increasing carbon dioxide in CCSM3. *J. Climate* 19, 2382–2397.
- 1253 Bryden, H., Imawaki, S., 2001. Ocean heat transport. In: Siedler, G., Church, J.,  
1254 Gould, J. (Eds.), *Ocean circulation and climate*. Vol. 77 of International Geophys-  
1255 ical Series. Academic Press, pp. 317–336.
- 1256 Chassignet, E. P., Hurlburt, H. E., Smedstad, O. M., Halliwell, G. R., Wallcraft,  
1257 A. J., Metzger, E. J., Blanton, B. O., Lozano, C., Rao, D. B., Hogan, P. J.,  
1258 Srinivasan, A., 2006. Generalized vertical coordinates for eddy-resolving global  
1259 and coastal forecasts. *Oceanography* 19, 20–31.
- 1260 Chassignet, E. P., Smith, L. T., Halliwell, G. T., Bleck, R., 2003. North Atlantic sim-  
1261 ulations with the Hybrid Coordinate Ocean Model (HYCOM): Impact of the ver-  
1262 tical coordinate choice, reference pressure, and thermobaricity. *J. Phys. Oceanogr.*  
1263 33, 2504–2526.
- 1264 Cunningham, S. A., Kanzow, T., Rayner, D., Baringer, M. O., Johns, W. E.,  
1265 Marotzke, J., Longworth, H. R., Grant, E. M., Hirschi, J. J.-M., Beal, L. M.,  
1266 Meinen, C. S., Bryden, H. L., 2007. Temporal variability of the Atlantic merid-  
1267 ional overturning circulation at 26.5°N. *Science* 317, 935–938.

- 1268 Dai, A., Trenberth, K. E., 2002. Estimates of freshwater discharge from continents:  
1269 Latitudinal and seasonal variations. *J. Hydrometeorology* 3, 660–687.
- 1270 Danabasoglu, G., 2008. On multidecadal variability of the Atlantic meridional over-  
1271 turning circulation in the Community Climate System Model version 3. *J. Climate*  
1272 21, 5524–5544.
- 1273 Danabasoglu, G., Yeager, S. G., Bailey, D., Behrens, E., Bentsen, M., Bi, D., Bi-  
1274 astoch, A., Böning, C., Bozec, A., Canuto, V. M., Cassou, C., Chassignet, E.,  
1275 Danilov, S., Diansky, N., Drange, H., Farneti, R., Fernandez, E., Fogli, P. G.,  
1276 Forget, G., Fujii, Y., Griffies, S. M., Gusev, A., Heimbach, P., Howard, A., Jung,  
1277 T., Kelley, M., Large, W. G., Leboissetier, A., Lu, J., Marsland, S. J., Masina,  
1278 S., Navarra, A., Nurser, A. J. G., Pirani, A., Salas y Mélia, D., Samuels, B. L.,  
1279 Scheinert, M., Sidorenko, D., Treguier, A.-M., Tsujino, H., Uotila, P., Valcke, S.,  
1280 Voltaire, A., Wang, Q., 2014. North Atlantic simulations in Coordinated Ocean-ice  
1281 Reference Experiments phase II (CORE-II). Part I: Mean states. *Ocean Modelling*  
1282 73, 76–107.
- 1283 Danabasoglu, G., Yeager, S. G., Kwon, Y.-O., Tribbia, J. J., Phillips, A. S., Hur-  
1284 rell, J. W., 2012. Variability of the Atlantic meridional overturning circulation in  
1285 CCSM4. *J. Climate* 25, 5153–5172.
- 1286 Delworth, T., Manabe, S., Stouffer, R. J., 1993. Interdecadal variations of the thermo-  
1287 haline circulation in a coupled ocean-atmosphere model. *J. Climate* 6, 1993–2011.
- 1288 Delworth, T. L., Mann, M. E., 2000. Observed and simulated multidecadal variability  
1289 in the Northern Hemisphere. *Clim. Dyn.* 16, 661–676.
- 1290 Delworth, T. L., Zeng, F., 2012. Multicentennial variability of the Atlantic Meridional



- 1291 Overturning Circulation and its climate influence in a 4000 year simulation of the  
1292 GFDL CM2.1 climate model. *Geophys. Res. Lett.* 39, L13702.
- 1293 Deshayes, J., Frankignoul, C., 2008. Simulated variability of the circulation in the  
1294 North Atlantic from 1953 to 2003. *J. Clim.* 21, 4919–4933.
- 1295 Doney, S. C., Yeager, S., Danabasoglu, G., Large, W. G., McWilliams, J. C., 2007.  
1296 Mechanisms governing interannual variability of upper-ocean temperature in a  
1297 global ocean hindcast simulation. *J. Phys. Oceanogr.* 37, 1918–1938.
- 1298 Dong, B., Sutton, R. T., 2005. Mechanism of interdecadal thermohaline circulation  
1299 variability in a coupled ocean - atmosphere GCM. *J. Climate* 18, 1117–1135.
- 1300 Ducet, N., Le Traon, P.-Y., Reverdin, G., 2000. Global high-resolution mapping of  
1301 ocean circulation from TOPEX/Poseidon and ERS-1 and -2. *J. Geophys. Res.* 105,  
1302 19477–19498.
- 1303 Eden, C., Willebrand, J., 2001. Mechanism of interannual to decadal variability of  
1304 the North Atlantic circulation. *J. Climate* 14, 2266–2280.
- 1305 Farneti, R., Downes, S. M., Griffies, S. M., Marsland, S. J., Behrens, E., Bentsen,  
1306 M., Bi, D., Biastoch, A., Böning, C., Bozec, A., Canuto, V. M., Chassignet, E.,  
1307 Danabasoglu, G., Danilov, S., Diansky, N., Drange, H., Fogli, P. G., Gusev, A.,  
1308 Hallberg, R. W., Howard, A., Ilicak, M., Jung, T., Kelley, M., Large, W. G.,  
1309 Leboissetier, A., Long, M., Lu, J., Masina, S., Mishra, A., Navarra, A., Nurser,  
1310 A. J. G., Patara, L., Samuels, B. L., Sidorenko, D., Tsujino, H., Uotila, P., Wang,  
1311 Q., Yeager, S. G., 2015. An assessment of Antarctic Circumpolar Current and  
1312 Southern Ocean meridional overturning circulation during 1958-2007 in a suite of  
1313 interannual CORE-II simulations. *Ocean Modelling* 93, 84–120.

- 1314 Farneti, R., Vallis, G. K., 2011. Mechanisms of interdecadal climate variability and  
1315 the role of ocean-atmosphere coupling. *Clim. Dyn.* 36, 289–308.
- 1316 Fox-Kemper, B., Danabasoglu, G., Ferrari, R., Griffies, S. M., Hallberg, R. W., Hol-  
1317 land, M. M., Maltrud, M. E., Peacock, S., Samuels, B. L., 2011. Parameterization  
1318 of mixed layer eddies. Part III: Implementation and impact in global ocean climate  
1319 simulations. *Ocean Modelling* 39, 61–78.
- 1320 Gent, P. R., 2011. The Gent-McWilliams parameterization: 20/20 hindsight. *Ocean*  
1321 *Modelling* 39, 2–9.
- 1322 Gent, P. R., McWilliams, J. C., 1990. Isopycnal mixing in ocean circulation models.  
1323 *J. Phys. Oceanogr.* 20, 150–155.
- 1324 Griffies, S. M., Bryan, K., 1997. Predictability of North Atlantic multidecadal climate  
1325 variability. *Science* 275, 181–184.
- 1326 Griffies, S. M., Winton, M., Samuels, B., Danabasoglu, G., Yeager, S., Marsland, S.,  
1327 Drange, H., Bentsen, M., 2012. Datasets and protocol for the CLIVAR WGOMD  
1328 Coordinated Ocean sea-ice Reference Experiments (COREs). WCRP Report No.  
1329 21/2012.
- 1330 Griffies, S. M., Yin, J., Durack, P. J., Goddard, P., Bates, S. C., Behrens, E.,  
1331 Bentsen, M., Bi, D., Biastoch, A., Böning, C. W., Bozec, A., Chassignet, E., Dan-  
1332 abasoglu, G., Danilov, S., Domingues, C. M., Drange, H., Farneti, R., Fernandez,  
1333 E., Greatbatch, R. J., Holland, D. M., Ilicak, M., Large, W. G., Lorbacher, K.,  
1334 Lu, J., Marsland, S. J., Mishra, A., Nurser, A. J. G., Salas y Mélia, D., Palter,  
1335 J. B., Samuels, B. L., Schröter, J., Schwarzkopf, F. U., Sidorenko, D., Treguier,  
1336 A. M., Tseng, Y.-H., Tsujino, H., Uotila, P., Valcke, S., Voltaire, A., Wang, Q.,

- 1337 Winton, M., Zhang, X., 2014. An assessment of global and regional sea level for  
1338 years 1993-2007 in a suite of interannual CORE-II simulations. *Ocean Modelling*  
1339 78, 35–89.
- 1340 Gusev, A. V., Diansky, N. A., 2014. Numerical simulation of the World ocean cir-  
1341 culation and its climatic variability for 1948-2007 using the INMOM. *Izvestiya,*  
1342 *Atmospheric and Oceanic Physics* 50, 1–12.
- 1343 Häkkinen, S., 1999. Variability of the simulated meridional transport in the North  
1344 Atlantic for the period 1951-1993. *J. Geophys. Res.* 104, 10991–11007.
- 1345 Häkkinen, S., Rhines, P. B., 2004. Decline of subpolar North Atlantic circulation  
1346 during the 1990s. *Science* 304, 555–559.
- 1347 Halliwell, G. R., 2004. Evaluation of vertical coordinate and vertical mixing algo-  
1348 rithms in the HYbrid-Coordinate Ocean Model (HYCOM). *Ocean Modelling* 7,  
1349 285–322.
- 1350 Hibler, W., 1979. A dynamic thermodynamic sea ice model. *J. Phys. Oceanogr.* 9,  
1351 815–846.
- 1352 Holland, M. M., Bitz, C. M., Hunke, E. C., Lipscomb, W. H., Schramm, J. L.,  
1353 2006. Influence of the sea ice thickness distribution on polar climate in CCSM3.  
1354 *J. Climate* 19, 2398–2414.
- 1355 Hunke, E. C., Lipscomb, W. H., 2010. CICE: the Los Alamos Sea Ice Model docu-  
1356 mentation and software users manual version 4.1. Los Alamos National Laboratory  
1357 Tech. Rep. LA-CC-06012, Los Alamos, NM.

- 1358 Hurrell, J. W., Hack, J. J., Shea, D., Caron, J. M., Rosinski, J., 2008. A new sea  
1359 surface temperature and sea ice boundary dataset for the Community Atmosphere  
1360 Model. *J. Climate* 21, 5145–5153.
- 1361 Johns, W. E., Baringer, M. O., Beal, L. M., Cunningham, S. A., Kanzow, T., Bryden,  
1362 H. L., Hitschi, J. J. M., Marotzke, J., Meinen, C. S., Shaw, B., Curry, R., 2011.  
1363 Continuous, array-based estimates of Atlantic Ocean heat transport at 26.5°N. *J.*  
1364 *Climate* 24, 2429–2449.
- 1365 Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell,  
1366 M., Saha, S., White, G., Woollen, J., Zhu, Y., Leetmaa, A., Reynolds, R., Chelliah,  
1367 M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K. C., Ropelewski, C., Wang,  
1368 J., Jenne, R., Joseph, D., 1996. The NCEP/NCAR 40-year reanalysis project.  
1369 *Bull. Amer. Meteor. Soc.* 77, 437–471.
- 1370 Kara, A. B., Hurlburt, H. E., Wallcraft, A. J., 2005. Stability-dependent exchange  
1371 coefficients for air-sea fluxes. *J. Atmospheric and Oceanic Technology* 22, 1080–  
1372 1094.
- 1373 Karspeck, A. R., Stammer, D., Köhl, A., Danabasoglu, G., Balmaseda, M., Smith,  
1374 D. M., Fujii, Y., Zhang, S., Giese, B., Tsujino, H., Rosati, A., 2015. Comparison  
1375 of the Atlantic meridional overturning circulation between 1960 and 2007 in six  
1376 ocean reanalysis products. *Clim. Dyn.*(in press).
- 1377 Kushnir, Y., 1994. Interdecadal variations in North Atlantic sea surface temperature  
1378 and associated atmospheric conditions. *J. Climate* 7, 141–157.
- 1379 Kwon, Y.-O., Frankignoul, C., 2012. Stochastically-driven multidecadal variability

- 1380 of the Atlantic meridional overturning circulation in CCSM3. *Clim. Dyn.* 38, 859–  
1381 876.
- 1382 Kwon, Y.-O., Frankignoul, C., 2014. Mechanisms of multidecadal Atlantic merid-  
1383 ional overturning circulation variability diagnosed in depth versus density space.  
1384 *J. Climate* 27, 9359–9376.
- 1385 Large, W. G., McWilliams, J. C., Doney, S. C., 1994. Oceanic vertical mixing: A re-  
1386 view and a model with a nonlocal boundary layer parameterization. *Rev. Geophys.*  
1387 32, 363–403.
- 1388 Large, W. G., Yeager, S., 2004. Diurnal to decadal global forcing for ocean and sea-  
1389 ice models: The data sets and flux climatologies. NCAR Tech. Note NCAR/TN-  
1390 460+STR.
- 1391 Large, W. G., Yeager, S. G., 2009. The global climatology of an interannually varying  
1392 air-sea flux data set. *Clim. Dyn.* 33, 341–364.
- 1393 Lavender, K. L., Davis, R. E., Owens, W. B., 2002. Observations of open-ocean  
1394 deep convection in the Labrador Sea from subsurface floats. *J. Phys. Oceanogr.*  
1395 32, 511–526.
- 1396 Le Traon, P.-Y., Nadal, F., Ducet, N., 1998. An improved mapping method of mul-  
1397 tisatellite altimeter data. *J. Atmos. Oceanic Technol.* 15, 522–534.
- 1398 Levitus, S., Boyer, T., Conrigh, M., Johnson, D., O'Brien, T., Antonov, J.,  
1399 Stephens, C., Garfield, R., 1998. World Ocean Database 1998, volume I: Intro-  
1400 duction.
- 1401 Liu, Z., 2012. Dynamics of interdecadal climate variability: A historical perspective.  
1402 *J. Climate* 25, 1963–1994.

- 1403 Lohmann, K., Drange, H., Bentsen, M., 2009a. A possible mechanism for the strong  
1404 weakening of the North Atlantic subpolar gyre in the mid-1990s. *Geophys. Res.*  
1405 *Lett.* 36, L15602.
- 1406 Lohmann, K., Drange, H., Bentsen, M., 2009b. Response of the North Atlantic  
1407 subpolar gyre to persistent North Atlantic oscillation like forcing. *Clim. Dyn.* 32,  
1408 273–285.
- 1409 Mellor, L. G., Kantha, L., 1989. An ice - ocean coupled model. *J. Geophys. Res.* 94,  
1410 10937–10954.
- 1411 Msadek, R., Dixon, K. W., Delworth, T. L., Hurlin, W., 2010. Assessing the pre-  
1412 dictability of the Atlantic meridional overturning circulation and associated fin-  
1413 gerprints. *Geophys. Res. Lett.* 37, L19608.
- 1414 Msadek, R., Johns, W. E., Yeager, S. G., Danabasoglu, G., Delworth, T. L., Rosati,  
1415 A., 2013. The Atlantic meridional heat transport at 26.5°N and its relationship  
1416 with the MOC in the RAPID array and the GFDL and NCAR coupled models. *J.*  
1417 *Climate* 26, 4335–4356.
- 1418 Pohlmann, H., Botzet, M., Latif, M., Roesch, A., Wild, M., Tschuck, P., 2004.  
1419 Estimating the decadal predictability of a coupled AOGCM. *J. Climate* 17, 4463–  
1420 4472.
- 1421 Pohlmann, H., Smith, D. M., Balmaseda, M. A., Keenlyside, N. S., Masina, S., Matei,  
1422 D., Müller, W. A., Rogel, P., 2013. Predictability of the mid-latitude Atlantic  
1423 meridional overturning circulation in a multi-model system. *Clim. Dyn.* 41, 775–  
1424 785.

- 1425 Robson, J., Sutton, R., Lohmann, K., Smith, D., Palmer, M. D., 2012. Causes of the  
1426 rapid warming of the North Atlantic Ocean in the 1990s. *J. Climate* 25, 4116–4134.
- 1427 Schlesinger, M. E., Ramankutty, N., 1994. An oscillation in the global climate system  
1428 of period 65–70 years. *Nature* 367, 723–726.
- 1429 Spiegel, E. A., Veronis, G., 1960. On the Boussinesq approximation for a compressible  
1430 fluid. *Astrophys. J.* 131, 442–447.
- 1431 Srokosz, M., Baringer, M., Bryden, H., Cunningham, S., Delworth, T., Lozier, S.,  
1432 Marotzke, J., Sutton, R., 2012. Past, present and future change in the Atlantic  
1433 meridional overturning circulation. *BAMS* 93, 1663–1676.
- 1434 Steele, M., Morley, R., Ermold, W., 2001. PHC: A global ocean hydrography with a  
1435 high quality Arctic Ocean. *J. Climate* 14, 2079–2087.
- 1436 Sun, S., Bleck, R., 2006. Multi-century simulations with the coupled GISS-HYCOM  
1437 climate model: control experiments. *Clim. Dyn.* 26, 407–428.
- 1438 Sun, S., Bleck, R., Rooth, C., Dukowicz, J., Chassignet, E., Killworth, P., 1999. In-  
1439 clusion of thermobaricity in isopycnic-coordinate ocean models. *J. Phys. Oceanog.*  
1440 29, 2719–2729.
- 1441 Sutton, R. W., Hodson, D. L. R., 2005. Atlantic Ocean forcing of North American  
1442 and European summer climate. *Science* 309, 115–118.
- 1443 Teng, H., Branstator, G., Meehl, G. A., 2011. Predictability of the Atlantic over-  
1444 turning circulation and associated surface patterns in two CCSM3 climate change  
1445 ensemble experiments. *J. Climate* 24, 6054–6076.

- 1446 Uppala, S. M., Kallberg, P. W., Simmons, A. J., Andrae, U., Da Costa Bechtold, V.,  
1447 Fiorino, M., Gibson, J. K., Haseler, J., Hernandez, A., Kelly, G. A., Li, X., Onogi,  
1448 K., Saarinen, S., Sokka, N., Allan, R. P., Andersson, E., Arpe, K., Balmaseda,  
1449 M. A., Beljaars, A. C. M., Van De Berg, L., Bidlot, J., Bormann, N., Caires, S.,  
1450 Chevallier, F., Dethof, A., Dragosavac, M., Fisher, M., Fuentes, M., Hagemann,  
1451 S., Holm, E., Hoskins, B. J., Isaksen, L., Janssen, P. A. E. M., Jenne, R., McNally,  
1452 A. P., Mahfouf, J.-F., Morcrette, J.-J., Rayner, N. A., Saunders, R. W., Simon,  
1453 P., Sterl, A., Trenberth, K. E., Untch, A., Vasiljevic, D., Viterbo, P., Woollen, J.,  
1454 2005. The ERA-40 reanalysis. *Q. J. R. Meteorol. Soc.* 131, 2961–3012.
- 1455 Wang, Q., Ilicak, M., Gerdes, R., Drange, H., Aksenov, Y., Bailey, D. A., Bentsen,  
1456 M., Biastoch, A., Bozec, A., Böning, C., Cassou, C., Chassignet, E., Coward,  
1457 A. C., Curry, B., Danabasoglu, G., Danilov, S., Fernandez, E., Fogli, P. G., Fujii,  
1458 Y., Griffies, S. M., Iovino, D., Jahn, A., Jung, T., Large, W. G., Lee, C., Lique,  
1459 C., Lu, J., Masina, S., Nurser, A. J. G., Rabe, B., Roth, C., Salas y Mélia, D.,  
1460 Samuels, B. L., Spence, P., Tsujino, H., Valcke, S., Voltaire, A., Wang, X., Yeager,  
1461 S. G., 2015. An assessment of the Arctic Ocean in a suite of interannual CORE-II  
1462 simulations: Sea ice and freshwater. *Ocean Modelling*(submitted).
- 1463 Yashayaev, I., 2007. Hydrographic changes in the Labrador Sea, 1960-2005. *Prog.*  
1464 *Oceanogr.* 73, 242–276.
- 1465 Yeager, S., 2015. Topographic coupling of the Atlantic overturning and gyre circula-  
1466 tions. *J. Phys. Oceanogr.* 45, 1258–1284.
- 1467 Yeager, S., Danabasoglu, G., 2014. The origins of late-twentieth-century variations  
1468 in the large-scale North Atlantic circulation. *J. Climate* 27, 3222–3247.



- 1469 Yeager, S., Karspeck, A., Danabasoglu, G., Tribbia, J., Teng, H., 2012. A decadal  
1470 prediction case study: Late Twentieth-century North Atlantic Ocean heat content.  
1471 J. Climate 25, 5173–5189.
- 1472 Yeager, S. G., Danabasoglu, G., 2012. Sensitivity of Atlantic meridional overturning  
1473 circulation variability to parameterized Nordic Sea overflows in CCSM4. J. Climate  
1474 25, 2077–2103.
- 1475 Zhang, R., 2010. Latitudinal dependence of Atlantic meridional overturning circula-  
1476 tion AMOC variations. Geophys. Res. Lett. 37, L16703.

1477 **List of Figures**

- 1478 1 AMOC annual-mean maximum transport time series at  $26.5^{\circ}\text{N}$  for the  
1479 1958–2007 period from the last cycle of simulations. The time series  
1480 are anomalies from the respective 50-year means given for each model  
1481 in parentheses in the labels. The thick gray lines represent the annual-  
1482 mean RAPID data from Cunningham et al. (2007). The 4-year mean  
1483 for the RAPID data is 18.6 Sv. MMM time series are included in all  
1484 panels as the dashed black lines. MMM does not include MRI-A. . . . 71
- 1485 2 AMOC annual-mean maximum transport time series at  $45^{\circ}\text{N}$  for the  
1486 1958–2007 period from the last cycle of simulations. The time series  
1487 are anomalies from the respective 50-year means given for each model  
1488 in parentheses in the labels. MMM time series are included in all  
1489 panels as the dashed black lines. MMM does not include MRI-A. . . . 72
- 1490 3 Model – model correlations for the AMOC maximum transport time  
1491 series at (a-c)  $26.5^{\circ}\text{N}$  and (d-f)  $45^{\circ}\text{N}$ . (left column) High-pass filtered;  
1492 (middle column) Low-pass filtered with trend; and (right column)  
1493 Low-pass filtered and detrended. A 7-year cutoff is used for the filters.  
1494 AMOC in depth and latitude space is used for the 1958–2007 period.  
1495 All negative correlations are included in the darkest blue color. . . . . 73

1496	4	AMOC EOF1 spatial distributions in depth (km) and latitude space	
1497		for the 1958–2007 period. The associated variances accounted by	
1498		EOF1 as a percentage of the total AMOC variance are also given.	
1499		The positive and negative contours indicate clockwise and counter-	
1500		clockwise circulations, respectively. In MIT, AWI, MRI-F, MRI-A,	
1501		FSU, BERGEN, GISS, GISS2, and FSU2, the AMOC distributions do	
1502		not include the high latitude North Atlantic and / or Arctic Oceans,	
1503		and hence are masked. No detrending is applied. . . . .	74
1504	5	AMOC PC1 time series corresponding to Fig. 4. The time series are	
1505		normalized to have unit variance, so that the EOF spatial pattern	
1506		magnitudes correspond to one standard deviation changes in the time	
1507		series. . . . .	75
1508	6	AMOC EOF1 spatial distributions in $\sigma_2$ ( $\text{kg m}^{-3}$ ) and latitude space	
1509		for the 1958–2007 period. The associated variances accounted by	
1510		EOF1 as a percentage of the total AMOC variance are also given.	
1511		The positive and negative contours indicate clockwise and counter-	
1512		clockwise circulations, respectively. INMOM distribution is not avail-	
1513		able. No detrending is applied. . . . .	76
1514	7	SST EOF1 spatial distributions for the 1958–2007 period for the	
1515		North Atlantic. The associated variances accounted by EOF1 as a	
1516		percentage of the total SST variance are also given. The panel to the	
1517		left of the color bar shows SST EOF1 calculated from the HadISST	
1518		dataset. No detrending is applied. . . . .	77

1519	8	SST PC1 time series corresponding to Fig. 7. The time series are	
1520		normalized to have unit variance, so that the EOF spatial pattern	
1521		magnitudes correspond to one standard deviation changes in the time	
1522		series. The time series from the HadISST dataset are included in all	
1523		panels as the black lines. . . . .	78
1524	9	March-mean MLD EOF1 spatial distributions for the 1958–2007 pe-	
1525		riod for the North Atlantic. The associated variances accounted by	
1526		EOF1 as a percentage of the total MLD variance are also given. MLD	
1527		is based on a $\Delta\rho = 0.125 \text{ kg m}^{-3}$ criterion. No detrending is applied.	
1528		The interior white areas (i.e., excluding west of $80^\circ\text{W}$ and east of	
1529		$10^\circ\text{E}$ ) indicate regions of no variability as the time-mean MLDs reach	
1530		the ocean bottom in some models. . . . .	79
1531	10	March-mean MLD PC1 time series corresponding to Fig. 9. The time	
1532		series are normalized to have unit variance, so that the EOF spatial	
1533		pattern magnitudes correspond to one standard deviation changes in	
1534		the time series. . . . .	80

1535	11	Time series of potential temperature anomalies averaged over the	
1536		150–1000 m depth range and within a central Labrador Sea region	
1537		bounded by $49^{\circ} - 56^{\circ}\text{W}$ and $56^{\circ} - 61^{\circ}\text{N}$ . The anomalies are with re-	
1538		spect to the 1958–2007 period. The black lines show the observational	
1539		data from Yashayaev (2007) with data missing for some years. May-	
1540		mean output from the models is used to roughly match the mostly	
1541		Spring-time observations. For each model, the first number in paren-	
1542		theses gives the root-mean-square model – observations difference of	
1543		their time series while the second number is the correlation coefficient	
1544		between the model and observational time series. Data from ACCESS,	
1545		FSU, GISS, GISS2, KIEL, and MIT are not available. . . . .	81
1546	12	Same as in Fig. 11, but for salinity anomalies. . . . .	82
1547	13	Same as in Fig. 11, but for density anomalies based on $\sigma_0$ . . . . .	83
1548	14	Time series of SPG SSH anomalies with respect to the 1993–2007	
1549		mean. SSH time series represent averages for the SPG region defined	
1550		as the area between $15^{\circ} - 60^{\circ}\text{W}$ and $48^{\circ} - 65^{\circ}\text{N}$ . The SSH anomaly	
1551		time series from AVISO dataset are also shown in each panel. The	
1552		AVISO time series include the ranges of the spatially- and annually-	
1553		averaged standard errors based on the monthly-mean data. The first	
1554		number in parentheses for each model gives the correlation coefficient	
1555		between the AVISO and that model’s SSH time series. The second	
1556		number in parentheses and the number for AVISO show the linear	
1557		trend for the 1993–2007 period in $\text{cm yr}^{-1}$ . . . . .	84

1558 15 Low-pass filtered, MMM time series of (top) AMOC maximum trans-  
1559 port at  $45^{\circ}\text{N}$ , March-mean MLD, and SPG BSF; and (bottom) AMOC  
1560 maximum transport at  $45^{\circ}\text{N}$  (same as in the top panel), AMOC max-  
1561 imum transport at  $26.5^{\circ}\text{N}$ , and SPG SSH. The top panel also includes  
1562 low-pass filtered NAO time series whose amplitude is multiplied by a  
1563 factor of two for clarity. MLD is calculated as an average for the LS  
1564 – Irminger Sea region defined as the area between  $15^{\circ} - 60^{\circ}\text{W}$  and  
1565  $48^{\circ} - 60^{\circ}\text{N}$ . The SPG BSF and SSH represent averages for the SPG  
1566 region defined by  $15^{\circ} - 60^{\circ}\text{W}$  and  $48^{\circ} - 65^{\circ}\text{N}$ . We note that negative  
1567 SPG BSF and SSH anomalies indicate strengthening of the cyclonic  
1568 SPG circulation. All time series are anomalies with respect to the  
1569 1958–2007 period. A 7-year cutoff is used for the low-pass filter. The  
1570 respective colored shadings denote one standard deviation spread of  
1571 the models' time series from those of the respective MMM. The spread  
1572 for the AMOC transport at  $45^{\circ}\text{N}$  is not repeated in the bottom panel  
1573 for clarity. MMM does not include MRI-A. Units are Sv for AMOC  
1574 and BSF;  $\times 100$  m for MLD; and cm for SSH. . . . . 85

- 1575 16 Low-pass filtered AMOC maximum transport at 45°N time series cor-  
 1576 relations with (first column) March-mean MLD, (second column) SPG  
 1577 BSF, (third column) SPG SSH, and (fourth column) NAO. The black  
 1578 lines in each panel show the MMM correlation functions evaluated as  
 1579 the mean of the individual model correlations. MMM does not include  
 1580 MRI-A. The correlations outside the shaded regions have confidence  
 1581 levels greater than 95% (see section 2 for calculation of confidence lev-  
 1582 els). Anomalies are with respect to the 1958–2007 period. A 7-year  
 1583 cutoff is used for the low-pass filter. AMOC index leads for positive  
 1584 lags. . . . . 86
- 1585 17 Years 1988–2007 mean AMOC plotted in depth (km) and latitude  
 1586 space from FSU, FSU2, and GISS2. The positive and negative con-  
 1587 tours indicate clockwise and counter-clockwise circulations, respectively. 87
- 1588 18 (a) Years 2004–2007 mean AMOC depth profiles at 26.5°N from FSU,  
 1589 FSU2, and GISS2 in comparison with the 4-year mean (April 2004 –  
 1590 March 2008) RAPID data; (b) Years 1988–2007 mean meridional heat  
 1591 transports for the Atlantic Ocean from the three models. In (b), the  
 1592 black line denoted by L&Y09 represents implied time-mean transport  
 1593 calculated by Large and Yeager (2009) with shading showing the im-  
 1594 plied transport range in individual years for the 1984–2006 period.  
 1595 Direct estimates with their uncertainty ranges from the RAPID data  
 1596 (square; Johns et al., 2011) and from Bryden and Imawaki (2001)  
 1597 (triangle; B&I01) are also shown. . . . . 88

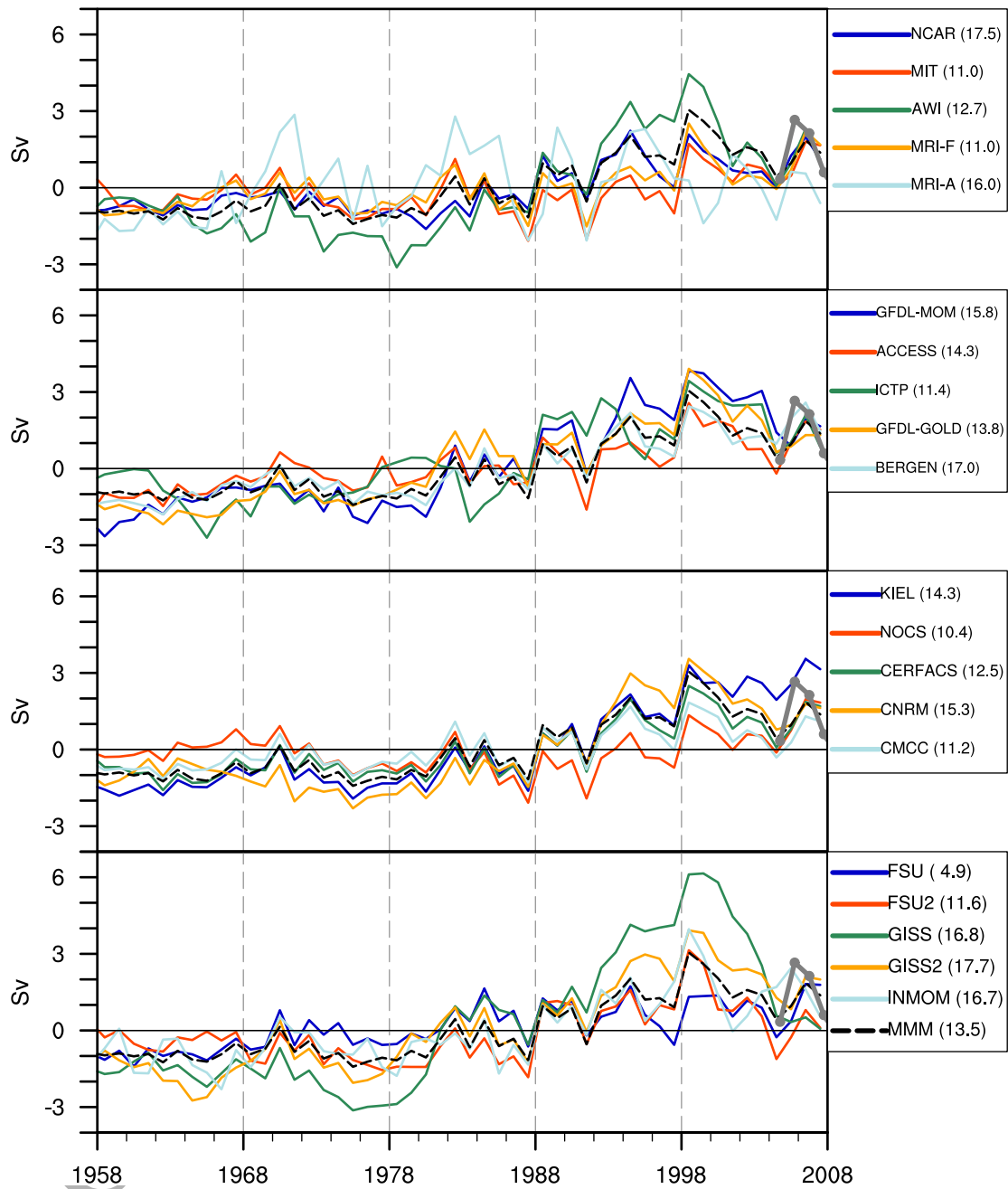


Figure 1: AMOC annual-mean maximum transport time series at  $26.5^{\circ}\text{N}$  for the 1958–2007 period from the last cycle of simulations. The time series are anomalies from the respective 50-year means given for each model in parentheses in the labels. The thick gray lines represent the annual-mean RAPID data from Cunningham et al. (2007). The 4-year mean for the RAPID data is 18.6 Sv. MMM time series are included in all panels as the dashed black lines. MMM does not include MRI-A.



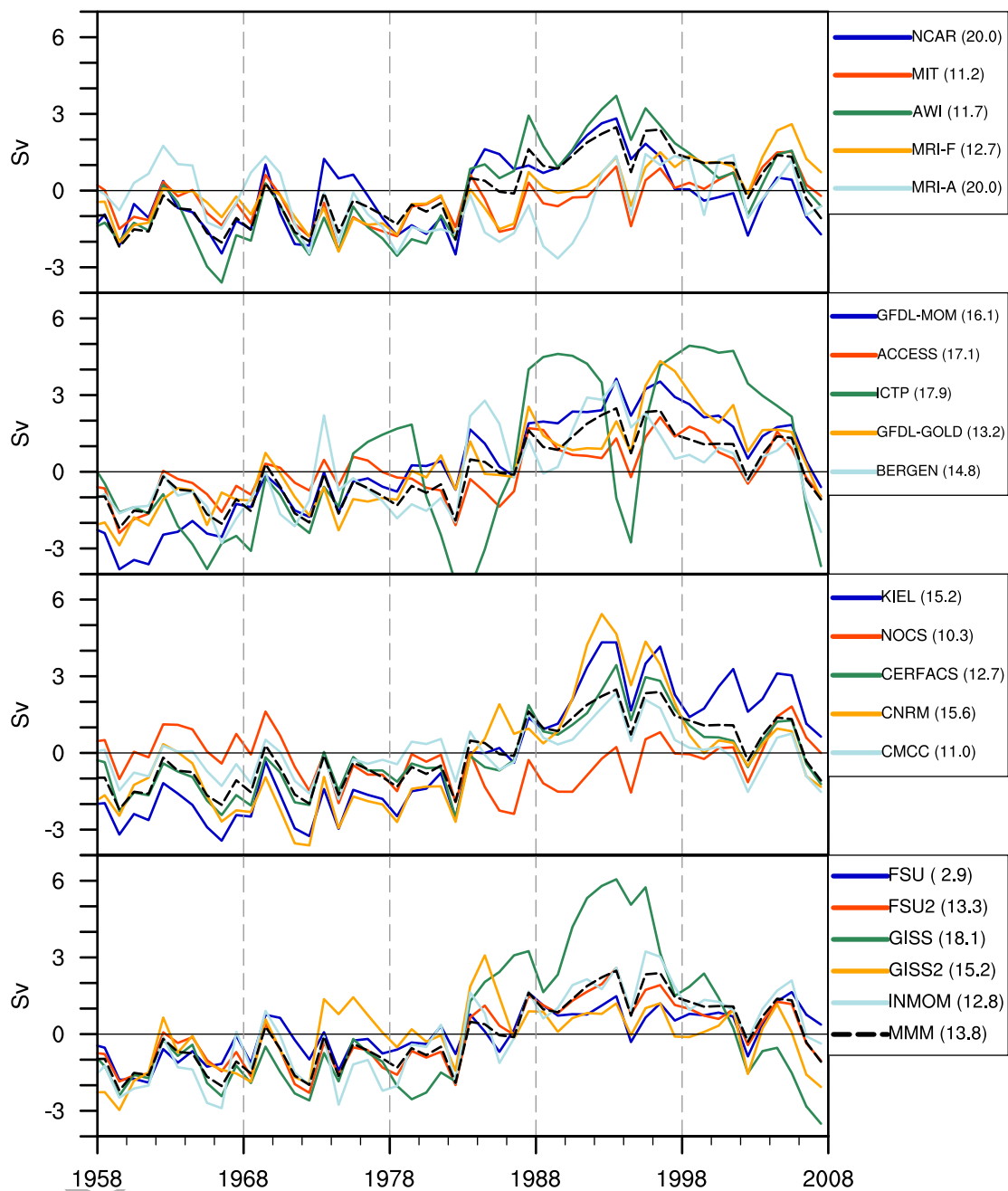


Figure 2: AMOC annual-mean maximum transport time series at 45°N for the 1958–2007 period from the last cycle of simulations. The time series are anomalies from the respective 50-year means given for each model in parentheses in the labels. MMM time series are included in all panels as the dashed black lines. MMM does not include MRI-A.

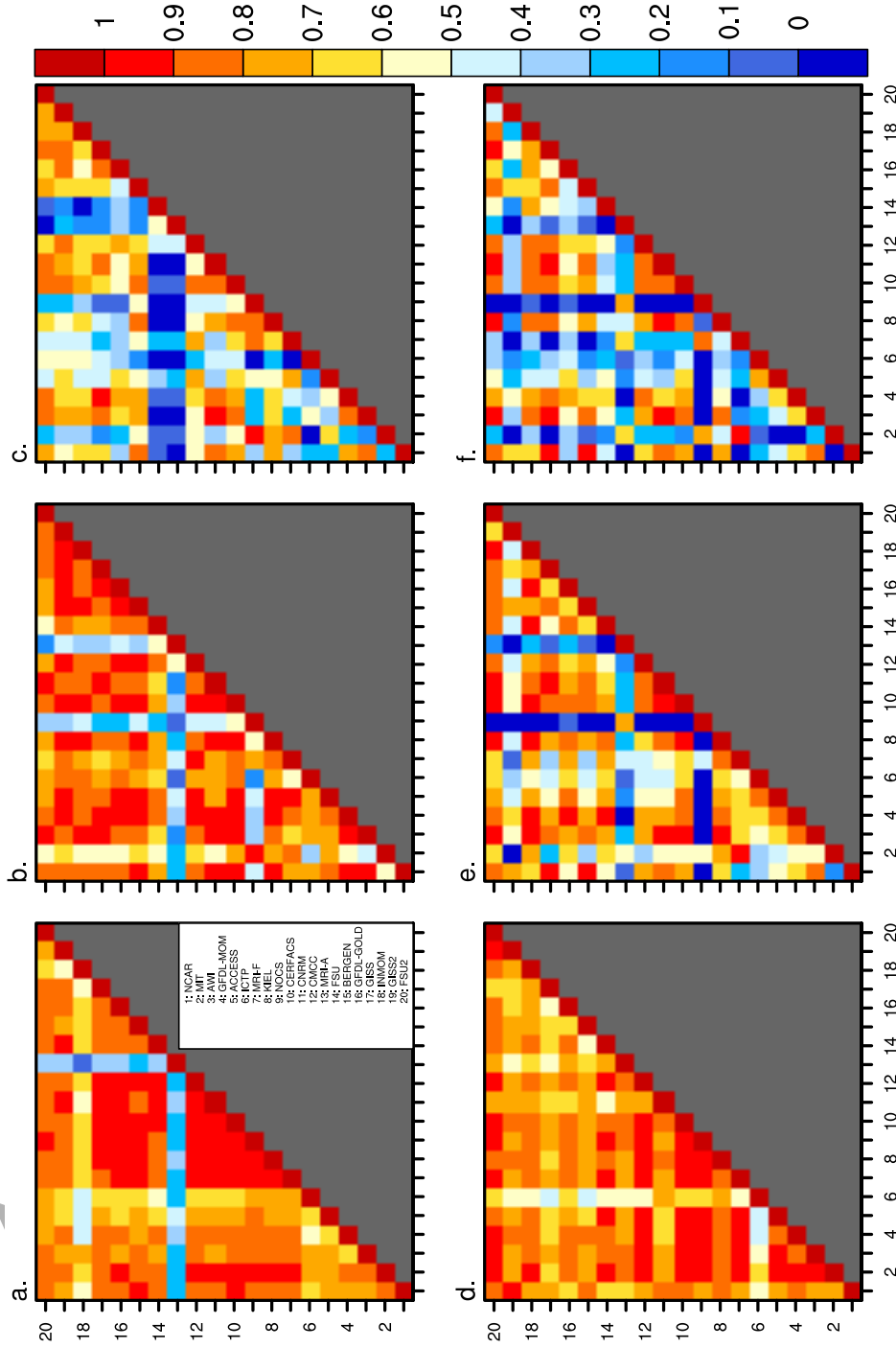


Figure 3: Model – model correlations for the AMOC maximum transport time series at (a-c) 26.5°N and (d-f) 45°N. (left column) High-pass filtered; (middle column) Low-pass filtered with trend; and (right column) Low-pass filtered and detrended. A 7-year cutoff is used for the filters. AMOC in depth and latitude space is used for the 1958–2007 period. All negative correlations are included in the darkest blue color.

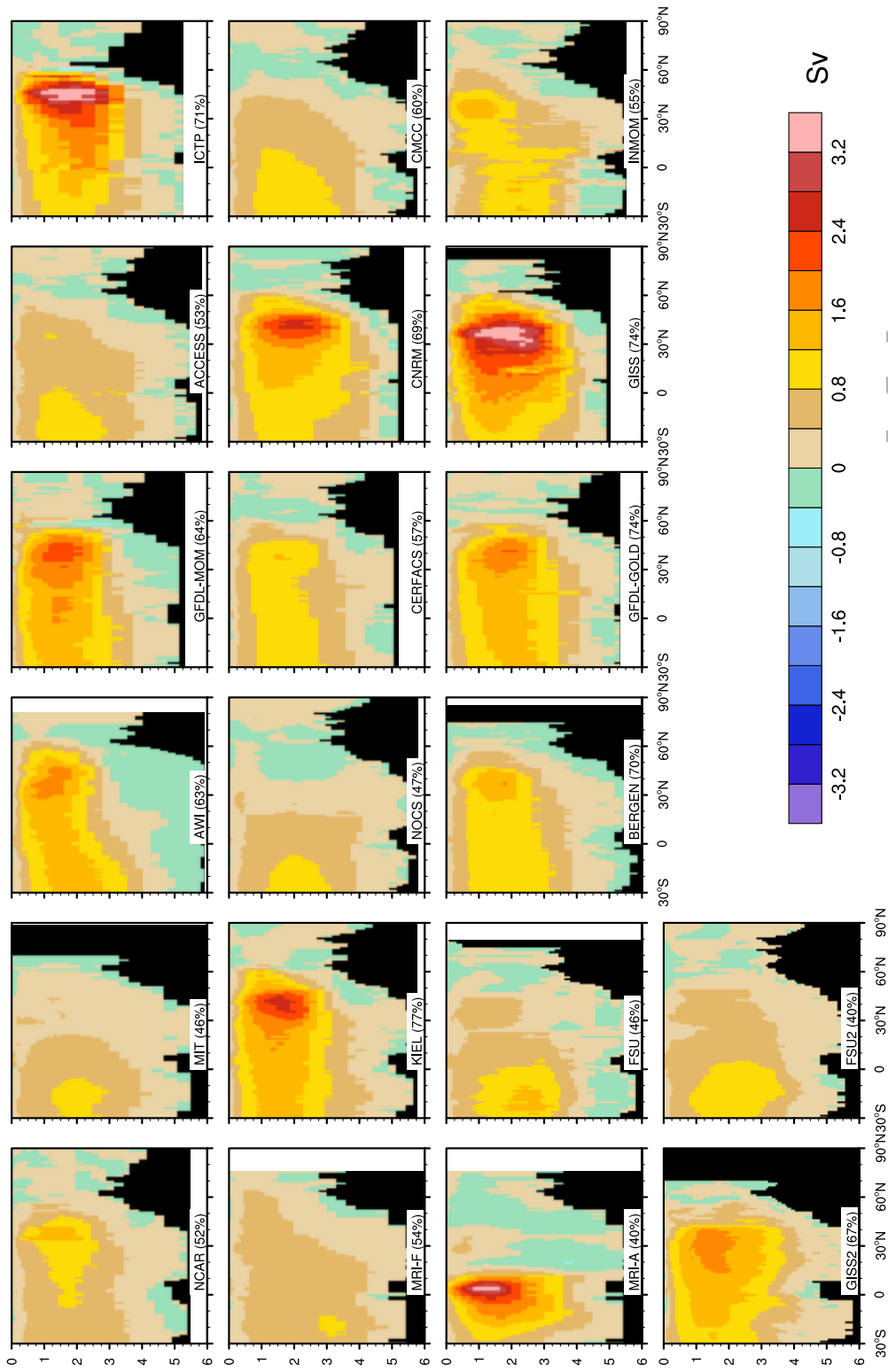


Figure 4: AMOC EOF1 spatial distributions in depth (km) and latitude space for the 1958–2007 period. The associated variances accounted by EOF1 as a percentage of the total AMOC variance are also given. The positive and negative contours indicate clockwise and counter-clockwise circulations, respectively. In MIT, AWI, MRI-F, MRI-A, FSU, BERGEN, GISS, GISS2, and FSU2, the AMOC distributions do not include the high latitude North Atlantic and / or Arctic Oceans, and hence are masked. No detrending is applied.

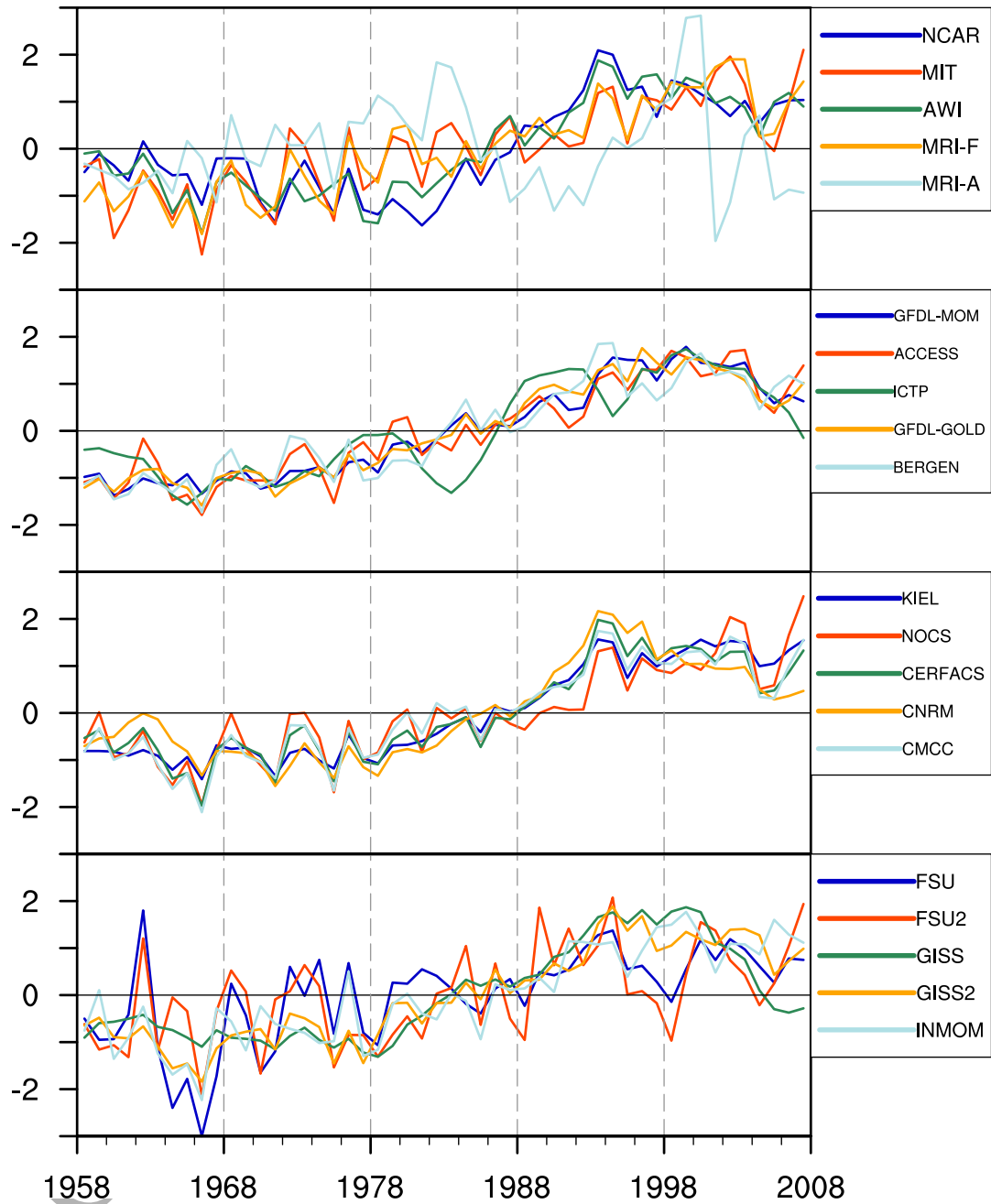


Figure 5: AMOC PC1 time series corresponding to Fig. 4. The time series are normalized to have unit variance, so that the EOF spatial pattern magnitudes correspond to one standard deviation changes in the time series.

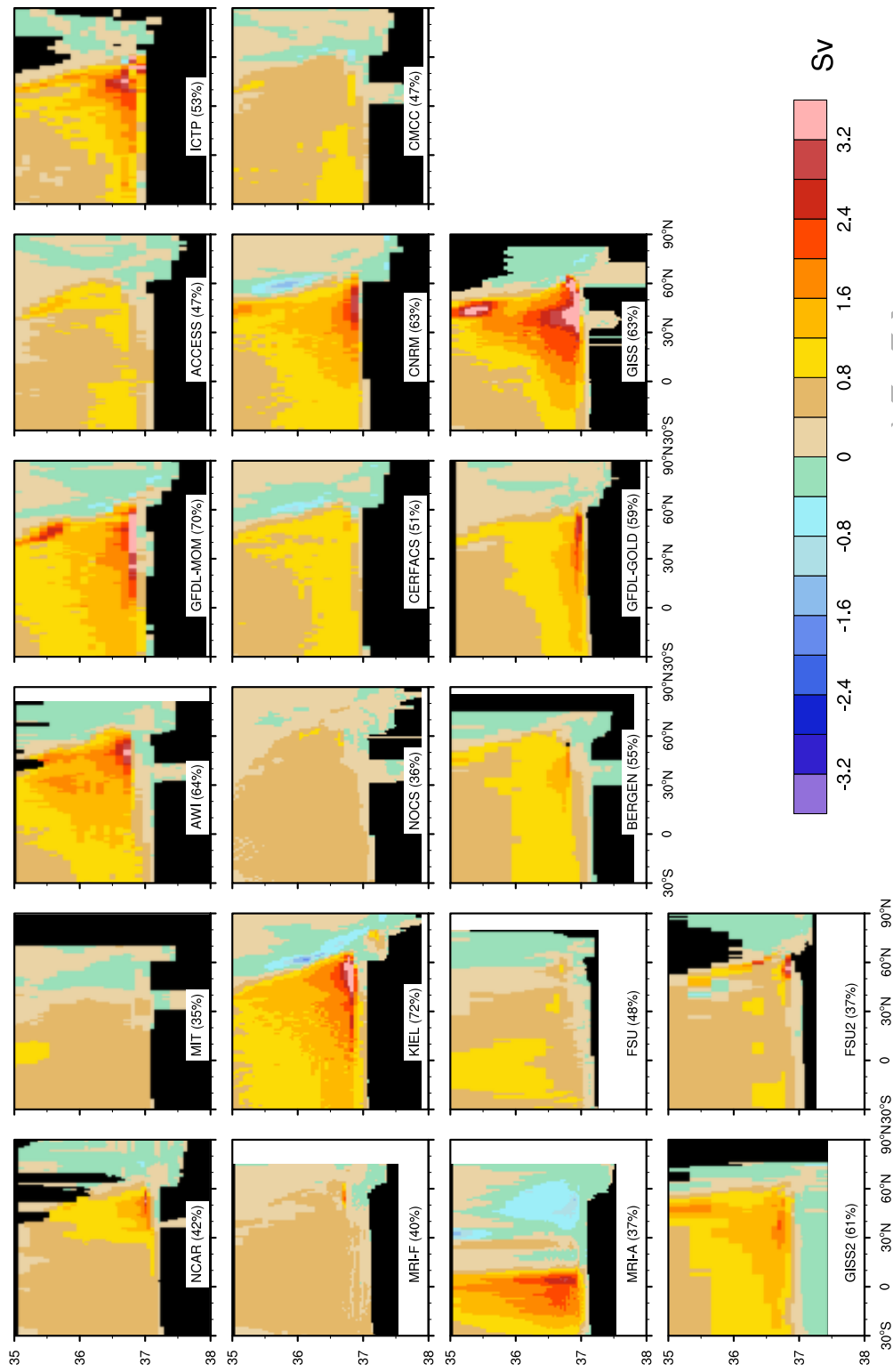


Figure 6: AMOC EOF1 spatial distributions in  $\sigma_2$  ( $\text{kg m}^{-3}$ ) and latitude space for the 1958–2007 period. The associated variances accounted by EOF1 as a percentage of the total AMOC variance are also given. The positive and negative contours indicate clockwise and counter-clockwise circulations, respectively. INMOM distribution is not available. No detrending is applied.

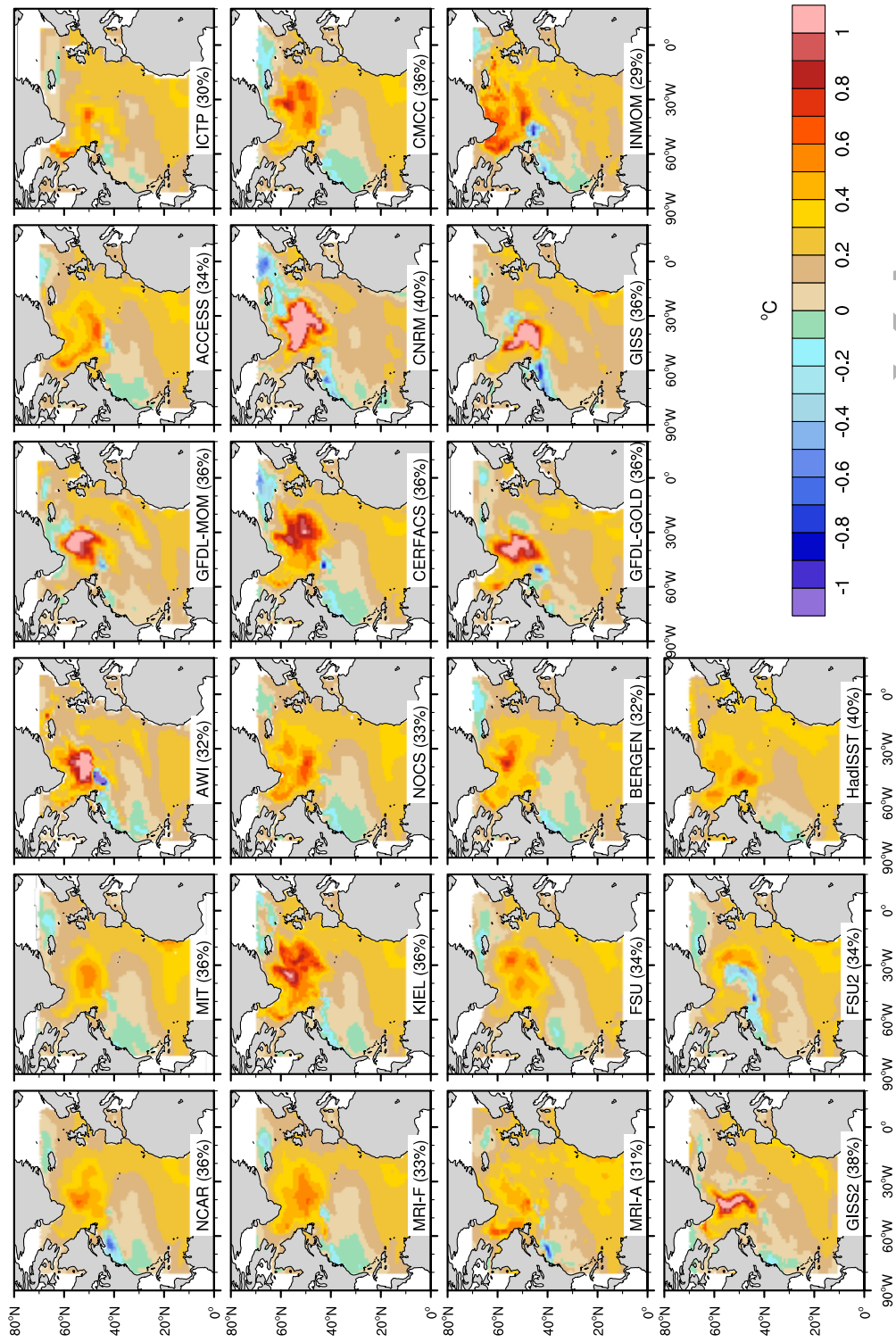


Figure 7: SST EOF1 spatial distributions for the 1958–2007 period for the North Atlantic. The associated variances accounted by EOF1 as a percentage of the total SST variance are also given. The panel to the left of the color bar shows SST EOF1 calculated from the HadISST dataset. No detrending is applied.

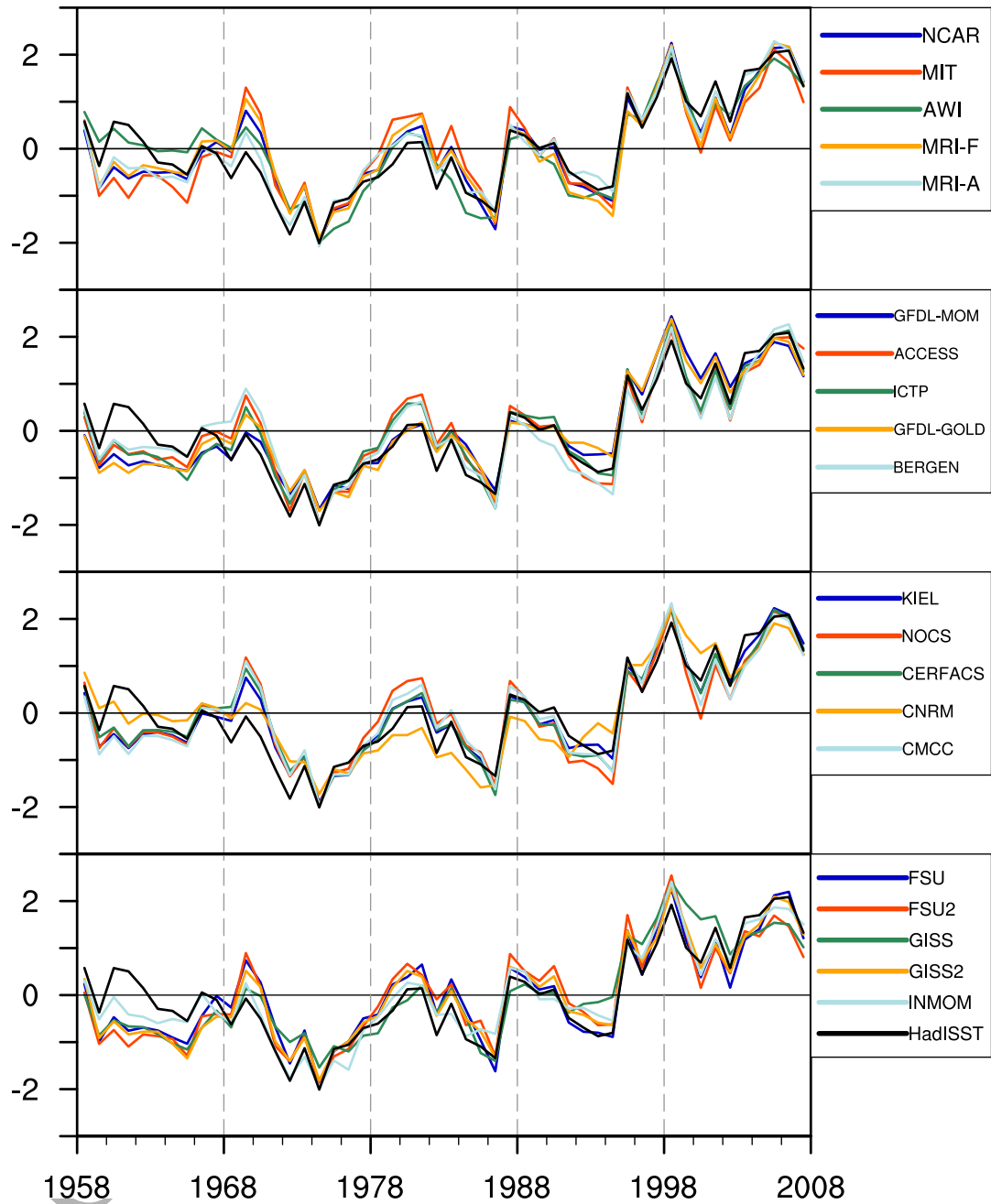


Figure 8: SST PC1 time series corresponding to Fig. 7. The time series are normalized to have unit variance, so that the EOF spatial pattern magnitudes correspond to one standard deviation changes in the time series. The time series from the HadISST dataset are included in all panels as the black lines.

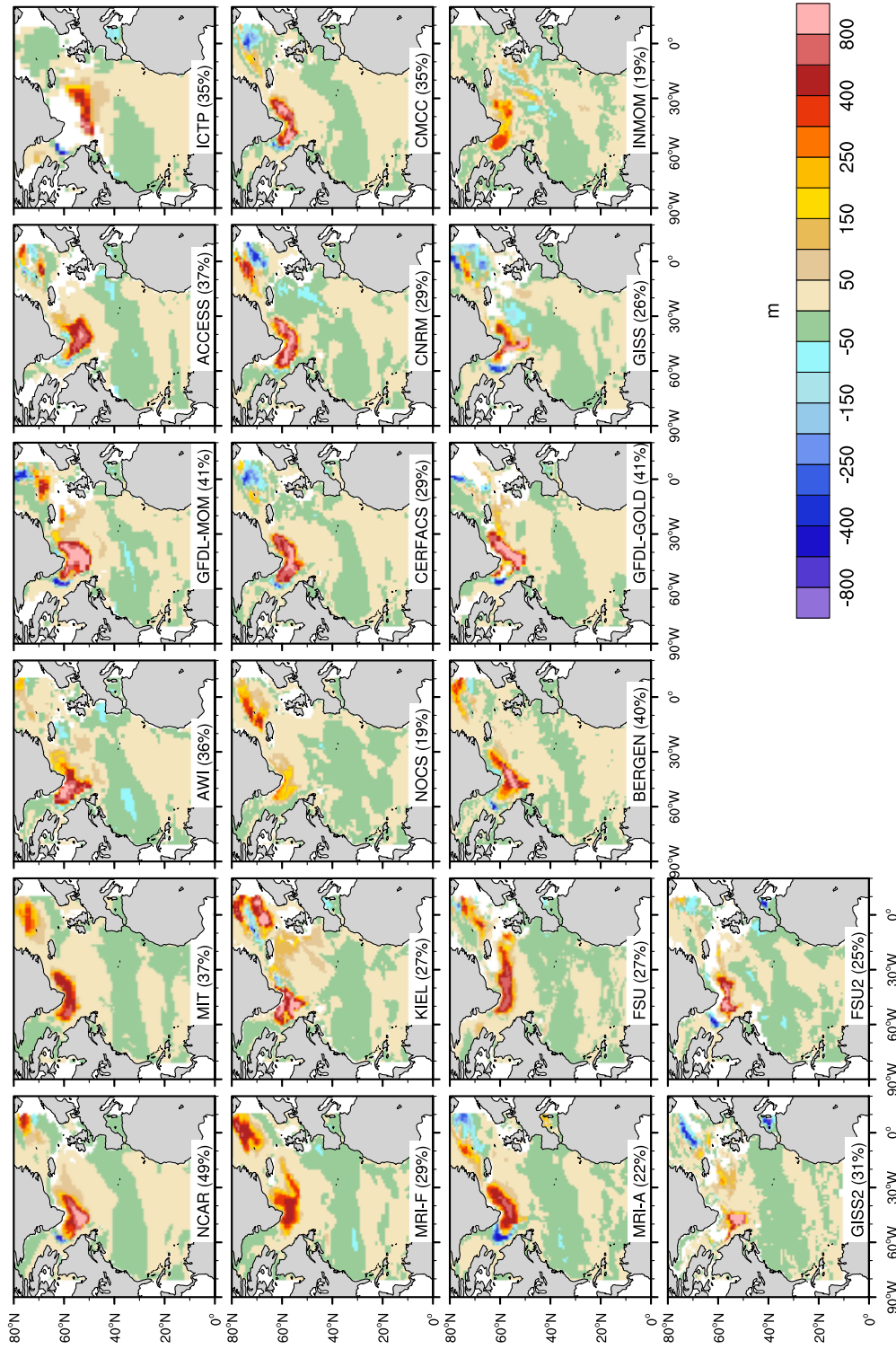


Figure 9: March-mean MLD EOF1 spatial distributions for the 1958–2007 period for the North Atlantic. The associated variances accounted by EOF1 as a percentage of the total MLD variance are also given. MLD is based on a  $\Delta\rho = 0.125 \text{ kg m}^{-3}$  criterion. No detrending is applied. The interior white areas (i.e., excluding west of  $80^\circ\text{W}$  and east of  $10^\circ\text{E}$ ) indicate regions of no variability as the time-mean MLDs reach the ocean bottom in some models.



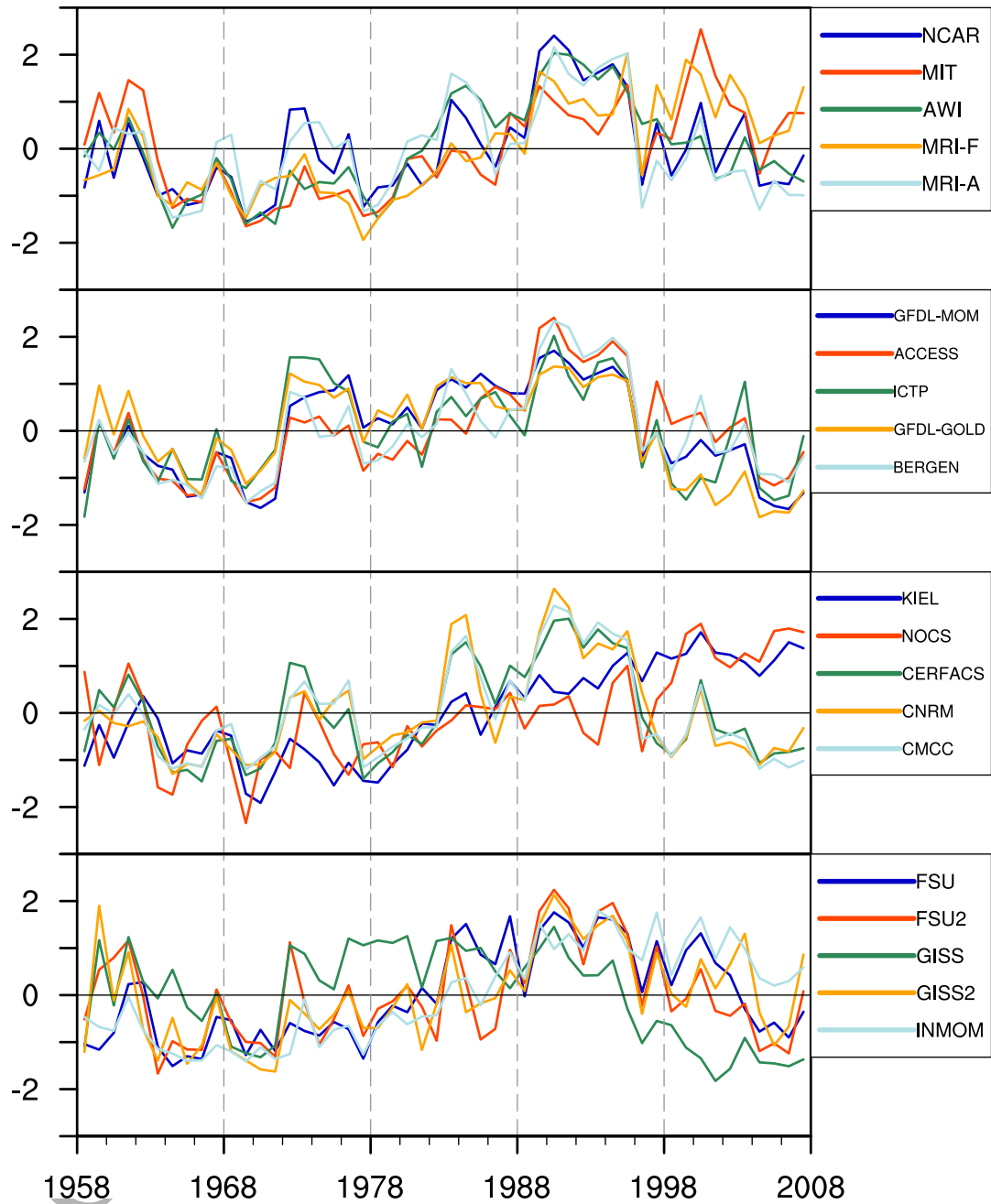


Figure 10: March-mean MLD PC1 time series corresponding to Fig. 9. The time series are normalized to have unit variance, so that the EOF spatial pattern magnitudes correspond to one standard deviation changes in the time series.

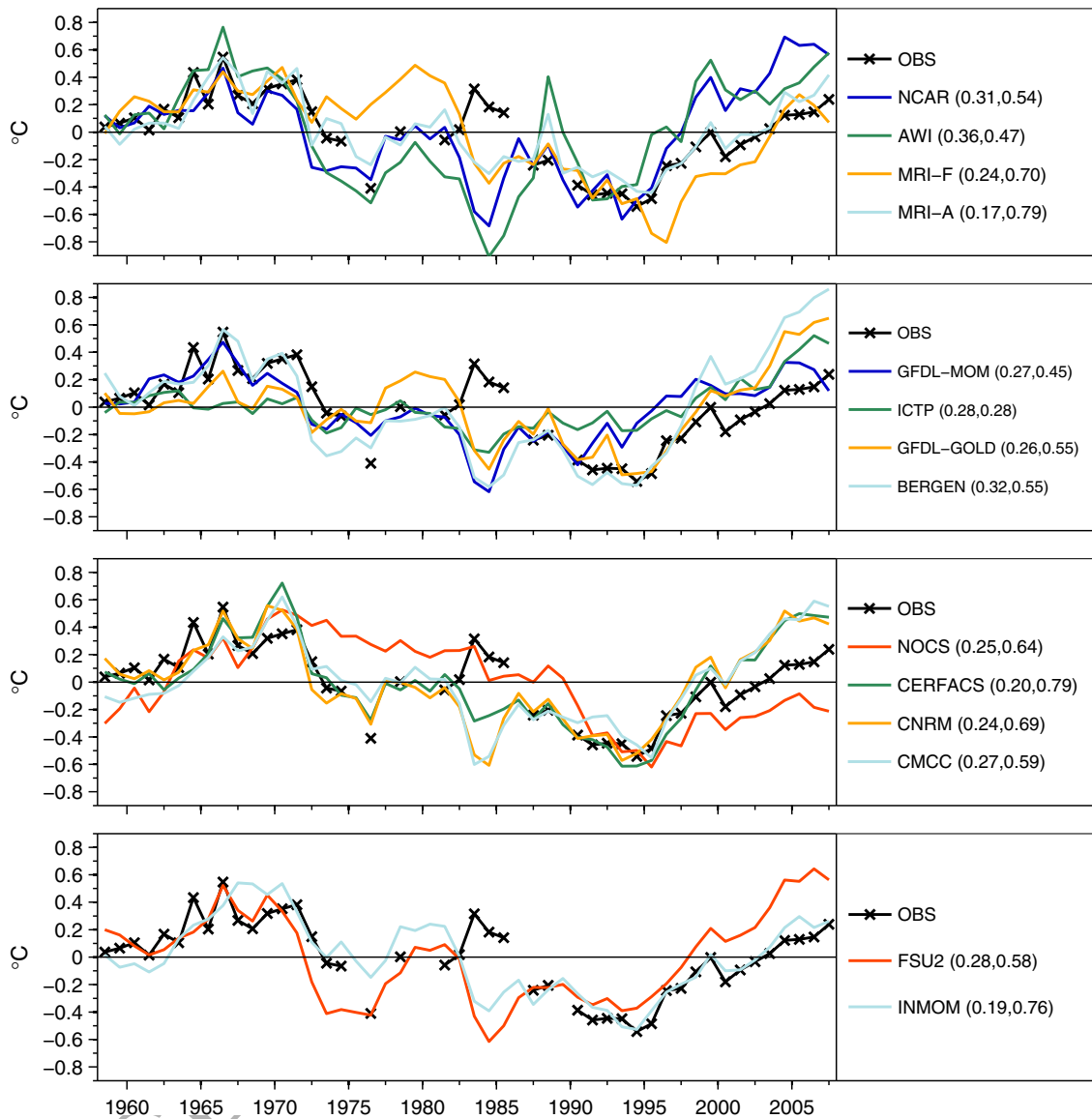


Figure 11: Time series of potential temperature anomalies averaged over the 150–1000 m depth range and within a central Labrador Sea region bounded by  $49^{\circ} - 56^{\circ}\text{W}$  and  $56^{\circ} - 61^{\circ}\text{N}$ . The anomalies are with respect to the 1958–2007 period. The black lines show the observational data from Yashayaev (2007) with data missing for some years. May-mean output from the models is used to roughly match the mostly Spring-time observations. For each model, the first number in parentheses gives the root-mean-square model – observations difference of their time series while the second number is the correlation coefficient between the model and observational time series. Data from ACCESS, FSU, GISS, GISS2, KIEL, and MIT are not available.

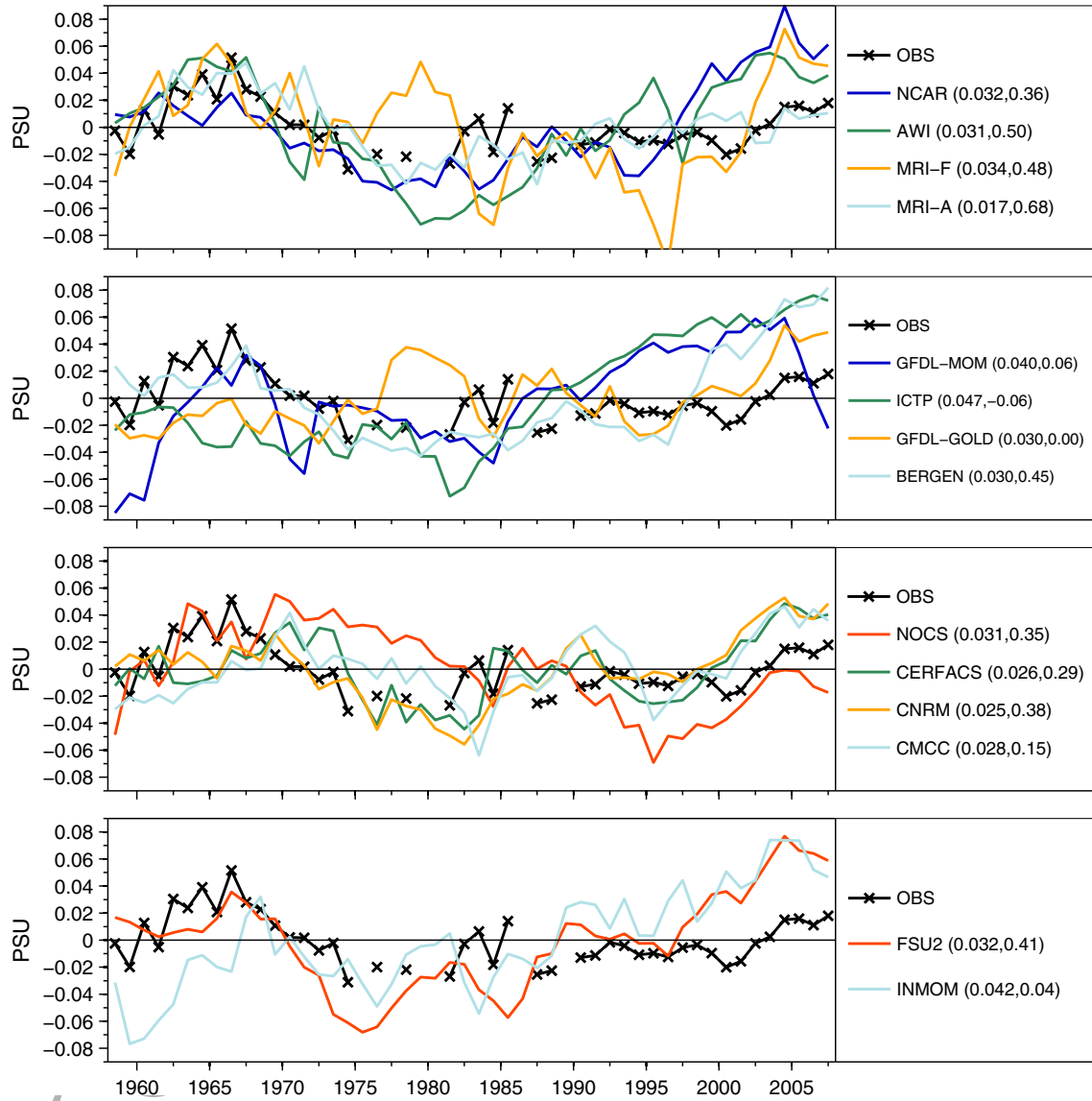


Figure 12: Same as in Fig. 11, but for salinity anomalies.

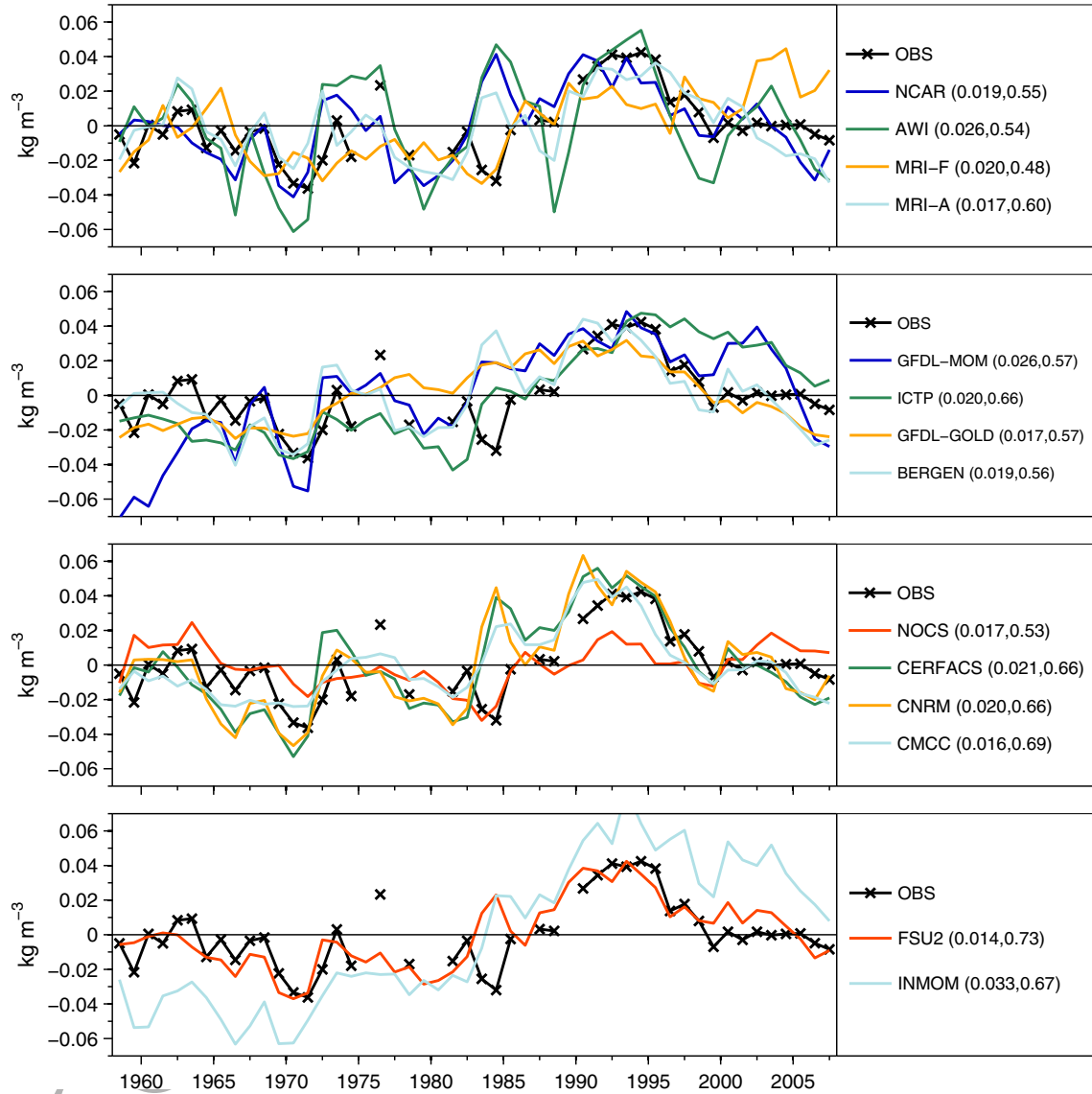


Figure 13: Same as in Fig. 11, but for density anomalies based on  $\sigma_0$ .

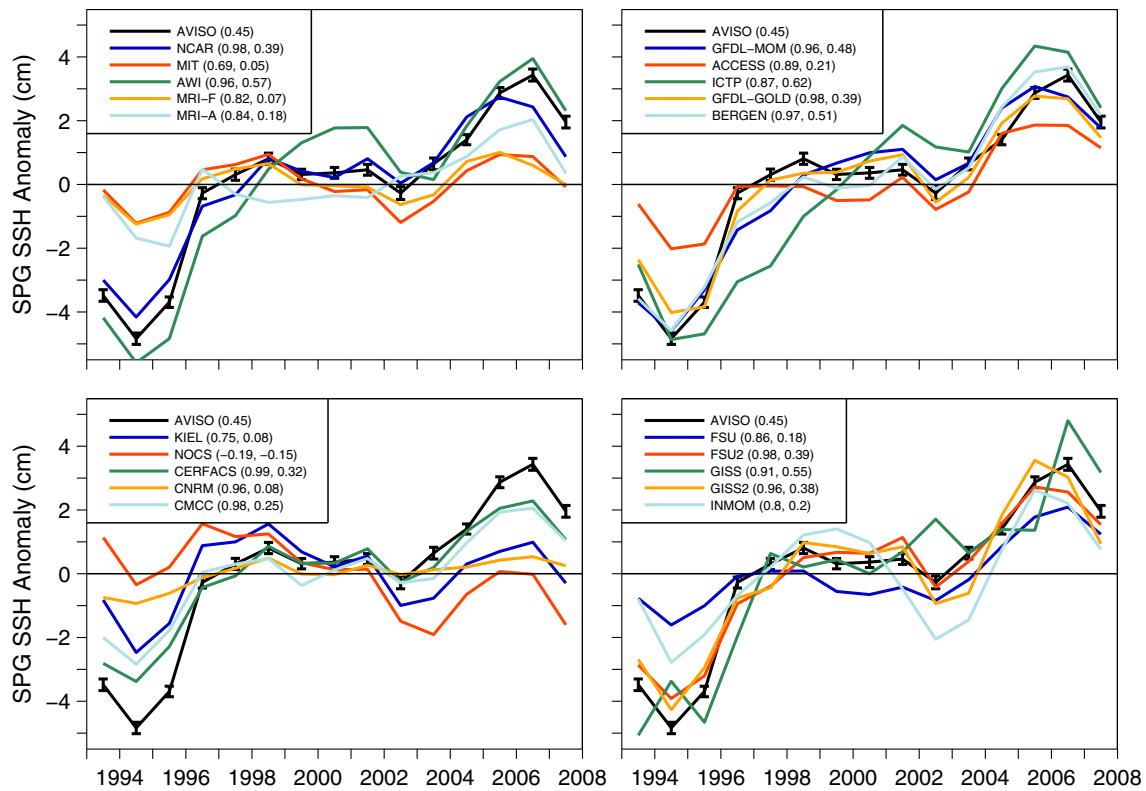


Figure 14: Time series of SPG SSH anomalies with respect to the 1993–2007 mean. SSH time series represent averages for the SPG region defined as the area between  $15^{\circ} - 60^{\circ}\text{W}$  and  $48^{\circ} - 65^{\circ}\text{N}$ . The SSH anomaly time series from AVISO dataset are also shown in each panel. The AVISO time series include the ranges of the spatially- and annually-averaged standard errors based on the monthly-mean data. The first number in parentheses for each model gives the correlation coefficient between the AVISO and that model's SSH time series. The second number in parentheses and the number for AVISO show the linear trend for the 1993–2007 period in  $\text{cm yr}^{-1}$ .

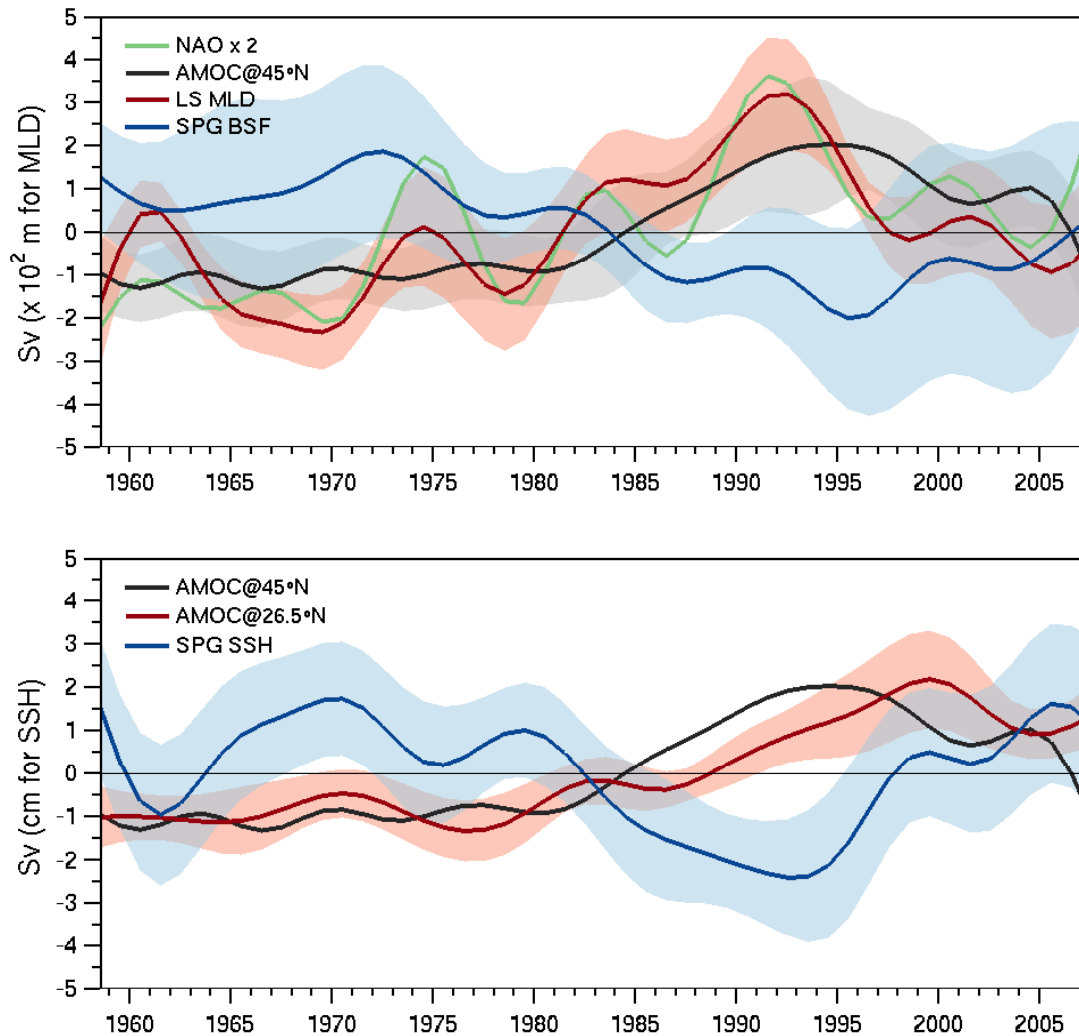


Figure 15: Low-pass filtered, MMM time series of (top) AMOC maximum transport at 45°N, March-mean MLD, and SPG BSF; and (bottom) AMOC maximum transport at 45°N (same as in the top panel), AMOC maximum transport at 26.5°N, and SPG SSH. The top panel also includes low-pass filtered NAO time series whose amplitude is multiplied by a factor of two for clarity. MLD is calculated as an average for the LS – Irminger Sea region defined as the area between 15° – 60°W and 48° – 60°N. The SPG BSF and SSH represent averages for the SPG region defined by 15° – 60°W and 48° – 65°N. We note that negative SPG BSF and SSH anomalies indicate strengthening of the cyclonic SPG circulation. All time series are anomalies with respect to the 1958–2007 period. A 7-year cutoff is used for the low-pass filter. The respective colored shadings denote one standard deviation spread of the models' time series from those of the respective MMM. The spread for the AMOC transport at 45°N is not repeated in the bottom panel for clarity. MMM does not include MRI-A. Units are Sv for AMOC and BSF;  $\times 100$  m for MLD; and cm for SSH.

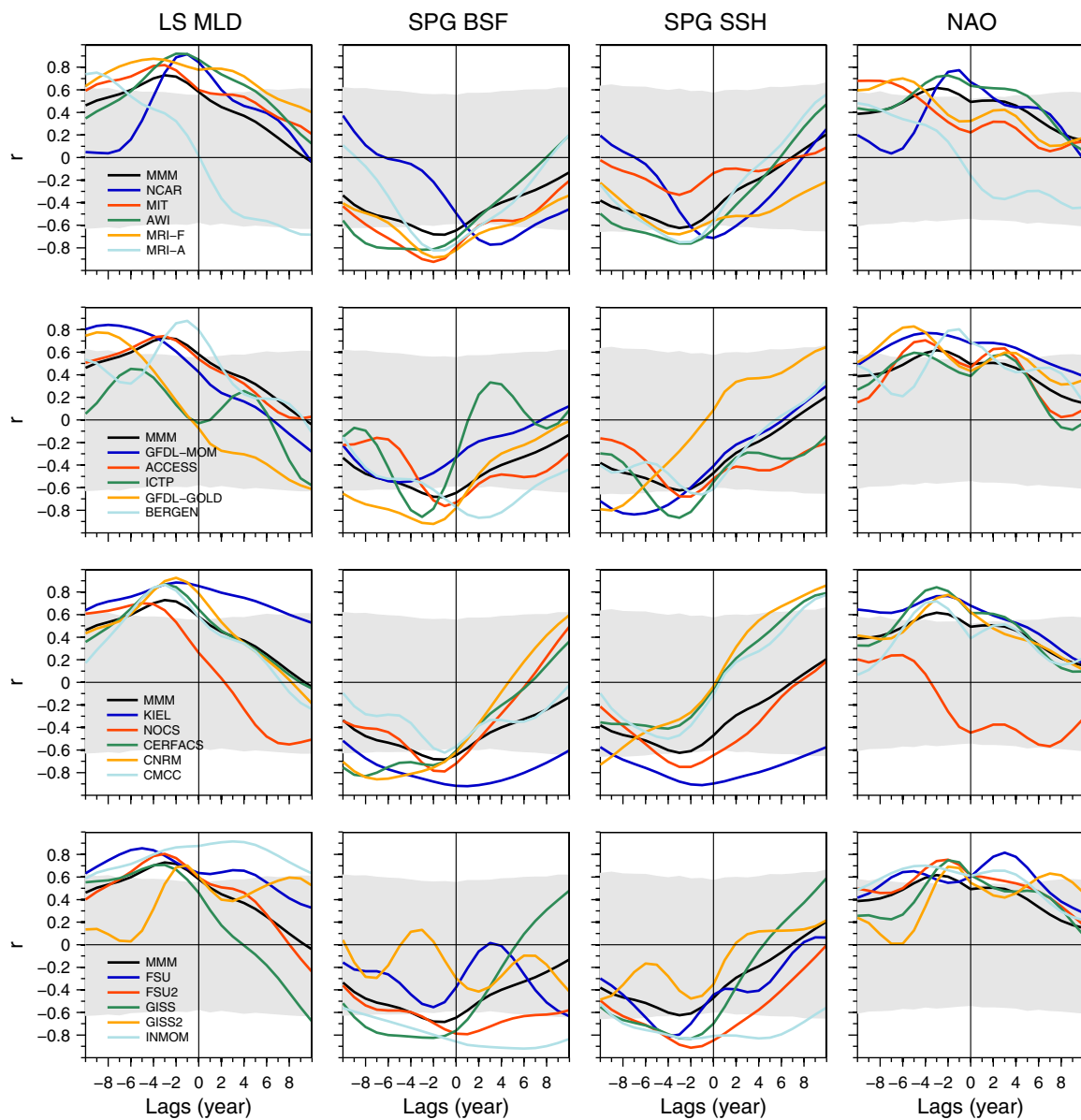


Figure 16: Low-pass filtered AMOC maximum transport at  $45^{\circ}\text{N}$  time series correlations with (first column) March-mean MLD, (second column) SPG BSF, (third column) SPG SSH, and (fourth column) NAO. The black lines in each panel show the MMM correlation functions evaluated as the mean of the individual model correlations. MMM does not include MRI-A. The correlations outside the shaded regions have confidence levels greater than 95% (see section 2 for calculation of confidence levels). Anomalies are with respect to the 1958–2007 period. A 7-year cutoff is used for the low-pass filter. AMOC index leads for positive lags.

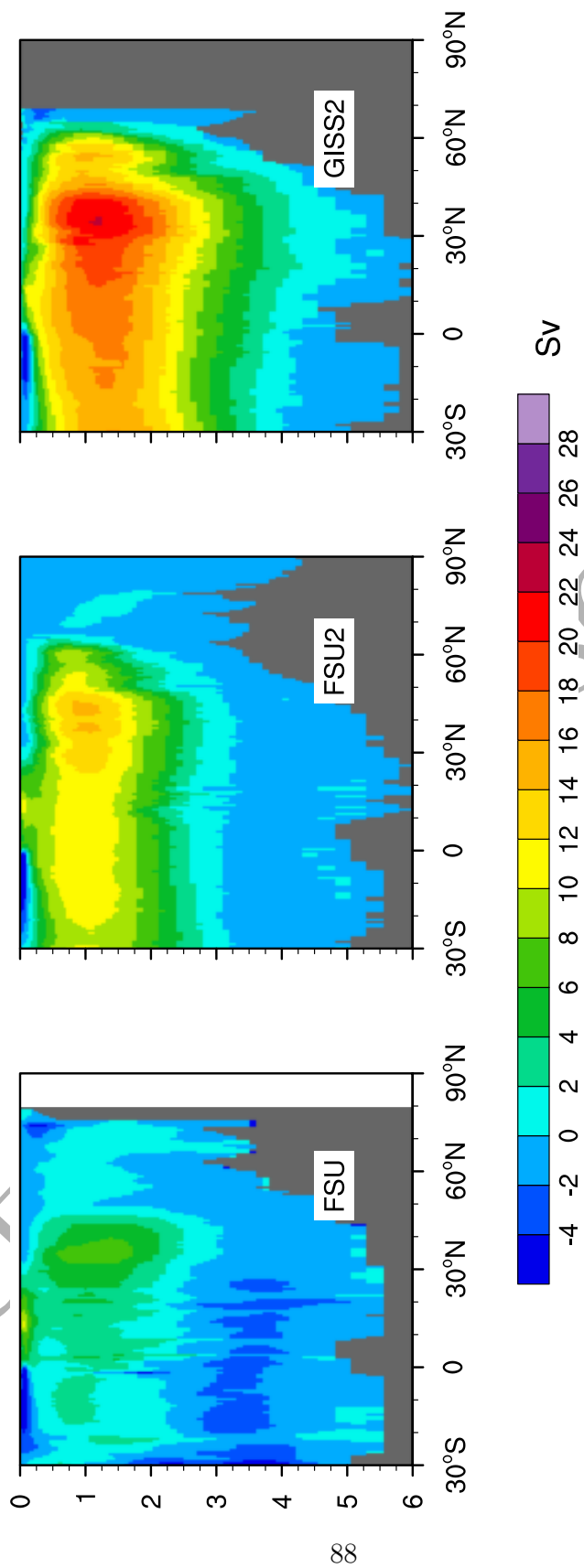


Figure 17: Years 1988–2007 mean AMOC plotted in depth (km) and latitude space from FSU, FSU2, and GISS2. The positive and negative contours indicate clockwise and counter-clockwise circulations, respectively.



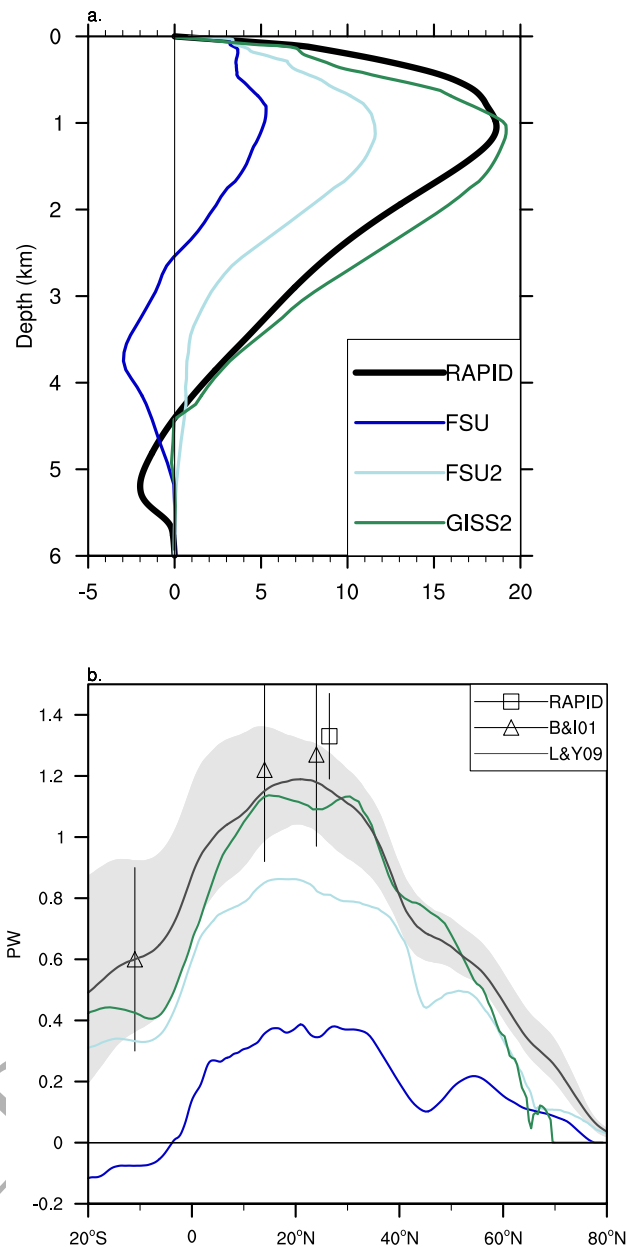


Figure 18: (a) Years 2004–2007 mean AMOC depth profiles at 26.5°N from FSU, FSU2, and GISS2 in comparison with the 4-year mean (April 2004 – March 2008) RAPID data; (b) Years 1988–2007 mean meridional heat transports for the Atlantic Ocean from the three models. In (b), the black line denoted by L&Y09 represents implied time-mean transport calculated by Large and Yeager (2009) with shading showing the implied transport range in individual years for the 1984–2006 period. Direct estimates with their uncertainty ranges from the RAPID data (square; Johns et al., 2011) and from Bryden and Imawaki (2001) (triangle; B&I01) are also shown.

Table 1: Summary of the ocean and sea-ice models in alphabetical order according to the participating group name (first column). The table includes the name of the combined ocean – sea-ice configuration (if any); the ocean model name and its version; the sea-ice model name and its version; vertical coordinate and number of layers / levels in parentheses; orientation of the horizontal grid with respect to the North Pole / Arctic; the number of horizontal grid cells (longitude  $\times$  latitude); and the horizontal resolution (longitude  $\times$  latitude). In MRI-A and MRI-F, the vertical levels shallower than 32 m follow the surface topography as in sigma-coordinate models. In FESOM, the total number of surface nodes is given under horizontal grid, because it has an unstructured grid. FESIM is imbedded in FESOM. H79 is Hibler (1979) and MK89 is Mellor and Kantha (1989).

Group	Configuration	Ocean model	Sea-ice model	Vertical	Orientation	Horiz. grid	Horiz. res.
ACCESS	ACCESS-OM	MOM 4p1	CICE 4	$z^*$ (50)	tripolar	$360 \times 300$	nominal $1^\circ$
AWI		FESOM	FESIM	$z$ (46)	displaced	126000	nominal $1^\circ$
BERGEN	NorESM-O	MICOM	CICE 4	$\sigma_2$ (51+2)	tripolar	$360 \times 384$	nominal $1^\circ$
CERFACS	ORCA1	NEMO 3.2	LIM 2	$z$ (42)	tripolar	$360 \times 290$	nominal $1^\circ$
CMCC	ORCA1	NEMO 3.3	CICE 4	$z$ (46)	tripolar	$360 \times 290$	nominal $1^\circ$
CNRM	ORCA1	NEMO 3.2	Gelato 5	$z$ (42)	tripolar	$360 \times 290$	nominal $1^\circ$
FSU		HYCOM 2.2.21	CSIM 5	hybrid (32)	displaced	$320 \times 384$	nominal $1^\circ$
FSU2		HYCOM 2.2.74	CICE 4	hybrid (32)	tripolar	$500 \times 382$	nominal $0.72^\circ$
GFDL-GOLD	ESM2G-ocean-ice	GOLD	SIS	$\sigma_2$ (59+4)	tripolar	$360 \times 210$	nominal $1^\circ$
GFDL-MOM	ESM2M-ocean-ice	MOM 4p1	SIS	$z^*$ (50)	tripolar	$360 \times 200$	nominal $1^\circ$
GISS		GISS Model E2-R		mass (32)	regular	$288 \times 180$	$1.25^\circ \times 1^\circ$
GISS2		HYCOM 0.9	SIS	hybrid (26)	regular	$360 \times 387$	nominal $1^\circ$
ICTP		MOM 4p1	SIS	$z^*$ (30)	tripolar	$180 \times 96$	nominal $2^\circ$
INMOM		INMOM		sigma (40)	displaced	$360 \times 340$	$1^\circ \times 0.5^\circ$
KIEL	ORCA05	NEMO 3.1.1	LIM 2	$z$ (46)	tripolar	$722 \times 511$	nominal $0.5^\circ$
MIT		MITgcm	H79	$z$ (50)	quadrupolar	$360 \times 292$	nominal $1^\circ$
MRI-A	(data assimilation)	MOVE/MRI.COM 3	MK89; CICE	$z$ (50)	tripolar	$360 \times 364$	$1^\circ \times 0.5^\circ$
MRI-F		MRI.COM 3	MK89; CICE	$z$ (50)	tripolar	$360 \times 364$	$1^\circ \times 0.5^\circ$
NCAR		POP 2	CICE 4	$z$ (60)	displaced	$320 \times 384$	nominal $1^\circ$
NOCS	ORCA1	NEMO 3.4	LIM 2	$z$ (75)	tripolar	$360 \times 290$	nominal $1^\circ$

Table 2: Summary of AMOC maximum transports and linear trends at 26.5°N (columns 3-5) and 45°N (columns 6-8). Models are listed in alphabetical order according to the participating group name (first column). The second column shows whether the AMOC EOF1 maximum occurs in the Northern Hemisphere (N), in the Southern Hemisphere (S), or near the equator (E). The mean transports represent 50-year means for the 1958–2007 period. The linear trends are calculated for the 1978–1998 and 1998–2007 periods for 26.5°N; and for the 1975–1995 and 1995–2007 periods for 45°N based on the annual-mean data. The trends that meet the 95% confidence level based on a two-sided Student’s t-test are shown in bold. The mean transports and trends are in Sv and Sv-decade<sup>-1</sup>, respectively. MMM does not include MRI-A.

Group	N/S/E	26.5°N			45°N		
		Mean	1978-1998 trend	1998-2007 trend	Mean	1975-1995 trend	1995-2007 trend
ACCESS	S	14.3	<b>0.33</b>	-1.08	17.1	0.32	-1.35
AWI	N	12.7	<b>1.52</b>	-3.27	11.7	<b>1.37</b>	-1.68
BERGEN	N	17.0	<b>0.64</b>	-0.34	14.8	<b>1.01</b>	-1.59
CERFACS	N	12.5	<b>0.62</b>	-1.02	12.7	<b>0.91</b>	-1.86
CMCC	S	11.2	<b>0.33</b>	-0.95	11.0	<b>0.48</b>	-1.51
CNRM	N	15.3	<b>1.15</b>	-2.30	15.6	<b>1.75</b>	-2.53
FSU	S	4.9	0.16	-0.02	2.9	<b>0.37</b>	-0.02
FSU2	S	11.6	<b>0.77</b>	-3.15	13.3	<b>0.82</b>	-1.21
GFDL-GOLD	N	13.8	<b>0.62</b>	-3.19	13.2	<b>0.75</b>	-2.67
GFDL-MOM	N	15.8	<b>1.08</b>	-2.85	16.1	<b>0.93</b>	-2.06
GISS	N	16.8	<b>1.62</b>	-8.13	18.1	<b>2.06</b>	-4.81
GISS2	N	17.7	<b>0.88</b>	-2.57	15.2	0.11	-1.25
ICTP	N	11.4	<b>0.66</b>	-2.63	17.9	0.54	-3.52
INMOM	N	16.7	<b>0.82</b>	-1.73	12.8	<b>1.01</b>	-1.52
KIEL	N	14.3	<b>0.85</b>	0.25	15.2	<b>1.50</b>	-1.03
MIT	S	11.0	0.13	0.15	11.2	<b>0.33</b>	0.15
MRI-A	E	16.0	0.09	0.20	20.0	0.32	-1.10
MRI-F	S	11.0	0.28	-0.30	12.7	<b>0.48</b>	0.27
NCAR	N	17.5	<b>0.66</b>	-0.38	20.0	<b>0.88</b>	-1.34
NOCS	S	10.4	0.17	0.72	10.3	0.03	0.29
MMM		13.5	<b>0.70</b>	-1.73	13.8	<b>0.82</b>	-1.54

Table 3: Simultaneous correlation and regression relationships between the AMOC maximum transports and meridional heat transports (MHT) at  $26.5^{\circ}\text{N}$  based on the annual-mean transports for 1958–2007. Models are listed in alphabetical order according to the participating group name (first column). The second column gives the correlation coefficients. The regression coefficients and the intercept values obtained when MHT is regressed onto AMOC are listed in the third and fourth columns, respectively.

Group	Correlation	Regression ( $\text{PW Sv}^{-1}$ )	Intercept (PW)
ACCESS	0.93	0.063	-0.095
AWI	0.98	0.065	0.011
BERGEN	0.94	0.055	0.032
CERFACS	0.95	0.061	0.022
CMCC	0.94	0.067	-0.094
CNRM	0.96	0.059	0.000
FSU	0.96	0.067	0.007
FSU2	0.91	0.058	0.082
GFDL-GOLD	0.96	0.064	-0.099
GFDL-MOM	0.96	0.058	-0.070
GISS	0.96	0.051	0.103
GISS2	0.95	0.047	0.177
ICTP	0.97	0.061	0.047
INMOM	0.86	0.042	-0.008
KIEL	0.97	0.056	0.053
MIT	0.92	0.063	-0.026
MRI-A	0.86	0.066	-0.068
MRI-F	0.94	0.067	-0.117
NCAR	0.95	0.062	-0.072
NOCS	0.93	0.068	-0.070