Decadal variability of oxygen uptake, export, and storage in the Labrador Sea from observations and CMIP6 models

Jannes Koelling [1,*], **Dariia Atamanchuk** [1], **Douglas WR Wallace** [1], **and**

Johannes Karstensen [2]

[1]*Dalhousie University, Department of Oceanography, Halifax, NS, Canada*
[2]*GEOMAR Helmholtz Centre for Ocean Research Kiel, Kiel, Germany*

Correspondence*:
Jannes Koelling
j.koelling@dal.ca

2 ABSTRACT

3   The uptake of dissolved oxygen from the atmosphere via air-sea gas exchange and its physical
4 transport away from the region of uptake are crucial for supplying oxygen to the deep ocean.
5 This process takes place in a few key regions that feature intense oxygen uptake, deep water
6 formation, and physical oxygen export. In this study we analyze one such region, the Labrador
7 Sea, utilizing the World Ocean Database (WOD) to construct a 65–year oxygen content time
8 series in the Labrador Sea Water (LSW) layer (0–2200 m). The data reveal decadal variability
9 associated with the strength of deep convection, with a maximum anomaly of 27 $mol\ m^{-2}$ in
10 1992. There is no long-term trend in the time series, suggesting that the mean oxygen uptake is
11 balanced by oxygen export out of the region.

12   We compared the time series with output from nine models of the Ocean Model Intercomparison
13 Project phase 1 in the Climate Model Intercomparison Project phase 6, (CMIP6-OMIP1), and
14 constructed a "model score" to evaluate how well they match oxygen observations. Most CMIP6-
15 OMIP1 models score around 50/100 points and the highest score is 57/100 for the ensemble
16 mean, suggesting that improvements are needed. All of the models underestimate the maximum
17 oxygen content anomaly in the 1990s. One possible cause for this is the representation of air-sea
18 gas exchange for oxygen, with all models underestimating the mean uptake by a factor of two or
19 more. Unrealistically deep convection and biased mean oxygen profiles may also contribute to
20 the mismatch. Refining the representation of these processes in climate models could be vital for
21 enhanced predictions of deoxygenation.

22   In the CMIP6-OMIP1 multi-model mean, oxygen uptake has its maximum in 1980–1992, followed
23 by a decrease in 1994–2006. There is a concurrent decrease in export, but oxygen storage also
24 changes between the two periods, with oxygen accumulated in the first period and drained out in
25 the second. Consequently, the change in oxygen export (5%) is much less than that in uptake
26 (28%), suggesting that newly ventilated LSW which remains in the formation region acts to buffer
27 the linkage between air-sea gas exchange and oxygen export.

1 INTRODUCTION

30 The Labrador Sea is one of the few regions in the world where oxygen can be transported directly into
31 the deep ocean, which acts as a supply for marine organisms throughout the global oceans. This supply of
32 dissolved oxygen to the deep is the result of oxygen uptake through air-sea gas exchange (Atamanchuk
33 et al., 2020), and its subsequent export out of the region (Koelling et al., 2022). Oxygen uptake occurs
34 primarily during wintertime, when the process of deep convection continually mixes the water column to
35 depths of more than 1000 m, bringing up low-oxygen water from below and maintaining an undersaturation
36 of surface waters relative to the atmosphere (Wolf et al., 2018). The resulting air-sea gradient in oxygen,
37 along with bubble injection driven by strong winds and wave breaking, leads to intense air-sea exchange of
38 oxygen during these winter months (Koelling et al., 2017). As a result the Labrador Sea, and the subpolar
39 North Atlantic as a whole, feature some of the strongest oxygen uptake rates anywhere in the world ocean
40 (Huang et al., 2018), with annual values in the Labrador Sea being on the order of 20 $mol\ m^{-2}\ y^{-1}$
41 (Atamanchuk et al., 2020). Thanks to this strong uptake and its link to deep water formation, the subpolar
42 North Atlantic, along with the Southern Ocean, is the main region where oxygen can enter the deep ocean,
43 making it one of the "lungs of the ocean" (Körtzinger et al., 2004; Portela et al., 2020). North Atlantic
44 Deep Water (NADW), a water mass formed in part as a result of deep convection in the Labrador Sea,
45 can be traced by its high oxygen concentrations in the formation region which decrease along southward
46 spreading pathways (Fig. 1). The formation of NADW in the subpolar North Atlantic also leads to a stark
47 difference in mean oxygen concentration between the Atlantic Ocean and the Pacific and Indian Oceans.

48 The Labrador Sea has been sampled more consistently than most open ocean regions over many decades,
49 with regular measurements dating back to the late 1940s. Furthermore, the main convection region in the
50 interior Labrador Sea is relatively small and homogeneous, such that individual measurements at only a few
51 stations can give useful information about variability in the basin (van Aken et al., 2011; Yashayaev and
52 Loder, 2016). Ocean Weather Ship Bravo was stationed in the center of the basin from the 1940s to 1970s,
53 and following occasional cruises in the 1950s to 1980s, the AR7W line across the basin from Labrador to
54 Greenland has been occupied almost every year since 1987 as part of the WOCE, CLIVAR and AZOMP
55 initiatives (Kieke and Yashayaev, 2015). The comprehensive record of oceanographic variables collected
56 by these projects has been used primarily in the context of the variability in physical properties, and their
57 relation to deep convection and the Atlantic Meridional Overturning Circulation (AMOC) (Yashayaev,
58 2007; Yashayaev and Loder, 2016). Fewer studies have focused on the variability in oxygen concentrations,
59 with notable exceptions including studies by Stendardo and Gruber (2012) and van Aken et al. (2011) that
60 reported long-term oxygen trends and variability in the subpolar North Atlantic until 2010 from the historic
61 shipboard data, and Rhein et al. (2017) who analyzed variability in oxygen concentrations since the early
62 1990s. Here, we utilize the data to construct a time series of oxygen content anomalies for the Labrador
63 Sea Water (LSW) layer (0–2200 m) in the central Labrador Sea from 1950 to 2015. We use this time series
64 to discuss how the decadal-scale variability of oxygen concentrations reported in previous studies relates
65 to variability in the uptake of oxygen during deep convection and its subsequent export along NADW
66 spreading pathways.

67 Changes in the strength of ocean ventilation have been proposed as one of the reasons for widespread
68 global ocean deoxygenation observed in recent decades (Ito et al., 2017; Schmidtko et al., 2017). However,
69 this is not reproduced accurately in climate models, which feature too little deoxygenation particularly in
70 the deep ocean, which may be related to difficulties with representing the processes underlying ventilation
71 (Oschlies et al., 2017, 2018; Buchanan and Tagliabue, 2021). To assess the ability of climate models
72 to reproduce observed variability in oxygen ventilation in the Labrador Sea, we use the time series

73 constructed from observational data to evaluate nine different models that participated in the Ocean
74 Model Intercomparison Project phase 1 component of the Climate Model Intercomparison Project phase
75 6 (hereinafter CMIP6-OMIP1). For this purpose, we develop a "model score" based on the ability of
76 CMIP6-OMIP1 models to reproduce the observed variability in the oxygen content time series and mean
77 profile from the WOD data, as well as gas exchange estimates taken from other studies. This score serves as
78 a simplified metric of the ability of climate models to represent the complex processes that underlie ocean
79 ventilation. Our assessment highlights processes which need to be better resolved in climate models in order
80 to more accurately portray oxygen cycling in the Labrador Sea. This analysis could inform improvements
81 for future model generations, and may be applicable to other regions with comparable dynamics.

82 Finally, we combine our analysis of the long-term variability in oxygen content with insights from
83 previous observational studies of air-sea gas exchange and oxygen export, and with the CMIP6-OMIP1
84 model data, in order to understand how air-sea gas exchange, oxygen storage, and oxygen export are related
85 on interannual and decadal time scales.

2 MATERIALS AND METHODS

2.1 Observational time series
86

87 To construct the oxygen inventory time series, we use quality controlled data from the World Ocean
88 Database (WOD; Boyer et al. (2013)) in the central Labrador Sea, which includes oxygen measurements
89 from as early as the 1950s. Some data from individual cruises are available as far back as 1927, but given
90 the large temporal gaps these are not used in our study. We use all measurements within a 200 km radius
91 of 56°49.3'N, 52°13.2'W from fixed platforms and shipboard hydrography, which are flagged as "good
92 data" in the WOD (see Fig. 1 for location). The data were averaged into annual values with 50 m vertical
93 bins, and cover the period from 1950 to 2015. We chose not to include more recent data from profiling
94 floats such as BGC-Argo in order to avoid biases resulting from the shift of predominantly ship-based data
95 typically collected during spring or summer to predominantly autonomous data collected year-round. Even
96 with the shipboard data, there may be some biases resulting from changing sampling schemes and times
97 between years, particularly in the early parts of the record, as well as year-to-year variability associated
98 with instrumental errors during individual cruises (Yashayaev et al., 2020). However, the focus here is
99 on the longer-term decadal signals rather than variability from one year to the next, and we assume these
100 effects are small compared to the decadal variability in properties, consistent with previous studies (van
101 Aken et al., 2011). To further test the representativeness of a single point and data collected once a year for
102 measuring interannual variability in the whole region, we used CMIP6-OMIP1 model output to construct a
103 synthetic time series derived from one grid point in the center of the convection region once a year and
104 compared it to the full time series using all data within a 200 km radius (Fig. S1). The comparison suggests
105 that taking measurements only at one location during summer each year does not introduce any significant
106 bias.

107 To construct the time series of oxygen content anomalies, the time-mean oxygen profile was first
108 subtracted from the annual WOD profiles, and the resulting anomalies were interpolated vertically between
109 measurements, and extrapolated to the surface and seafloor, ommitting years where no measurements were
110 available below 2000 m. The anomalies were then integrated over the 0–2200 m depth range to obtain a
111 time series of annual oxygen content anomalies in this depth range which we assume to be the maximum
112 depth typically occupied by Labrador Sea Water (Yashayaev and Loder, 2016). The time series used for
113 further analysis and comparison with the CMIP6-OMIP1 data was constructed by smoothing the annual

114 time series with a 3-year running mean filter to reduce some of the noise associated with the sampling
115 issues discussed above, and then filling any remaining gaps in the data with linear interpolation.

116 To estimate the uncertainty of the time series, we first calculated the standard deviation σ in each one-year,
117 50 m depth bin. These annual values were then averaged to obtain a mean standard deviation profile for
118 the entire record as an approximation of the true statistical standard deviation. The standard error was
119 calculated from this mean profile and the number of observations in each bin as $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{N}}$. Bins with no
120 measurements were treated as having N = 1, i.e. the standard error is equal to the standard deviation. This
121 method was chosen to avoid underestimating the error in years with fewer observations, where the standard
122 deviation of the samples will be lower than the true standard deviation of the statistical population. The
123 resulting error estimate for the oxygen inventory ranges from 1.59 $mol\ m^{-2}$ for recent years with more
124 extensive coverage to 9.65 $mol\ m^{-2}$ for years with limited or no measurements.

2.2 CMIP6-OMIP1 models

126 To evaluate the processes governing the oxygen budget in the Labrador Sea, we also used data from nine
127 ocean models that were part of the Coupled Model Intercomparison Project phase 6 (CMIP6), accessed
128 from `https://esgf-node.llnl.gov/search/cmip6/`. The models selected were those that
129 participated in phase 1 of the ocean model intercomparison project (OMIP1) and included a biogeochemical
130 component (Orr et al., 2017) with monthly output of the variables of interest. The resulting selection of
131 models was CCC CanESM5, CMCC ESM2, CNRM ESM2-1, IPSL CM6A-LR, MIROC E2SL, MRI
132 ESM2.0, NCAR CESM2, NCC NorESM-LM, and NOAA GFDL-CM4, and we used the $r1i1p1f1$ variant
133 where available, or the lowest number variant otherwise, as listed in table 1. The model outputs used
134 for the calculations are oxygen concentration (variable name $o2$), and surface air-sea gas exchange of
135 oxygen (variable name $fgo2$). The time period considered in this study is 1950–2009, which is the window
136 overlapping with the period of consistent observational data (1950–2015), and also coincides with almost
137 a full cycle of the 1948–2009 CORE-II atmospheric forcing (Large and Yeager, 2009) that is used in all
138 OMIP1 runs (Orr et al., 2017). Based on the comparison in Fig. S1 we use data from a single grid point in
139 the center of the convection region to be representative of the convective patch.

2.3 Oxygen budget calculation

141 As a biologically active element, the change in the inventory of oxygen for a given volume of the open
142 ocean is controlled by both physical processes, such as air-sea gas exchange, lateral advection, and vertical
143 export, and net community metabolism, which represents the net effect of biological production minus
144 autotrophic and heterotrophic respiration (Martz et al., 2009; Koelling et al., 2017).

$$\frac{\partial \int_0^z O_2}{\partial t} = F_{gasex} - F_{lat} - F_{vert} + F_{bio} \qquad (1)$$

145 where $\Delta \int_0^z O_2$ is the change in oxygen content between the surface and depth z in a given time period,
146 and the terms on the right hand side are the changes due to gas exchange, lateral exchanges, vertical
147 transport, and net community metabolism, respectively. In the above definition the vertical and lateral
148 terms, F_{vert} and F_{lat}, are positive for a flow of oxygen out of the volume considered.

149 One mechanism for vertical oxygen transport is the deepening of the mixed layer, which can have a
150 significant effect in the upper ocean as water with different properties is entrained into the mixed layer
151 (Martz et al., 2009). Vertical exchanges also occur as a result of deep wintertime convection, which can

152 extend the mixed layer to depths exceeding 2000 m in the Labrador Sea, and through diffusive fluxes at the
153 base of the integration depth. In this study, we integrate the oxygen content to the depth reached by the
154 maximum historical convection, 2200 m (Yashayaev and Loder, 2016), so the effect of deep convection, as
155 well as shallow summertime entrainment, is eliminated. Thus, the F_{vert} term in our calculation represents
156 only the exchange of oxygen that may occur with the underlying water masses by turbulent mixing, which
157 should be small in the deep ocean. The F_{lat} term represents lateral exchanges between the interior Labrador
158 Sea and the surrounding boundary current (Wolf et al., 2018; Koelling et al., 2022), as well as its export
159 along interior pathways (Talley and McCartney, 1982; Yashayaev and Loder, 2016). We refer to the sum of
160 the F_{vert} and F_{lat} terms as *oxygen export*, as they are equivalent to the amount of oxygen that is exported
161 by physical processes, leaving the volume of the Labrador Sea that can be affected by convection again
162 in the following winters. However, it is more accurately the net export, as it includes both the removal of
163 relatively high oxygen Labrador Sea Water (Koelling et al., 2022) and its replacement with lower oxygen
164 water either from exchange with the boundary current or vertical mixing (Wolf et al., 2018).

165   Previous results studying parts of the oxygen budget in the Labrador Sea suggest that the biological
166 component is relatively small compared to gas exchange and export in the mean balance (Koelling et al.,
167 2017; Wolf et al., 2018). To estimate its variability on interannual to decadal timescales, we use data from
168 the Carbon-based Productivity Model (Westberry et al., 2008), a satellite-based net primary production
169 product, along with a global estimate of oxygen utilization rates in the deep ocean (Karstensen et al., 2008),
170 to calculate annual Net Community Production (NCP, Fig. S2). Although this product only covers the
171 period since 2002, and therefore does not resolve the historical period analyzed here, interannual variability
172 is small (range: 0 mol m^{-2} to 1.16 mol m^{-2} of O_2), suggesting that the contribution of the biological
173 component and its variability to decadal changes are negligible for the 0-2200 m integrated budget in this
174 region of particularly strong physical forcing.

175   In our investigation of the oxygen content changes in the following sections, we therefore assume the
176 year-to-year change in oxygen content, also referred to as oxygen storage throughout this study, to be driven
177 only by air-sea gas exchange and export. In section 4, using CMIP6-OMIP1 model data, we rearrange eq. 1
178 to infer the oxygen export from model output of oxygen storage and gas exchange.

179 **2.4 Model score**

180   In order to evaluate the ability of different models to reproduce the observed variability in the Labrador
181 Sea, we construct a model score designed to represent the processes that are relevant for deep ocean
182 ventilation. The source code for calculating the score is provided at https://github.com/jkoell/labrador-sea-
183 oxygen
184 Similar analyses have been conducted in other studies for climate variables on global scales (Gleckler et al.,
185 2008; Anav et al., 2013), as well as for more regional assessments of ocean biogeochemistry, similar to
186 the present study (Rickard et al., 2016; Laurent et al., 2021). Unlike most of these previous works, our
187 analysis focuses mostly on interannual changes instead of seasonal cycles. We also choose a score that
188 is based solely on the comparison with observational data, rather than a relative model score or ranking
189 as used in other studies (Gleckler et al., 2008; Anav et al., 2013). This method is chosen to make the
190 score independent of the number or types of model used, making it easier to compare future models to
191 the results from this study. For example, the score calculation could be applied to high-resolution regional
192 models of the subpolar North Atlantic, or future generations of CMIP (e.g. Koenigk et al. (2021)), and a
193 higher score than the models analyzed here would imply better performance. However, we note that a score

194 calculated relative to the worst-performing model for each category, following Anav et al. (2013), would
195 yield qualitatively similar results.

196   There are five factors considered, with scores for each ranging from 0 to 20 points, for a maximum total
197 of 100 points which would indicate a perfect match with observations. These are summarized below, with
198 details on score calculation and reasoning for each:

199 • **The Pearson correlation coefficient (R) of 0-2200 m inventory anomaly with observational time**
200   **series**: This metric measures the extent to which the model time series reproduces the phase of the
201   decadal variability evident in the observational data. Scoring is calculated as $S_{corr} = \frac{R^2}{0.81} \times 20$ for
202     $R > 0$, and $S_{corr} = 0$ for $R \leq 0$

203 • **Maximum 0–2200 m inventory anomaly**: Another measure of how well the observational time series
204   is reproduced, focused on the amplitude of the 1990s LSW layer oxygen content anomaly. To allow for
205   slight variations in timing, the amplitude is calculated as the maximum between 1990 and 1995, minus
206   the minimum from 2004 to 2009. Scoring is based on the ratio of the amplitudes from observations
207   and models, with 20 points for a ratio of 1. The observational maximum used for the calculation is
208     $Max_{obs} = 38.51 \; mol \; m^{-2}$.

209 • **Standard deviation for the 2200m-bottom inventory anomaly**: Used to evaluate if there is excess
210   variability below the LSW layer. In observations, oxygen variability in this depth range is lower
211   compared to changes above. However, some models have unrealistically large anomalies near the
212   bottom, likely associated with excessively deep-reaching convection (see section 4.1). Scoring is based
213   on the ratio of the standard deviations from observations and models, with 20 points for a ratio of 1. The
214   standard deviation of the observational time series used for the calculation is $\sigma_{obs} = 4.23 \; mol \; m^{-2}$.

215 • **Average absolute difference of the time-mean profile with observations**: Designed to measure the
216   bias of the mean oxygen distributions compared to observations. This is split up into two depth ranges,
217   0–2200 m and 2200 m–bottom, with 0 to 10 points assigned for each. 10 points: $\Delta = 0 \; \mu mol \; L^{-1}$; 0
218   points: $\Delta \geq 20 \; \mu mol \; L^{-1}$

219 • **Mean annual gas exchange**: In the central Labrador Sea, annual net oxygen uptake is an important
220   driver for variability in the oxygen inventory. We use observational estimates from the central Labrador
221   Sea from two previous studies, for a total of five years: $22.1 \; mol \; m^{-2}$ from Atamanchuk et al.
222   (2020), and $21.36 \; mol \; m^{-2}$, $18.82 \; mol \; m^{-2}$, $22.91 \; mol \; m^{-2}$, $19.97 \; mol \; m^{-2}$, $33.87 \; mol \; m^{-2}$, and
223     $19.59 \; mol \; m^{-2}$ (Wolf et al. (2018), supplementary table S1, Liang flux with CCMP winds). The mean
224   of these five estimates is $22.7 \; mol \; m^{-2}$, with a standard deviation of $5.2 \; mol \; m^{-2}$. 20 points for a
225   difference between $\pm \frac{\sigma_{obs}}{2}$, 0 points for a difference $\geq \sigma_{obs} \times 4$.

Mathematically, the scores as described above are given by

$$S_{corr} = \begin{cases} 20 \times \frac{R^2}{0.81} & \text{if } R > 0 \\ 0 & \text{if } R \leq 0 \end{cases}$$

$$S_{LSWmax} = \begin{cases} 20 \times \frac{Max_{model}}{Max_{obs}} & \text{if } Max_{model} \leq Max_{obs} \\ 20 \times \frac{Max_{obs}}{Max_{model}} & \text{if } Max_{model} > Max_{obs} \end{cases}$$

$$S_{\sigma} = \begin{cases} 20 \times \frac{\sigma_{model}}{\sigma_{obs}} & \text{if } \sigma_{model} \leq \sigma_{obs} \\ 20 \times \frac{\sigma_{obs}}{\sigma_{model}} & \text{if } \sigma_{model} > \sigma_{obs} \end{cases}$$

$$S_{bias} = (10 - |\Delta(O_2(obs - mod, 0 - 2200m)|/2) + (10 - |\Delta(O_2(obs - mod, 2200m - bottom)|/2)$$

$$S_{GE} = 20 - (||GE_{obs}| - |GE_{mod}|| - \frac{\sigma(GE_{obs})}{2}) \times \frac{20}{4\sigma(GE_{obs})}$$

$$S_{total} = S_{corr} + S_{LSWmax} + S_{\sigma} + S_{bias} + S_{GE} \qquad (2)$$

Note that before the calculation of the final score S_{total}, individual score components below 0 are set to 0, and scores above 20 to 20.

As noted in previous studies, the selection of metrics by which to score or rank models will always be somewhat arbitrary (Anav et al., 2013). We choose an equal weighting of all five metrics to avoid subjective biases in how important each one is, although there will still be some bias in our choice of which metrics to include at all. Gleckler et al. (2008) suggested that an overall performance metric for climate models is not very informative, because there are many offsetting errors in different components and different regions, and the performance of a given model will depend on the specific application. The approach we took here is to calculate a score that is specific to one mechanism - deep ocean ventilation with oxygen - in one key region - the Labrador Sea - rather than a more comprehensive global ranking. This should ensure that the score calculated here more meaningfully represents the fidelity of different models in simulating the overall effect of this particular process.

3 OBSERVED OXYGEN CONTENT IN THE LABRADOR SEA

3.1 Variability of the oxygen inventory

The time evolution of the oxygen inventory anomaly in the Labrador Sea derived from all available historical shipboard data is shown in Fig. 2. The time series exhibits decadal variability, with maxima in the 1960s, 1970s, and 1990s, and minima in 1970, 1980, and 2010. These maxima have been reported in previous studies on Labrador Sea oxygen concentrations, and are closely linked to extended periods of particularly strong convection events (van Aken et al., 2011). A period of consistently increasing oxygen inventory anomalies occurred during the record deep convection of the early 1990s (Yashayaev, 2007), followed by a steady decrease in the late 1990s, and a minimum in 2010. There is an increase in oxygen content at the end of the record, which is associated with the re-emergence of deep convection starting in winter 2014/2015 (Yashayaev and Loder, 2016; Rhein et al., 2017) that led to enhanced oxygen uptake (Koelling et al., 2017; Atamanchuk et al., 2020). During the buildup to the maximum oxygen content anomaly from 1980 to 1992, there was an average increase, or oxygen storage, of 2.82 $mol\ m^{-2}$ per year, and the subsequent decrease occurred at a rate of 2.04 $mol\ m^{-2}$ per year from 1994–2006.

251   Although the inventory change could intuitively be assumed to be linked to gas exchange, the relationship
252   between the two is not necessarily strictly proportional. An analysis of oxygen inventories from Argo
253   floats showed that even during times of strong convection and larger than usual gas exchange, there is
254   little year-to-year change in the oxygen inventory (Wolf et al., 2018). Indeed, the total increase in oxygen
255   content over a 12-year period discussed above is only 1.5 times the uptake from the atmosphere for one
256   single year of 22.1 $mol\ m^{-2}$ observed by Atamanchuk et al. (2020). Estimating the effect of eddies such
257   as Irminger Rings advecting lower oxygen water into the central Labrador Sea, Wolf et al. (2018) found
258   that they could account for oxygen losses between 3.98 $mol\ m^{-2}$ and 11.34 $mol\ m^{-2}$ per year . Koelling
259   et al. (2022) showed that about 50% of the oxygen taken up in 2017 is rapidly exported southward into
260   the Deep Western Boundary Current, and hypothesized that additional oxygen export is associated with
261   different pathways of LSW spreading, such as an interior flow towards the Irminger Sea that has been
262   observed in transient tracer studies (Sy et al., 1997) and hydrographic data (Talley and McCartney, 1982;
263   Yashayaev, 2007). Together with our time series, these results suggest that to first order, the annual net
264   oxygen uptake from the atmosphere is largely balanced by the export of newly ventilated water out of
265   the formation region, and its replacement with lower oxygen water from the surrounding regions, and the
266   variability evident in Fig. 2 represents the cumulative effect of the small annual residual between these two
267   processes. The continuous loss of oxygen through outflowing LSW and its replacement with lower-oxygen
268   water may in fact be crucial for maintaining the Labrador Sea as a net oxygen sink year after year, as it
269   ensures that the wintertime mixed layer is far from equilibrium with the atmosphere even during years of
270   weaker convection (Wolf et al., 2018).

3.2   The Labrador Sea's oxygen reservoir

272   Given the dominant balance between gas exchange and export in the annual oxygen budget, we conclude
273   that the oxygen content increase from 1980–1992, and decrease from 1994–2006, is a result of uptake from
274   the atmosphere exceeding the losses from oxygen export during the former period, and export exceeding
275   uptake during the latter. The first of these periods, referred to in the following as Period 1, encompasses the
276   deepest reaching convection on record, which resulted in the coldest, freshest, and most oxygenated class
277   of LSW ever observed (Yashayaev, 2007; van Aken et al., 2011). Following the peak of this convective
278   activity in the early 1990s, the Labrador Sea experienced a period of weaker than usual convection during
279   1994–2006 coinciding with the decrease in oxygen content, which we will refer to as Period 2.

280   This time scale of the buildup and subsequent drainage of the oxygen reservoir may be related to the
281   ocean's "memory" in the region, which is a feature that has been observed in the Labrador Sea Water's
282   physical properties (Straneo et al., 2003; Yashayaev, 2007). Although there is a significant export of
283   LSW every year (Yashayaev and Loder, 2016), some of the newly formed deep water remains in the
284   formation region, so that the changes set by air-sea forcing during one convective season have an effect
285   on the initial conditions during the following winter. The result is a preconditioning of the water column,
286   making it easier for convection to penetrate deeper in a year following deep-reaching convection, and more
287   difficult following shallower convection (Yashayaev and Loder, 2017). This process contributes to the LSW
288   reservoir in the interior of the basin becoming progressively colder, fresher, and more oxygenated during
289   prolonged periods of increasingly deep reaching convection (van Aken et al., 2011). At the same time,
290   there is a continuous exchange between the interior convection region and the warmer, saltier boundary
291   current, as well as exchanges in the inner subpolar gyre (Straneo, 2006; Yashayaev, 2007). This results
292   in a positive advective heat flux (Yashayaev and Loder, 2009), which acts to increase stratification in the
293   central Labrador Sea and erode the pool of homogenized deep water (Tagklis et al., 2020). This exchange,
294   much of which happens along density surfaces, is also the main mechanism by which LSW from the

295 interior is added to the outflowing boundary current (Georgiou et al., 2020; Zou et al., 2020). Interannual
296 variability in the lateral exchange between the boundary current and interior Labrador Sea is related to the
297 boundary-interior gradient of layer thickness, and the resulting advective heat gain is mostly controlled by
298 the temperature of LSW (Straneo, 2006).

299   This variability in boundary-interior exchange could also affect the oxygen budget, as heat content and
300 oxygen inventory between 0–2200 m in the central Labrador Sea are highly correlated (R = -0.83; not
301 shown). This exchange may explain the leveling off and reversal of the inventory changes evident in Fig. 2:
302 If the exchange of oxygen, like heat, is proportional to its reservoir in the interior, a continuous increase in
303 oxygen content would lead to an increase in the oxygen lost through exchanges with the boundary current.
304 Eventually the export of oxygen would exceed gas exchange, with the uptake from the atmosphere no
305 longer sufficient to replenish the oxygen that is exported out of the region, resulting in a decrease in oxygen
306 content. This could occur either due to the continued intensification of the lateral oxygen exchange, or a
307 decrease in oxygen uptake from the atmosphere. Since some LSW can stay in the formation region for
308 several years (Straneo et al., 2003), the elevated oxygen export would persist for some time even after
309 gas exchange weakened. The buildup of oxygen during Period 1, and following decrease during Period 2,
310 would be consistent with this mechanism. During Period 1, several consecutive years of deep-reaching
311 convection and intense oxygen uptake could explain the continued increase in oxygen content, which in
312 turn would drive an increased boundary-interior oxygen exchange. In Period 2, the exchange would remain
313 elevated even once the oxygen uptake weakened following peak convection, thus slowly draining the excess
314 oxygen accumulated during Period 1.

315   However, from the oxygen content time series alone it is not possible to know for certain whether the
316 mechanism proposed above is at play in the Labrador Sea. The variability could also be explained by
317 changes in gas exchange with constant export, changes in export with constant gas exchange, or some
318 other variability in the gas exchange-export balance that does not follow the pattern outlined above. In
319 order to fully disentangle the relationship between decadal variability in oxygen uptake, gas exchange,
320 and inventory change, sustained long-term observations of all three would be needed. While this may
321 be possible in the future using the much broader coverage of *in situ* measurements in recent years (von
322 Oppeln-Bronikowski et al., 2021; Atamanchuk et al., 2022; Sharp et al., 2022), at present a multi-decade
323 observational record only exists for the oxygen inventory. In the meantime, data from ocean climate models
324 may be a useful tool to better understand these processes, which is the focus of the following section.

4 OXYGEN BUDGET IN CMIP6-OMIP1

325 In this section, we use data from nine different CMIP6 models in order to analyze the balance between the
326 different terms in the oxygen budget. We use output from all models that participated in the OMIP1 phase
327 and provided fields of oxygen and air-sea flux of oxygen, as described in section 2. Section 4.1 focuses on
328 evaluating how well the models reproduce the observational record, and uses the score defined in Section
329 2.4 to assess model performance. In section 4.2, we use the ensemble mean in order to diagnose differences
330 in gas exchange, oxygen export, and oxygen storage between the two periods defined in section 3.2.

4.1 Model Evaluation

332   Figure 3A shows the central Labrador Sea's oxygen inventory anomaly between 0-2200 m taken from
333 nine different CMIP6-OMIP1 models that include ocean biogeochemistry, as well as the mean of all models
334 (see Section 2 for details). While there is some spread between the modeled oxygen inventory in any given
335 year, they all exhibit variability similar to observations on decadal time scales. In particular, a maximum in

336 oxygen content in the mid-1990s is evident in all models, suggesting that the impact of the record-deep
337 convection during this time (Yashayaev, 2007) is reproduced to some extent in CMIP6-OMIP1. However,
338 the magnitude of the peak oxygen content anomaly is reduced by about half compared to the observational
339 record.

340   A possible explanation for this discrepancy could be the representation of gas exchange in climate models.
341 The biogeochemical protocols for participants of OMIP (Orr et al., 2017) state that modelers should use
342 the gas exchange parameterization of Wanninkhof (1992) with the updated coefficients of (Wanninkhof,
343 2014). This parameterization is constrained to reproduce global air-sea fluxes of carbon, but has been
344 shown to substantially underestimate oxygen uptake in the Labrador Sea because it does not include a
345 bubble-mediated term, which makes up a significant fraction of the total gas exchange for oxygen in this
346 region (Atamanchuk et al., 2020; Koelling et al., 2017). Figure 3B shows the time series of the annual
347 gas exchange from the CMIP6-OMIP1 models, with time-mean values provided in Table S1. Indeed, the
348 models seem to significantly underestimate the gas exchange observed in the real ocean: The highest mean
349 annual gas exchange from any of the models is 10.91 $mol\ m^{-2}$ (NOAA GFDL), which is less than half
350 of the uptake inferred by Atamanchuk et al. (2020) and Wolf et al. (2018), but comparable to an early
351 estimate by Körtzinger et al. (2008) that used the Wanninkhof (1992) parameterization. The mean annual
352 gas exchange for the ensemble mean is 7.22 $mol\ m^{-2}$, about a third of observation-based estimates. This
353 strong discrepancy with the observed data highlights again the need for critical evaluation of gas exchange
354 parameterizations for oxygen in climate models as called for in previous studies (Atamanchuk et al., 2020;
355 Seltzer et al., 2023).

356   Another issue common in CMIP is that modelled convection depths in the Labrador Sea are often
357 much deeper than in observations (Heuzé, 2017, 2021), which may be related to their inability to resolve
358 submesoscale processes (Tagklis et al., 2020). As a result, Labrador Sea Water in these models penetrates
359 deeper than observed in the real ocean, and the deep overflow water masses below LSW are much less
360 pronounced than in observations. This is evident in Fig. 3c, as neither the low-oxygen Northeast Atlantic
361 Deep Water (NEADW) layer from about 1500–3000 m depth, nor the near-bottom oxygen maximum
362 associated with Denmark Strait Overflow Water (Yashayaev and Loder, 2017), are discernible in the model
363 mean oxygen profile. The effect of unrealistically deep-reaching convection can also be seen in oxygen
364 concentration anomalies (Fig. 4). In particular, the IPSL CM6a, MRI ESM2, and NOAA GFDL models
365 have the maximum anomalies in oxygen concentrations near the bottom, rather than in the 1000–2000 m
366 depth range typically occupied by LSW in the real ocean, and all three of these models have been shown
367 to have much deeper convection compared to observations (Adcroft et al., 2019; Boucher et al., 2020;
368 Jackson and Petit, 2022). This is another factor that can contribute to the underestimation of oxygen content
369 anomalies in the LSW layer (Fig. 3a), as the variability in LSW oxygen in these models is spread over
370 the whole water column, rather than just 0–2200 m. The lack of realistic overflow water masses could
371 also be contributing to gas exchange being too weak. During some years, convection can penetrate into
372 the salty, low-oxygen NEADW layer underlying LSW, incorporating it into the newly formed water mass
373 (Yashayaev, 2007). This entrainment of old waters with lower oxygen is an important driver for maintaining
374 undersaturation of surface waters throughout the convective season (Wolf et al., 2018). Thus, the lack of a
375 clear oxygen minimum below LSW (Fig. 3c) could result in surface oxygen concentrations being too close
376 to saturation, reducing the magnitude of the calculated oxygen uptake.

377   Overall, while there generally is multidecadal variability in oxygen content at similar time scales to
378 observations, many CMIP6-OMIP1 models seem to struggle with representing the more detailed aspects
379 of the Labrador Sea's oxygen dynamics. This includes weaker amplitude of the multidecadal signal (Fig.

380 3a), reduced mean oxygen uptake (Fig. 3b), biases in the mean oxygen profile (Fig. 3c), and excessively
381 deep convection (Fig. 4). In order to assess each model's performance in these areas, we calculated a score
382 based on five factors, as described in section 2.4, and the scores for each metric as well as the total score
383 are given in Table 2.

384   The majority of the models score close to 50 out of 100 points, and none higher than 53/100. Based on
385 the different score components (Table 2) the exact nature of the mismatch with observations is different
386 for each model. For example, MRI ESM2 scores the highest on the correlation with the LSW time series,
387 but also has a low amplitude for the LSW maximum, too much variability below the LSW layer, and the
388 weakest mean gas exchange. Conversely, the NCAR CESM2 model does well for the amplitude of the
389 maximum and the mean gas exchange, but the correlation with the observational time series is weaker.
390 Using the ensemble mean of all models results in a better score than any individual model, consistent
391 with the expectation that the mean of an ensemble of models generally is closer to the "true" signal than
392 individual members (Collins et al., 2014). However, even the ensemble mean only achieves a score of 57,
393 with shortcomings evident particularly in the amplitude of the oxygen inventory maximum and the mean
394 gas exchange.

395   While the exact scores are somewhat arbitrary, these results highlight the need for improvements in
396 modelling the oxygen cycle in this key region for deep ocean ventilation. Given that models are our
397 only tool for projections of future ocean deoxygenation, we echo Oschlies et al. (2018)'s sentiment that
398 discrepancies between models and observations need to be resolved in order to understand how oxygen
399 levels may change in the future. We hope that the model score created here can contribute to that cause,
400 providing an easy and consistent way to evaluate how key metrics related to deep ocean ventilation in the
401 Labrador Sea are represented in ocean models.

402 4.2   Balance between oxygen storage, gas exchange, and export

403   Despite the CMIP6-OMIP1 model's shortcomings discussed above, they can be useful for testing the
404 links between gas exchange, oxygen content change, and export in the Labrador Sea. In particular, the
405 increase in the oxygen inventory associated with the 1987-1994 class of LSW in Period 1, and subsequent
406 decrease in Period 2, is present in all models (Fig. 3a), albeit with reduced magnitude. Below, we assess
407 how the different terms of the oxygen budget in the Labrador Sea vary during this record deep convection
408 period in the CMIP6-OMIP1 models.

409   To test how the multidecadal oxygen content variability is controlled by gas exchange and oxygen
410 export, we calculate annual oxygen storage rates and gas exchange using the ensemble mean of the nine
411 CMIP6-OMIP1 models, and infer the export as the residual by rearranging equation 1. Fig. 5A shows the
412 time series of the annual mean gas exchange and oxygen storage from the model outputs. Oxygen storage is
413 the year-to-year change in the oxygen inventory, i.e. the first derivative of the time series shown in Fig. 3A.
414 The long-term mean is $7.22 \ mol \ m^{-2} \ y^{-1}$ for the gas exchange and $0 \ mol \ m^{-2} \ y^{-1}$ for the storage term.
415 Thus, the mean of the export, calculated as the residual, is also $7.22 \ mol \ m^{-2} \ y^{-1}$: On long time scales,
416 the net annual oxygen uptake is balanced by a net export of oxygen. Interannual variability in gas exchange
417 is positively correlated with variability in storage (R = 0.66 for the multi-model mean, Fig. 5B) as well as
418 export (R = 0.64, Fig. 5C), while there is relatively little correlation between export and storage (R = -0.11,
419 not shown). This suggests that both storage and export are driven mostly by variability in oxygen uptake,
420 with both the local oxygen inventory and the amount of oxygen that is exported increasing during years with
421 more intense air-sea gas exchange, and decreasing when less oxygen is taken up. There is also a relationship
422 between the oxygen content anomaly (Fig. 3A) and oxygen export, with a correlation of R = 0.73. This

423 would be consistent with a lateral exchange of oxygen that is proportional to the boundary-interior gradient,
424 as proposed in section 3.2.

425   On decadal time scales, variability in oxygen storage can be an important factor in controlling the linkage
426 between uptake and export. Period 1 and 2 are shaded in Fig. 5A in green and red, respectively, and
427 mean values of the three terms over each period are shown in table 3, with the values for individual
428 models summarized in table S2. By definition, the storage term is generally positive in the first period
429 ($1.07 \ mol \ m^{-2} \ y^{-1}$) and negative in the latter (-$1.22 \ mol \ m^{-2} \ y^{-1}$), although there is also interannual and
430 inter-model variability around this mean. The period of positive storage changes is accompanied by the
431 largest gas exchange over the model record ($9.60 \ mol \ m^{-2} \ y^{-1}$), while oxygen uptake during the second
432 period is $6.89 \ mol \ m^{-2} \ y^{-1}$, slightly less than the long-term average. The difference in gas exchange
433 exceeds the difference in oxygen storage because the oxygen export is also stronger in Period 1 than Period
434 2, changing from $8.52 \ mol \ m^{-2} \ y^{-1}$ to $8.11 \ mol \ m^{-2} \ y^{-1}$. The modelled data suggest that the period
435 of record strong convection in the 1980s and early 1990s featured the record strongest gas exchange and
436 oxygen storage (Fig. 5A), as well as the record oxygen export (not shown).

437   Although both oxygen uptake and oxygen export are decreased in Period 2 compared to their record
438 values in Period 1, the change in the latter is much smaller: A $2.71 \ mol \ m^{-2} \ y^{-1}$ change in gas exchange
439 only corresponds to a change in oxygen export by $0.40 \ mol \ m^{-2} \ y^{-1}$, or 14% of the decreased uptake.
440 The remainder is balanced by a decrease in oxygen storage of $2.29 \ mol \ m^{-2} \ y^{-1}$, 85% of oxygen uptake
441 change, as the oxygen reservoir switches from accumulating excess oxygen to exporting more than is
442 added. This suggests that although the dominant balance in the long-term mean is between gas exchange
443 and export, oxygen storage in the interior Labrador Sea can be important for modulating their variability on
444 decadal time scales, acting as a buffer between increased uptake and increased export. The switch between
445 the two periods from positive to negative oxygen storage, and stronger to weaker gas exchange, are robust
446 across all models (table S2). The change in export is negative in most models, but there is a wider range of
447 behaviors, likely reflecting differences in model physics such as ocean circulation.

448   Our analysis of the CMIP6-OMIP1 data thus suggests that although much of the oxygen taken up from
449 the atmosphere is exported out of the formation region, there is also an accumulation, or storage, of oxygen
450 in the interior Labrador Sea during periods of high oxygen uptake, consistent with a multi-year residence
451 time of LSW (Straneo et al., 2003). This oxygen storage acts as a buffer, partially decoupling the export of
452 oxygen out of the region from the uptake via gas exchange, dampening the impact of variations in oxygen
453 uptake on export, and ensuring a more steady deep water oxygen supply to the rest of the ocean.

5   IMPLICATIONS

5.1   Ventilation and deoxygenation in CMIP6-OMIP1
454

455   Our analysis shows that contemporary CMIP6-OMIP1 models struggle to accurately portray the
456 magnitude of oxygen content changes, the mean vertical profile of oxygen, and mean air-sea gas exchange
457 in the Labrador Sea (section 4.1, table 2), suggesting a need for improvement of modeled oxygen cycles in
458 this region.

459   This is particularly true given the importance of the Labrador Sea and the subpolar North Atlantic
460 as a whole for ventilating the global ocean. A major shortcoming appears to be the representation of
461 air-sea gas exchange, with CMIP6-OMIP1 models estimating the annual net oxygen uptake to be only
462 about a third of observed values. This is in line with previous studies in the region that had shown the

463 gas exchange parameterization used in CMIP6-OMIP1 models to severely underestimate oxygen uptake
464 during deep convection in winter (Koelling et al., 2017; Atamanchuk et al., 2020; Seltzer et al., 2023),
465 and could be responsible for the reduced magnitude of oxygen content changes compared to observations.
466 Another contributing factor may be the difficulties that some CMIP6-OMIP1 models have with modelling
467 convection to realistic depths (Fig. 4) and representing overflow water masses (Fig. 3C), which can lead to
468 biases in oxygen distributions throughout the water column, including surface oxygen being too close to
469 saturation during deep convection.

470 Since gas exchange is balanced primarily by oxygen export out of the region this weaker oxygen uptake
471 would also result in too little supply of oxygen from the deep water formation area to the rest of the global
472 ocean in CMIP6-OMIP1 models. Given that the bubble-mediated gas flux is important for air-sea gas
473 exchange in the other major deep ocean ventilation region, the Southern Ocean (Bushinsky et al., 2017), it
474 may be expected that oxygen uptake, storage, and export there are likewise underestimated. Consequently,
475 the mean strength of deep ocean ventilation in both subpolar regions as currently represented in global
476 climate models, as well its variability, may be too small. This would be consistent with a recent finding that
477 climate models underestimate ventilation changes (Buchanan and Tagliabue, 2021), and could partially
478 explain discrepancies in ocean deoxygenation between models and observations (Oschlies et al., 2017,
479 2018). Furthermore, if the oxygen supply to the deep ocean is underestimated in global climate models,
480 then its consumption by marine organisms would have to also be too weak in order to avoid spurious trends
481 in global oxygen concentrations in the interior ocean. This might have important implications not only for
482 future ocean deoxygenation, but also for model-based studies of the biological carbon pump (BCP), as too
483 little oxygen consumption would correspond to too little carbon remineralization in the deep ocean. An
484 improved representation of deep ocean ventilation with oxygen therefore seems imperative for predicting
485 the evolution of both deoxygenation and the BCP in a changing ocean.

5.2 Gas exchange and oxygen export

487 Despite shortcomings, one aspect that the CMIP6-OMIP1 models do reproduce fairly consistently in the
488 Labrador Sea is the interannual and decadal variability of the oxygen content. Although the magnitude is
489 too weak, the correlation of the time series with the observational dataset is 0.73 for the ensemble mean
490 (table S1), and all models show positive oxygen storage from 1980 to 1992, and negative storage from 1994
491 to 2006 (table S2). Given that the oxygen content change is related to variability in air-sea gas exchange
492 (Section 4.2), this may suggest that although the magnitude of the gas exchange is too low in the models,
493 its temporal variability (Fig. 3B) more accurately reproduces that of the real ocean.

494 The mean balance in the oxygen budget in both observations and models is largely between gas exchange
495 and export: There is almost no long-term change in storage over the 60-year period that was analyzed
496 (sections 3.1, 4.2), meaning that the average amount of oxygen exported is equivalent to the average
497 amount taken up from the atmosphere. However, most of the decadal variability in gas exchange seen in the
498 CMIP6-OMIP1 models is imprinted onto the oxygen content, rather than the export. Between the strong
499 convection period 1980–1992 and the weak convection period 1994–2006, the mean annual gas exchange
500 decreases by $2.71\ mol\ m^{-2}$ (28% change from Period 1 mean) in CMIP6-OMIP1 models, 85% of which
501 is balanced by the change in the annual storage rate ($2.29\ mol\ m^{-2}$), while only 14% of the decreased
502 gas exchange is reflected in the oxygen export ($0.40\ mol\ m^{-2}$, 5% of Period 1 mean). Thus, while the
503 interannual to decadal variability in oxygen uptake does affect the export, its magnitude is dampened by
504 the accumulation and subsequent drainage of oxygen in the central Labrador Sea. Körtzinger et al. (2004),
505 using one of the first deployments of biogeochemical Argo floats, referred to the wintertime oxygen

506 uptake in the Labrador Sea as the ocean "taking a deep breath". Our results show that the multi-year
507 residence time of LSW in the interior of the basin leads to the ocean "holding its breath", storing some of
508 the excess oxygen taken up during strong convection years, and exporting it in subsequent years.

509   This relative insensitivity of the export to atmospheric forcing reported here differs from estimates based
510 on the change in layer thickness of LSW, which suggest almost a factor of 3 difference in the volume of
511 LSW export between strong and weak convection years (Yashayaev and Loder, 2016). The difference may
512 stem from the fact that their analysis highlights the export of LSW over a 1-year period following the
513 previous convection season, while the oxygen-based estimate discussed here includes the export of older
514 LSW as well. During particularly strong convection periods, stronger gas exchange does lead to increased
515 export, but also increases the oxygen content in the formation region since some of the newly formed LSW
516 stays in the region from one year to the next (Straneo et al., 2003). This "storage" affects the export in
517 subsequent years, such that some of the LSW and oxygen export from a strong convection period continues
518 even after convection has weakened and the oxygen uptake has decreased.

519   On the other hand, if convection in the Labrador Sea were to continuously weaken or stop altogether as
520 a result of climate change, as has been suggested could occur (e.g. Oltmanns et al. (2018)), the oxygen
521 uptake, and therefore also the export, would likely decrease dramatically. The shallower mixed layer
522 would reach equilibrium quickly, limiting the amount of oxygen that can be taken up, while the absence of
523 convection would also impede any supply of this atmospheric oxygen to the deep ocean. Due to the time
524 lag between oxygen uptake and export introduced by the Labrador Sea's oxygen reservoir, there would
525 still be a net flow of oxygen out of the basin for some time, but its rate would continuously decrease as the
526 oxygen stored in the interior Labrador Sea slowly drains out along LSW export pathways. Eventually, the
527 deep oxygen reservoir would disappear, and the Labrador Sea would no longer act as a source of oxygen
528 for the deep ocean, with potentially grave impacts for marine ecosystems as the ocean loses its breath.

529   One key question that remains is how uptake and export of oxygen in the Labrador Sea are related
530 to the strength of the Atlantic Meridional Oveturning Circulation (AMOC). The strength of the AMOC
531 varies on decadal time scales (Moat et al., 2020; Jackson et al., 2022; Le Bras et al., 2023), and future
532 projections with CMIP models suggest that it will weaken over the next century (Bakker et al., 2016;
533 Weijer et al., 2020). In the Labrador Sea, this would mean a decreased strength of both the outflowing
534 boundary current, which is responsible for some of the oxygen export (Koelling et al., 2022), and the flow
535 of lower-oxygen water into the basin. In an analysis of decadal changes in oxygen concentrations in the
536 neighboring Irminger Sea, Feucher et al. (2022) found that changes in ocean circulation play an important
537 role in controlling oxygen variability. In contrast, our analysis of CMIP6-OMIP1 output suggests that the
538 oxygen export does not play a major role for the oxygen content in the interior Labrador Sea, with most
539 of the variability in both oxygen content and export set by changes in the uptake from the atmosphere.
540 However, while this is true for the multi-model mean, individual models do have a larger correlation
541 between oxygen storage and export, suggesting that this finding is dependent on model physics. A more
542 detailed analysis of the differences in circulation between the models is beyond the scope of the current
543 study, but could provide further insight into whether there is a linkage between the AMOC and oxygen
544 transports. Our understanding of these processes will also benefit from the addition of oxygen sensors to
545 the OSNAP mooring array, which will enable calculation of cross-basin oxygen transports associated with
546 the overturning circulation (Atamanchuk et al., 2022).

547   In order to be able to monitor and predict the future variability in ventilation in this basin and the
548 subpolar North Atlantic as a whole, the long-term context provided by the present study will need to
549 be combined with more detailed analyses of the present-day oxygen uptake and export on basin-wide

550 scales. In recent years, there have been regular oxygen measurements from gliders in the Labrador Sea
551 (von Oppeln-Bronikowski et al., 2021), the use of moored oxygen sensors in the SPNA has increased
552 dramatically (Atamanchuk et al., 2022), and with the proliferation of BGC-Argo floats throughout the
553 ocean, full global 4-D fields of oxygen based on observational data are now becoming available (Sharp
554 et al., 2022). This wealth of data concentrated in a small but key area of the ocean could be exploited in
555 the future to study ocean ventilation by constraining the uptake, storage, and export of oxygen in much
556 more detail than was possible in the present study, and guide improvement in how these processes are
557 implemented in climate models. Given our current lack of knowledge about the mechanisms driving
558 deoxygenation and changes in deep ocean ventilation (Oschlies et al., 2018), improving our understanding
559 of these topics, and their representation in models, will be imperative for reliably predicting how they may
560 change in the future.

CONFLICT OF INTEREST STATEMENT

561 The authors declare that the research was conducted in the absence of any commercial or financial
562 relationships that could be construed as a potential conflict of interest.

AUTHOR CONTRIBUTIONS

563 Data analysis for this study was carried out by JKo. The manuscript was prepared by JKo, with contributions
564 from all other coauthors.

FUNDING

ACKNOWLEDGMENTS

SUPPLEMENTAL DATA

574 Supplementary Material contains figures S1 and S2 and tables S1 and S2

DATA AVAILABILITY STATEMENT

575 The datasets analyzed for this study are freely available online and can be accessed at
576 https://www.ncei.noaa.gov/products/world-ocean-database (WOD) and https://esgf-node.llnl.gov/search/cmip6/
577 (CMIP6). The code used for calculating the model score described in Section 2.4 is available at
578 https://github.com/jkoell/labrador-sea-oxygen.

REFERENCES

579 Adcroft, A., Anderson, W., Balaji, V., Blanton, C., Bushuk, M., Dufour, C. O., et al. (2019). The gfdl
580   global ocean and sea ice model om4.0: Model description and simulation features. *Journal of Advances*
581   *in Modeling Earth Systems* 11, 3167–3211. doi:10.1029/2019MS001726

582 Anav, A., Friedlingstein, P., Kidston, M., Bopp, L., Ciais, P., Cox, P., et al. (2013). Evaluating the Land
583   and Ocean Components of the Global Carbon Cycle in the CMIP5 Earth System Models. *Journal of*
584   *Climate* 26, 6801 – 6843. doi:0.1175/JCLI-D-12-00417.1

585 Atamanchuk, D., Koelling, J., Send, U., and Wallace, D. (2020). Rapid transfer of oxygen to the deep
586   ocean mediated by bubbles. *Nature Geoscience* , 1–6doi:10.1038/s41561-020-0532-2

587 Atamanchuk, D., Palter, J., Palevsky, H., Le Bras, I., Koelling, J., and Nicholson, D. (2022). Linking
588   Oxygen and Carbon Uptake with the Meridional Overturning Circulation Using a Transport Mooring
589   Array. In *Frontiers in Ocean Observing: Documenting Ecosystems, Understanding Environmental*
590   *Changes, Forecasting Hazards*, eds. E. Kappel, S. Juniper, S. Seeyave, E. Smith, and M. Visbeck. 9.
591   doi:10.5670/oceanog.2021.supplement.02-03

592 Bakker, P., Schmittner, A., Lenaerts, J. T. M., Abe-Ouchi, A., Bi, D., van den Broeke, M. R., et al. (2016).
593   Fate of the Atlantic Meridional Overturning Circulation: Strong decline under continued warming and
594   Greenland melting. *Geophysical Research Letters* 43, 12,252–12,260. doi:10.1002/2016GL070457

595 Boucher, O., Servonnat, J., Albright, A. L., Aumont, O., Balkanski, Y., Bastrikov, V., et al. (2020).
596   Presentation and evaluation of the ipsl-cm6a-lr climate model. *Journal of Advances in Modeling Earth*
597   *Systems* 12, e2019MS002010. doi:10.1029/2019MS002010. E2019MS002010 10.1029/2019MS002010

598 Boyer, T. P., Antonov, J. I., Baranova, O. K., Coleman, C., Garcia, H. E., Grodsky, A., et al. (2013). World
599   ocean database 2013 doi:10.7289/V5NZ85MT

600 Buchanan, P. J. and Tagliabue, A. (2021). The Regional Importance of Oxygen Demand and Supply for
601   Historical Ocean Oxygen Trends. *Geophysical Research Letters* 48, e2021GL094797. doi:10.1029/
602   2021GL094797. E2021GL094797 2021GL094797

603 Bushinsky, S. M., Gray, A. R., Johnson, K. S., and Sarmiento, J. L. (2017). Oxygen in the Southern
604   Ocean From Argo Floats: Determination of Processes Driving Air-Sea Fluxes. *Journal of Geophysical*
605   *Research: Oceans* 122, 8661–8682. doi:10.1002/2017JC012923

606 Collins, M., Knutti, R., Arblaster, J., Dufresne, J.-L., Fichefet, T., Friedlingstein, P., et al. (2014). Long-term
607   Climate Change: Projections, Commitments and Irreversibility. In *Climate Change 2013 – The Physical*
608   *Science Basis: Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental*
609   *Panel on Climate Change* (Cambridge University Press). 1029–1136. doi:10.1017/CBO9781107415324.
610   024

611 Feucher, C., Portela, E., Kolodziejczyk, N., and Thierry, V. (2022). Subpolar gyre decadal variability
612   explains the recent oxygenation in the Irminger Sea. *Communications Earth & Environment* 3, 279.
613   doi:10.1038/s43247-022-00570-y

614 Georgiou, S., Ypma, S. L., Brüggemann, N., Sayol, J.-M., van der Boog, C. G., Spence, P., et al. (2020).
615   Direct and indirect pathways of convected water masses and their impacts on the overturning dynamics
616   of the Labrador Sea. *Journal of Geophysical Research: Oceans* n/a, e2020JC016654. doi:10.1029/
617   2020JC016654

618 Gleckler, P. J., Taylor, K. E., and Doutriaux, C. (2008). Performance metrics for climate models. *Journal*
619   *of Geophysical Research: Atmospheres* 113. doi:10.1029/2007JD008972

620 Heuzé, C. (2017). North Atlantic deep water formation and AMOC in CMIP5 models. *Ocean Science* 13,
621   609–622. doi:10.5194/os-13-609-2017

Heuzé, C. (2021). Antarctic Bottom Water and North Atlantic Deep Water in CMIP6 models. *Ocean Science* 17, 59–90. doi:10.5194/os-17-59-2021

Huang, J., Huang, J., Liu, X., Li, C., Ding, L., and Yu, H. (2018). The global oxygen budget and its future projection. *Science Bulletin* 63, 1180–1186. doi:https://doi.org/10.1016/j.scib.2018.07.023

Ito, T., Minobe, S., Long, M. C., and Deutsch, C. (2017). Upper ocean O2 trends: 1958–2015. *Geophysical Research Letters* 44, 4214–4223. doi:10.1002/2017GL073613

Jackson, L. and Petit, T. (2022). North Atlantic overturning and water mass transformation in CMIP6 models. *Climate Dynamics* , 1–21doi:10.1007/s00382-022-06448-1

Jackson, L. C., Biastoch, A., Buckley, M. W., Desbruyères, D. G., Frajka-Williams, E., Moat, B., et al. (2022). The evolution of the North Atlantic meridional overturning circulation since 1980. *Nature Reviews Earth & Environment* 3, 241–254. doi:10.1038/s43017-022-00263-2

Karstensen, J., Stramma, L., and Visbeck, M. (2008). Oxygen minimum zones in the eastern tropical Atlantic and Pacific oceans. *Progress in Oceanography* 77, 331–350. doi:10.1016/j.pocean.2007.05.009

Kieke, D. and Yashayaev, I. (2015). Studies of Labrador Sea Water formation and variability in the subpolar North Atlantic in the light of international partnership and collaboration. *Progress in Oceanography* 132, 220–232. doi:10.1016/j.pocean.2014.12.010. Oceanography of the Arctic and North Atlantic Basins

Koelling, J., Atamanchuk, D., Karstensen, J., Handmann, P., and Wallace, D. W. R. (2022). Oxygen export to the deep ocean following Labrador Sea Water formation. *Biogeosciences* 19, 437–454. doi:10.5194/bg-19-437-2022

Koelling, J., Wallace, D. W. R., Send, U., and Karstensen, J. (2017). Intense oceanic uptake of oxygen during 2014–2015 winter convection in the Labrador Sea. *Geophysical Research Letters* 44, 7855–7864. doi:10.1002/2017GL073933

Koenigk, T., Fuentes-Franco, R., Meccia, V. L., Gutjahr, O., Jackson, L. C., New, A. L., et al. (2021). Deep mixed ocean volume in the Labrador Sea in HighResMIP models. *Climate Dynamics* 57, 1895–1918. doi:10.1007/s00382-021-05785-x

Körtzinger, A., Schimanski, J., Send, U., and Wallace, D. (2004). The Ocean Takes a Deep Breath. *Science* 306, 1337–1337. doi:10.1126/science.1102557

Körtzinger, A., Send, U., Wallace, D. W., Karstensen, J., and DeGrandpre, M. (2008). Seasonal cycle of O2 and pCO2 in the central Labrador Sea: Atmospheric, biological, and physical implications. *Global Biogeochemical Cycles* 22

Large, W. and Yeager, S. (2009). The global climatology of an interannually varying air–sea flux data set. *Climate dynamics* 33, 341–364. doi:10.1007/s00382-008-0441-3

Laurent, A., Fennel, K., and Kuhn, A. (2021). An observation-based evaluation and ranking of historical Earth system model simulations in the northwest North Atlantic Ocean. *Biogeosciences* 18, 1803–1822. doi:10.5194/bg-18-1803-2021

Le Bras, I. A.-A., Willis, J., and Fenty, I. (2023). The Atlantic Meridional Overturning Circulation at 35°N From Deep Moorings, Floats, and Satellite Altimeter. *Geophysical Research Letters* 50, e2022GL101931. doi:10.1029/2022GL101931

Martz, T. R., DeGrandpre, M. D., Strutton, P. G., McGillis, W. R., and Drennan, W. M. (2009). Sea surface pCO2 and carbon export during the Labrador Sea spring-summer bloom: An in situ mass balance approach. *Journal of Geophysical Research: Oceans* 114. doi:https://doi.org/10.1029/2008JC005060

Moat, B. I., Smeed, D. A., Frajka-Williams, E., Desbruyères, D. G., Beaulieu, C., Johns, W. E., et al. (2020). Pending recovery in the strength of the meridional overturning circulation at 26 N. *Ocean Science* 16, 863–874. doi:10.5194/os-16-863-2020

Oltmanns, M., Karstensen, J., and Fischer, J. (2018). Increased risk of a shutdown of ocean convection posed by warm North Atlantic summers. *Nature Climate Change* 8, 300–304. doi:10.1038/s41558-018-0105-1

Orr, J. C., Najjar, R. G., Aumont, O., Bopp, L., Bullister, J. L., Danabasoglu, G., et al. (2017). Biogeochemical protocols and diagnostics for the CMIP6 Ocean Model Intercomparison Project (OMIP). *Geoscientific Model Development* 10, 2169–2199. doi:10.5194/gmd-10-2169-2017

Oschlies, A., Brandt, P., Stramma, L., and Schmidtko, S. (2018). Drivers and mechanisms of ocean deoxygenation. *Nature Geoscience* 11, 467–473. doi:10.1038/s41561-018-0152-2

Oschlies, A., Duteil, O., Getzlaff, J., Koeve, W., Landolfi, A., and Schmidtko, S. (2017). Patterns of deoxygenation: sensitivity to natural and anthropogenic drivers. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 375, 20160325. doi:10.1098/rsta.2016.0325

Portela, E., Kolodziejczyk, N., Vic, C., and Thierry, V. (2020). Physical Mechanisms Driving Oxygen Subduction in the Global Ocean. *Geophysical Research Letters* 47, e2020GL089040. doi:10.1029/2020GL089040

Rhein, M., Steinfeldt, R., Kieke, D., Stendardo, I., and Yashayaev, I. (2017). Ventilation variability of Labrador Sea Water and its impact on oxygen and anthropogenic carbon: a review. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 375, 20160321. doi:10.1098/rsta.2016.0321

Rickard, G. J., Behrens, E., and Chiswell, S. M. (2016). CMIP5 earth system models with biogeochemistry: An assessment for the southwest Pacific Ocean. *Journal of Geophysical Research: Oceans* 121, 7857–7879. doi:10.1002/2016JC011736

Schmidtko, S., Stramma, L., and Visbeck, M. (2017). Decline in global oceanic oxygen content during the past five decades. *Nature* 542, 335–339. doi:10.1038/nature21399

Seltzer, A. M., Nicholson, D. P., Smethie, W. M., Tyne, R. L., Roy, E. L., Stanley, R. H. R., et al. (2023). Dissolved gases in the deep North Atlantic track ocean ventilation processes. *Proceedings of the National Academy of Sciences* 120, e2217946120. doi:10.1073/pnas.2217946120

Sharp, J. D., Fassbender, A. J., Carter, B. R., Johnson, G. C., Schultz, C., and Dunne, J. P. (2022). GOBAI-O$_2$: temporally and spatially resolved fields of ocean interior dissolved oxygen over nearly two decades. *Earth System Science Data Discussions* 2022, 1–46. doi:10.5194/essd-2022-308

Stendardo, I. and Gruber, N. (2012). Oxygen trends over five decades in the North Atlantic. *Journal of Geophysical Research: Oceans* 117. doi:10.1029/2012JC007909

Straneo, F. (2006). Heat and Freshwater Transport through the Central Labrador Sea. *Journal of Physical Oceanography* 36, 606 – 628. doi:10.1175/JPO2875.1

Straneo, F., Pickart, R. S., and Lavender, K. (2003). Spreading of Labrador sea water: an advective-diffusive study based on Lagrangian data. *Deep Sea Research Part I: Oceanographic Research Papers* 50, 701–719. doi:10.1016/S0967-0637(03)00057-8

Sy, A., Rhein, M., Lazier, J. R., Koltermann, K. P., Meincke, J., Putzka, A., et al. (1997). Surprisingly rapid spreading of newly formed intermediate waters across the North Atlantic Ocean. *Nature* 386, 675–679. doi:10.1038/386675a0

Tagklis, F., Bracco, A., Ito, T., and Castelao, R. (2020). Submesoscale modulation of deep water formation in the Labrador Sea. *Scientific reports* 10, 17489. doi:10.1038/s41598-020-74345-w

Talley, L. and McCartney, M. (1982). Distribution and Circulation of Labrador Sea Water. *J. Phys. Oceanography* 1. doi:10.1175/1520-0485(1982)012⟨1189:DACOLS⟩2.0.CO;2

van Aken, H. M., Femke de Jong, M., and Yashayaev, I. (2011). Decadal and multi-decadal variability of Labrador Sea Water in the north-western North Atlantic Ocean derived from tracer distributions: Heat

711   budget, ventilation, and advection. *Deep Sea Research Part I: Oceanographic Research Papers* 58,
712   505–523. doi:https://doi.org/10.1016/j.dsr.2011.02.008

713 von Oppeln-Bronikowski, N., de Young, B., Atamanchuk, D., and Wallace, D. (2021). Glider-based
714   observations of CO2 in the Labrador Sea. *Ocean Science* 17, 1–16. doi:10.5194/os-17-1-2021

715 Wanninkhof, R. (1992). Relationship between wind speed and gas exchange over the ocean. *Journal of*
716   *Geophysical Research: Oceans* 97, 7373–7382. doi:10.1029/92JC00188

717 Wanninkhof, R. (2014). Relationship between wind speed and gas exchange over the ocean revisited.
718   *Limnology and Oceanography: Methods* 12, 351–362. doi:10.4319/lom.2014.12.351

719 Weijer, W., Cheng, W., Garuba, O. A., Hu, A., and Nadiga, B. T. (2020). CMIP6 Models Predict Significant
720   21st Century Decline of the Atlantic Meridional Overturning Circulation. *Geophysical Research Letters*
721   47, e2019GL086075. doi:10.1029/2019GL086075

722 Westberry, T., Behrenfeld, M. J., Siegel, D. A., and Boss, E. (2008). Carbon-based primary productivity
723   modeling with vertically resolved photoacclimation. *Global Biogeochemical Cycles* 22. doi:10.1029/
724   2007GB003078

725 Wolf, M. K., Hamme, R. C., Gilbert, D., Yashayaev, I., and Thierry, V. (2018). Oxygen Saturation
726   Surrounding Deep Water Formation Events in the Labrador Sea From Argo-O2 Data. *Global*
727   *Biogeochemical Cycles* 32, 635–653. doi:10.1002/2017GB005829

728 Yashayaev, I. (2007). Hydrographic changes in the Labrador Sea, 1960-2005. *Progress in Oceanography*
729   75. doi:10.1016/j.pocean.2007.04.015

730 Yashayaev, I. and Loder, J. W. (2009). Enhanced production of Labrador Sea water in 2008. *Geophysical*
731   *Research Letters* 36. doi:10.1029/2008GL036162

732 Yashayaev, I. and Loder, J. W. (2016). Recurrent replenishment of Labrador Sea Water and associated
733   decadal-scale variability. *Journal of Geophysical Research: Oceans* 121, 8095–8114. doi:10.1002/
734   2016JC012046

735 Yashayaev, I. and Loder, J. W. (2017). Further intensification of deep convection in the Labrador Sea in
736   2016. *Geophysical Research Letters* 44, 1429–1438. doi:10.1002/2016GL071668

737 Yashayaev, I., Peterson, I., and Wang, Z. (2020). Meteorological, Sea Ice, and Physical Oceanographic
738   Conditions in the Labrador Sea during 2019. *DFO Can. Sci. Advis. Sec. Res. Doc*

739 Zou, S., Lozier, M. S., Li, F., Abernathey, R., and Jackson, L. (2020). Density-compensated overturning in
740   the Labrador Sea. *Nature Geoscience* 13, 121–126. doi:10.1038/s41561-019-0517-1

FIGURE CAPTIONS
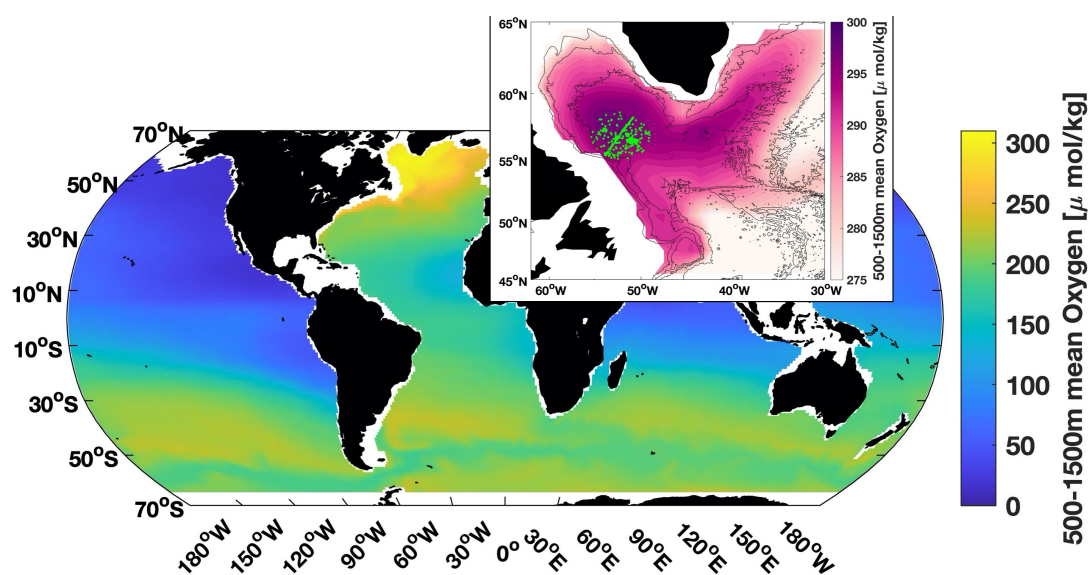
Figure 1. Global mean dissolved oxygen concentration between 500–1500 m depth from the GOBAI-O$_2$ gridded oxygen product (Sharp et al., 2022). Inset shows detail in the Labrador Sea, with green symbols representing measurements used to construct the time series in Section 2.
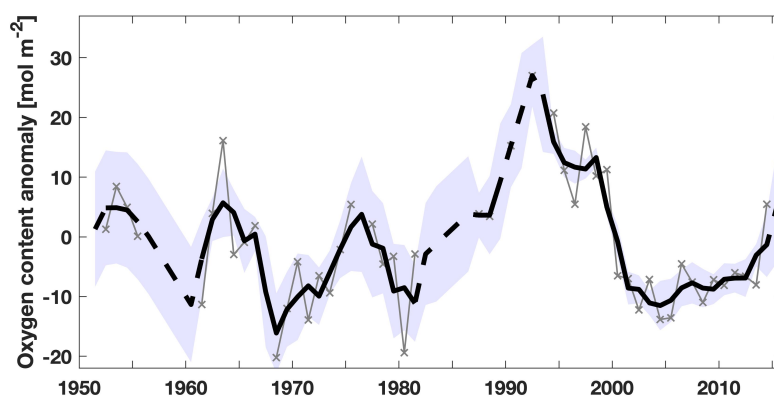

Figure 2. Time series of oxygen inventory anomaly from WOD data. Thin gray lines and symbols show annual values. The thick black line shows the time series smoothed with a 3-year running mean, and is dashed where gaps were filled. Blue shading represents the estimated error.
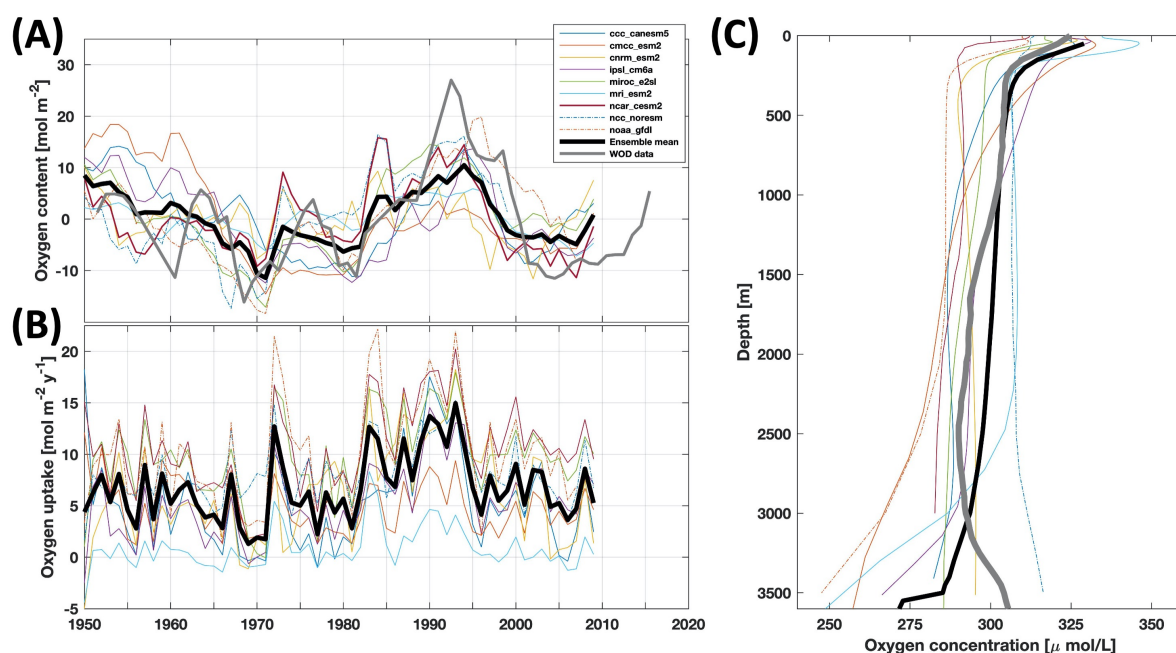
Figure 3. Comparison of CMIP6-OMIP1 model data from the central Labrador Sea to observations. In all panels, the thick black line shows the mean of all models, and in panels (A) and (C) the grey line shows the observational data from WOD. (A) Oxygen content anomaly (B) Annual mean air-sea gas exchange of oxygen (C) Mean profiles of oxygen concentration
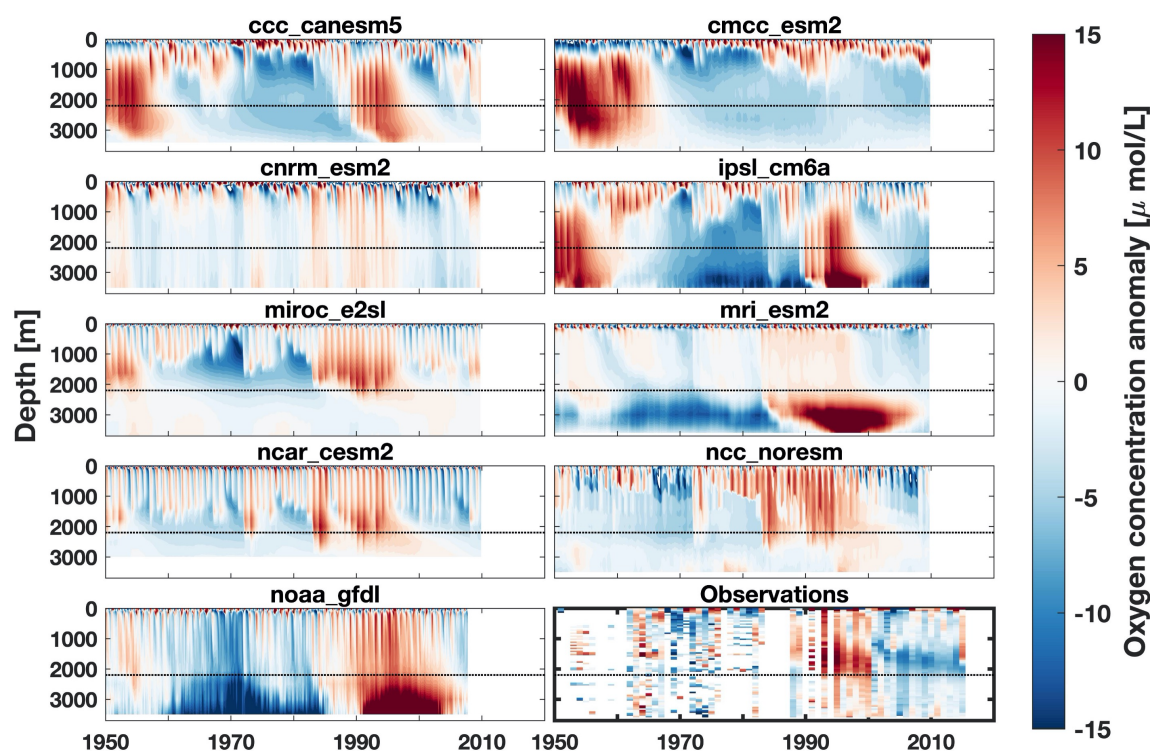
Figure 4. Oxygen concentration anomaly contour plots for all CMIP6-OMIP1 models used, and the observational estimate. Anomalies are calculated relative to the mean at each depth level. The thin dashed line in each panel is at 2200 m, the integration depth for the oxygen inventory time series.
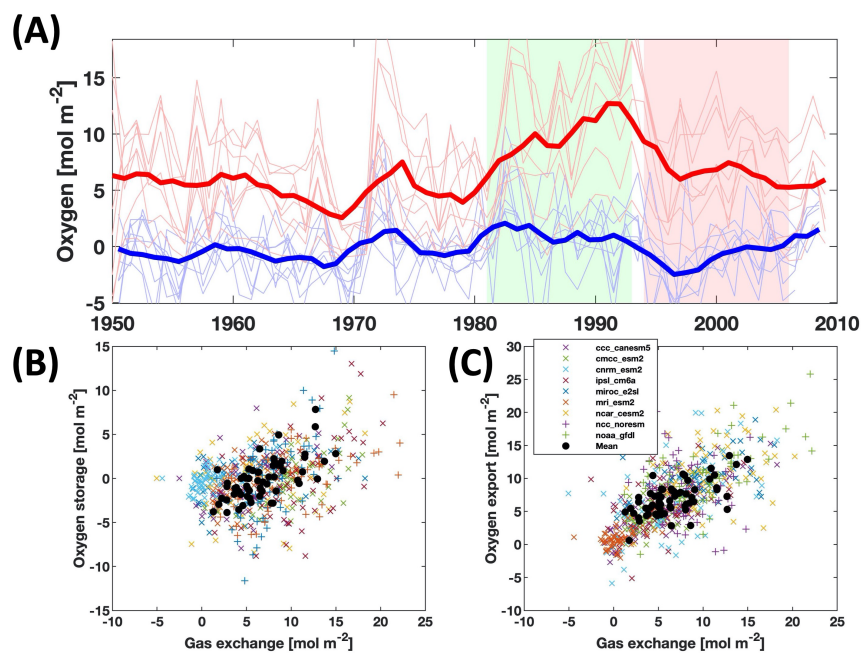
Figure 5. (A) Time series of annual gas exchange (red) and oxygen storage (blue) from the CMIP6-OMIP1 models used in this study. Thin lines show individual models, and thick lines show the ensemble mean, smoothed with a 5-year running mean filter. Mean values for the two periods are given in Table 3. (B) Relationship between annual gas exchange and oxygen storage (C) Relationship between annual gas exchange and export. In panels (B) and (C), individual models are shown as colored symbols, and black circles show the multi-model mean.

TABLES

Table 1. Summary of models used in this study. All model runs used are OMIP1.

Center	CCC	CMCC	CNRM	IPSL	MIROC	MRI	NCAR	NCC	NOAA GFDL
Model	CanESM5	ESM2	ESM2-1	CM6A-LR	E2SL	ESM2.0	CESM2	NorESM2-LM	GFDL-CM4
Variant	r1i1p2f1	r1i1p1f1	r1i1p1f2	r1i1p1f1	r1i1p1f2	r1i1p1f1	r1i1p1f1	r1i1p1f1	r1i1p1f1

Table 2. Scores for each of the five components used in constructing the model score, and total score for each model. The factors are correlation of the LSW layer (0–2200 m) oxygen content with observations, the amplitude of the 1990s maximum in LSW oxygen content, standard deviation of the oxygen content below the LSW layer (2200 m–bottom), mean absolute difference between the modeled and observed oxygen profile for the LSW layer and 2200m bottom layer, and mean gas exchange. Values for the ensemble mean obtained by averaging together the output fields from all models are given at the bottom. The values used for calculating each component of the score are given in table S1 in the supplementary material.

Model	Corr.	LSW Max	bottom σ	Mean O2 diff	Gas exch.	Score
CCC CanESM5	5	11	15	13	6	**50**
CMCC ESM2	2	5	13	8	5	**33**
CNRM ESM2	0	8	11	16	7	**42**
IPSL CM6A	6	11	7	15	6	**45**
MIROC E2SL	11	9	6	14	11	**51**
MRI ESM2	16	6	7	11	2	**42**
NCAR CESM2	8	14	7	11	11	**51**
NCC NorESM	11	15	10	8	9	**53**
NOAA GFDL	15	11	5	3	11	**45**
Ensemble mean	13	8	16	13	7	**57**

Table 3. Terms of the oxygen budget (eq. 1) for Period 1 and Period 2

Time periods	Storage [$mol\ m^{-2}\ y^{-1}$]	Uptake [$mol\ m^{-2}\ y^{-1}$]	Export [$mol\ m^{-2}\ y^{-1}$]
1980-1992	1.07	9.60	8.52
1994-2006	-1.22	6.89	8.11