Ocean biogeochemistry in the Canadian Earth System Model version 5.0.3: CanESM5 and CanESM5-CanOE

James R. Christian[1,2], Kenneth L. Denman[2,3], Hakase Hayashida[3,4], Amber M. Holdsworth[1], Warren G. Lee[2], Olivier G.J. Riche[3,5], Andrew E. Shao[2,3], Nadja Steiner[1,2], and Neil C. Swart[2]

1 Fisheries and Oceans Canada, Sidney, BC, Canada

2 Canadian Centre for Climate Modelling and Analysis, Victoria, BC, Canada

3 School of Earth and Ocean Sciences, University of Victoria, Victoria, BC, Canada

4 now at the Institute for Marine and Antarctic Studies, University of Tasmania, Hobart, Tasmania, Australia

5 now at Fisheries and Oceans Canada, Mont Joli, Québec, Canada

Abstract. The ocean biogeochemistry components of two new versions of the Canadian Earth

System Model ~~v. 5~~ are presented and compared to observations and other models. CanESM5

employs the same ~~biogeochemistry module~~ ocean biology model as CanESM2 whereas

CanESM5-CanOE ("Canadian Ocean Ecosystem model") is a new, more complex

~~biogeochemistry module~~model developed for CMIP6, with multiple food chains, flexible

phytoplankton elemental ratios, and a prognostic iron cycle. This new model is described in

detail and the outputs (distributions of major tracers such as oxygen, dissolved inorganic carbon,

and alkalinity, the iron and nitrogen cycles, plankton biomass, and historical trends in CO_2

uptake and export production) compared to CanESM5 and CanESM2, as well as to observations

and other CMIP6 models. Both CanESM5 models show gains in skill relative to CanESM2,

which are attributed primarily to improvements in ocean circulation. CanESM5-CanOE shows

improved skill relative to CanESM5 ~~in some areas~~for most major tracers at most depths.

CanESM5-CanOE includes a prognostic iron cycle, and maintains high nutrient / low

chlorophyll conditions in the expected regions (in CanESM2 and CanESM5, iron limitation is

specified as a temporally static 'mask'). Surface nitrate concentrations are biased low in the

subarctic Pacific and equatorial Pacific, and high in the Southern Ocean, in both CanESM5 and

CanESM5-CanOE. Export production in CanESM5-CanOE is among the lowest for CMIP6

models; in CanESM5 it is among the highest, but shows the most rapid decline after about 1980.

CanESM5-CanOE shows some ability to simulate aspects of plankton community structure that a

single-species model can not (e.g., seasonal dominance of large cells), but is biased towards ~~has~~

~~relatively~~ low concentrations of zooplankton and detritus relative to phytoplankton.~~, and a high~~

~~and relatively constant living phytoplankton fraction of total particulate organic matter. In most~~

~~regions, large and small phytoplankton show decoupled seasonal cycles with greater abundance~~

42 ~~of large phytoplankton in the productive seasons.~~ Cumulative ocean uptake of anthropogenic

43 carbon dioxide through 2014 is lower in both CanESM5 models than in observation-based

44 estimates (145 PgC) or the model ensemble mean (144 PgC), and is lower in CanESM5-CanOE

45 (122 PgC) than in CanESM5 (132 PgC).

1. **Introduction**

The Canadian Centre for Climate Modelling and Analysis has been developing coupled models with an interactive carbon cycle for more than a decade (Arora et al., 2009; 2011; Christian et al., 2010). The Canadian Earth System Model version 5 (CanESM5 ~~(,~~ Swart et al., 2019a) is an updated version of CanESM2 (Arora et al., 2011), with a~~n entirely~~ new ocean model ~~and an atmosphere with the same T63 horizontal resolution and important improvements in atmospheric physics. The CanESM5 ocean is~~ based on the Nucleus for European Modelling of the Ocean (NEMO) system version 3.4. The ocean biogeochemistry modules were developed in-house~~, although parameterizations for some processes were adapted from the native PISCES biogeochemistry model (Aumont et al., 2015)~~. CanESM5 uses the same ocean ~~biogeochemistry~~ biology model as CanESM1 (Christian et al., 2010) and CanESM2 (Arora et al., 2011), the Canadian Model of Ocean Carbon (CMOC; Zahariev et al., 2008)~~, adapted for use within NEMO~~. An additional model was developed for CMIP6, called the Canadian Ocean Ecosystem model (CanOE). The biological components of CanOE are of substantially greater complexity than CMOC, including multiple food chains, flexible phytoplankton elemental ratios, and a prognostic iron (Fe) cycle. ~~Carbon chemistry, gas exchange and solubility of carbon dioxide (CO_2) and oxygen are identical between the two and follow the protocols specified by the Ocean Model Intercomparison Project - Biogeochemistry (OMIP-BGC) (Orr et al., 2017).~~ The two coupled models are known as CanESM5 and CanESM5-CanOE, respectively. ~~There are no feedbacks between biology and the physical ocean model, so the physical climate of CanESM5 and CanESM5-CanOE is identical in experiments with prescribed atmospheric CO_2 concentration.~~

70

The reasons for developing both models are, firstly, to evaluate the effect of changes in ocean

circulation between CanESM2 and CanESM5 on ocean biogeochemistry by running the new

climate model with the same ocean biogeochemistry, and secondly because CanOE is

substantially more expensive computationally (as it has 19 tracers vs 7 the total computation cost

is 2-3 times greater). Having CMOC as an option allowed us to run many Most CMIP6

experiments were run with CanESM5 only, as ocean biogeochemistry is not central to their

purpose. Many aAdditional tracers requested by the Ocean Model Intercomparison Project -

Biogeochemistry (OMIP-BGC) including abiotic and natural dissolved inorganic carbon (DIC),

DI^{14}C, CFCs and SF$_6$ (see Orr et al., 2017) were run only in CanESM5, since adding these

tracers on top of the larger set of biological components in CanOE would have been

prohibitively expensive. The CMIP6 experiments published for CanESM5-CanOE are listed in

Supplementary Table S1.

83

CMOC is a nutrient-phytoplankton-zooplankton-detritus (NPZD) model with highly

parameterized representations of phytoplankton Fe limitation, dinitrogen (N$_2$) fixation and

denitrification, and calcification and calcite dissolution (Zahariev et al., 2008; Supplementary

Figure S1). In CanESM1 and CanESM2, CMOC did not include oxygen;. In CanESM5, CMOC

now includes oxygen as a purely 'downstream' tracer that does not affect other biogeochemical

processes. , whereas iIn CanESM5-CanOE, denitrification is prognostic and dependent on the

concentration of oxygen. Among the less satisfactory aspects of CMOC biogeochemistry are,

firstly, that Fe limitation is specified as a static 'mask' that does not change with climate (being it

is calculated from the present-day climatological distribution of nitrate, based on the assumption

93 that regions without iron limitation will have complete drawdown of surface nitrate at some point

94 in the year), and secondly, that denitrification is parameterized so that nitrogen (N) is conserved

95 within each vertical column, i.e., collocated with N_2 fixation in tropical and subtropical open-

96 ocean regions (Zahariev et al., 2008; Riche and Christian, 2018). This latter simplification

97 produced excessive accumulations of nitrate in Eastern Boundary Current (EBC) regions where

98 most denitrification ~~actually~~ occurs. CMOC also has a tendency to produce rather stark extremes

99 of high and low primary and export production (Zahariev et al., 2008), a well-known problem of

100 NPZD models (Armstrong, 1994; Friedrichs et al., 2007). Our intent in developing CanOE was

101 to alleviate, or at least reduce, these biases, by including multiple food chains, a prognostic Fe

102 cycle, and prognostic denitrification. Dinitrogen fixation is still parameterized, but the CanOE

103 parameterization includes Fe (but not P) limitation, whereas in CMOC N_2 fixation tends to grow

104 without bound in a warming ocean as there is no P or Fe limitation (Riche and Christian, 2018).

105 ~~Calcification is represented by a prognostic detrital calcite pool with its own sinking rate (distinct~~

106 ~~from that of organic detritus), and calcite burial depends on the saturation state. In CMOC~~

107 ~~calcification is parameterized by a temperature dependent "rain ratio" and 100% burial of calcite~~

108 ~~that reaches the seafloor is assumed.~~

109

110 In this paper we present a detailed model description for CanOE and an evaluation of both

111 CanESM5 and CanESM5-CanOE relative to observational data products and other available

112 models. CMOC has been well described previously (Zahariev et al., 2008) and the details are not

113 reiterated here. In some cases, CanESM2 results are also shown to illustrate which differences in

114 the model solutions arise largely from the evolution of the physical ocean model, and which are

115 specifically associated with different representations of biogeochemistry. An overall evaluation

116 of the CanESM5 physical ocean model is given in Swart et al. (2019a). Here, we focus on

117 biogeochemical variables, and have evaluated model performance in three main areas: (1) the

118 distribution of major tracers like oxygen, DIC and alkalinity, and the resulting saturation state for

119 $CaCO_3$ minerals, (2) the iron cycle and its interaction with the nitrogen cycle, and (3) plankton

120 community structure and the concentration and export of particulates. We first address the major

121 chemical species that are common to both models (and almost all other Earth System Models) to

122 determine whether a more complex biology model measurably improves skill, and whether the

123 updated circulation model improves skill relative to CanESM2. Then we examine the areas

124 where our two models differ: the presence of a prognostic iron cycle and multiple food chains in

125 CanOE. More specifically, does CanESM5-CanOE reproduce the geographic distribution of

126 High-Nutrient, Low-Chlorophyll (HNLC) regions? Does the large phytoplankton / large

127 zooplankton food chain become dominant under nutrient-rich conditions, and how does having

128 multiple detrital size classes affect particle flux and remineralization length scale? Following this

129 model evaluation, we present historical trends in ocean anthropogenic CO_2 uptake, export

130 production, and total volume of low-oxygen waters over the historical (1850-2014) experiment.

131 Possible future changes under Shared Socioeconomic Pathway experiments will be addressed in

132 subsequent publications.

133

134 **2. Model Description**

135

136 CanESM5 (Swart et al., 2019a) is an updated version of CanESM2 (Arora et al., 2011), with an

137 entirely new ocean. The atmosphere model has the same T63 horizontal resolution, and contains

138 some important improvements in atmospheric physics (Swart et al., 2019a). The land surface

139 (Canadian Land Surface Scheme) and terrestrial carbon cycle (Canadian Terrestrial Ecosystem

140 Model) models are substantially the same as in CanESM2 with minor modifications as described

141 by Arora et al. (2020). The CanESM5 ocean is based on the NEMO modelling system version

142 3.4, with a horizontal resolution of 1°, telescoping to 1/3° in the tropics, and 45 vertical levels

143 ranging in thickness from ~6 m near the surface to ~250 m in the deep ocean (Swart et al.,

144 2019a). All physical climate model components are the same in CanESM5 and CanESM5-

145 CanOE. There are no feedbacks between biology and the physical ocean model, so the physical

146 climate of CanESM5 and CanESM5-CanOE is identical in experiments with prescribed

147 atmospheric CO_2 concentration.

148

149 The NEMO system is a publicly available archive of codes based on the OPA (Océan

150 PArallelisé) ocean model (Madec and Imbard, 1996; Guilyardi and Madec, 1997) and the

151 Tracers in Ocean Paradigm (TOP) module for tracer advection and mixing. It comes with two

152 options for biogeochemistry: PISCES (Pelagic Interactions Scheme for Carbon and Ecosystem

153 Studies) and LOBSTER (LODyC Ocean Biogeochemical System for Ecosystem and Resources).

154 Our ocean biogeochemistry modules CanOE and the NEMO implementation of CMOC are built

155 around the basic code structure of PISCES within the Tracers in Ocean Paradigm (TOP) module,

156 using NEMO v3.4.1, but have also been implemented in NEMO 3.6 for regional downscaling

157 applications (Holdsworth et al., 2021).

158

159 The biology, carbon chemistry, gas exchange and light attenuation components have all been

160 modified to various degrees. In a few cases PISCES parameterizations, or slightly modified

161 versions thereof, were adopted. CanOE uses PISCES three-band light attenuation while NEMO-

162 ~~CMOC uses broadband attenuation of photosynthetically active radiation (PAR) for consistency~~

163 ~~with the published version of CMOC.~~ Carbon chemistry ~~was modified to be consistent with~~ is

164 based on the Best Practices Guide (Dickson et al., 2007) and the OMIP-BGC data request (Orr et

165 al., 2017) and are identical in CanESM5 and CanESM5-CanOE. All calculations are done on the

166 total scale and the recommended formulae for the equilibrium constants are employed. ~~The~~

167 ~~PISCES conventions for convergence of c~~The carbon chemistry solver ~~calculations were~~

168 ~~retained, the greater~~ was run for a fixed number of iterations (ten in the surface layer, and five in

169 the subsurface layers in CanESM5-CanOE)~~offering greater accuracy in calculating pCO_2 and gas~~

170 ~~exchange. In subsurface layers (where the only function of the carbon chemistry is to calculate~~

171 ~~burial of calcite in the sediments) the number of iterations is fixed at five~~. CanESM5 ~~uses the~~

172 ~~same carbon chemistry but~~ does not solve the carbon chemistry equations in the subsurface

173 layers. OMIP-BGC formulations for CO_2 and O_2 solubility and gas exchange are employed. It is

174 important to note here that the carbon chemistry and gas exchange formulations used in

175 CanESM2 (and other CMIP5 models) are slightly different than those used in CMIP6. However,

176 this difference is of little functional significance, i.e., it will have a negligible impact on the

177 distribution of $[CO3^{--}]$ compared to the differences in DIC and alkalinity distribution. The

178 initialization fields for nitrate, DIC and alkalinity were also different in CanESM2. This will

179 affect the total ocean inventory of DIC but not the spatial distribution if the model is well

180 equilibrated.

181

182 The CanOE biology model is a substantially new model based on the cellular regulation model

183 of Geider et al. (1998). There are two phytoplankton ~~functional groups~~size classes, and each

184 group has four state variables: C, N, Fe and chlorophyll. Photosynthesis is decoupled from cell

185   production and photosynthetic rate is a function of the cell's internal N and Fe quotas. Each

186   functional group has a specified minimum and maximum N quota and Fe quota, and nutrient

187   uptake ceases when the maximal cell quota is reached. Chlorophyll synthesis is a function of N

188   uptake and increases at low irradiance. There are also two size classes each of zooplankton and

189   detritus. Small zooplankton graze on small phytoplankton, while large zooplankton graze on both

190   large phytoplankton and small zooplankton. Small detritus sinks at 2 m d^{-1} and large detritus at

191   30 m d^{-1} (in CanESM5 there is a single detrital pool with a sinking rate of 8 m d^{-1}). Model

192   parameters and their values are listed in Table 1. A schematic of the model is shown in Figure 1.

193

194   **2.1 Photosynthesis and Phytoplankton Growth**

195

196   For simplicity and clarity, the equations are shown here for a single phytoplankton species, and

197   do not differ structurally for small and large phytoplankton. Some parameter values differ for the

198   two phytoplankton groups; all parameter values are listed in Table 1.

199

200   Temperature dependence of photosynthetic activity is expressed by the Arrhenius equation

201    $T_f = \exp(-\frac{E_{ap}}{R}(\frac{1}{T} - \frac{1}{T_{\text{ref}}}))$ (1)

202   where E_{ap} is an enzyme activation energy that corresponds approximately to that of RuBisCo (cf.

203   Raven and Geider 1988), R is the gas constant (8.314 J mol^{-1} K^{-1}), and temperature T and

204   reference temperature T_{ref} are in Kelvin. Maximal rates of nutrient (either N or Fe, but

205   generically referred to here with the superscript X) uptake are given by

206    $V_{max}^X = V_{ref}^X T_f (\frac{Q_{max}^X - Q^X}{Q_{max}^X - Q_{min}^X})^{0.05}$ (2)

207   where V_{max}^X is the maximal uptake rate in mg of nutrient X per mg of cell C, X can represent N

208   or Fe, Q is the nutrient cell quota and Q_{min} and Q_{max} its minimum and maximum values, and V_{ref}^X

209   is a (specified) basal rate at $T=T_{ref}$ and $Q=Q_{min}$. These maximum rates are then reduced according

210   to the ambient nutrient concentration, i.e.

211      $V^N = V_{max}^N (L_{NH4} + (1 - L_{NH4})L_{NO3})$ (3a)

212   where $L_{\text{NH4}} = \frac{N_a}{K_{\text{NaX}}+N_a}$ and $L_{\text{NO3}} = \frac{N_i}{K_{\text{NiX}}+N_i}$, with N_i and N_a indicating nitrate and ammonium

213   respectively, and

214      $V^{Fe} = V_{max}^{Fe} \left(\frac{\text{Fe}}{K_{\text{FeX}}+\text{Fe}} \right)$ (3b)

215   where X indicates large or small phytoplankton (Table 1). The maximal carbon-based growth

216   rate is given by

217      $P_{max}^C = P_{ref}^C T_f \min\{\frac{Q^N-Q_{min}^N}{Q_{max}^N-Q_{min}^N} \cdot \frac{Q^{Fe}-Q_{min}^{Fe}}{Q_{max}^{Fe}-Q_{min}^{Fe}}\}$ (4)

218   where P_{ref}^C is the rate at the reference temperature T_{ref} under nutrient-replete conditions

219   ($Q=Q_{max}$). The light-limited growth rate is then given by

220      $P_{phot}^C = P_{max}^C \left(1 - e^{-\alpha_{chl}E\theta_C / P_{max}^C} \right)$ (5)

221   where E is irradiance and θ_C is the chlorophyll-to-carbon ratio. The rate of chlorophyll synthesis

222   is

223      $\rho_{chl} = \theta_{max}^N \frac{P_{phot}^C}{E\alpha_{chl}\theta}$ (6)

224   These rates are then used to define a set of state equations for phytoplankton carbon (C_p),

225   nitrogen (N_p), iron (Fe_p), and chlorophyll (M).

226      $\frac{dC_p}{dt} = (P_{phot}^C - \zeta V_N)C_p - (G + C_{XS}) - m_1 C_p - m_2 C_p^2 - k_{XU} C_{INTR}$ (7)

227   where ζ is the respiratory cost of biosynthesis, G is the grazing rate (equation 12), C_{XS} is the

228   excess (above the ratio in grazer biomass) carbon in grazing losses (see below equation 16a), m_1

229   and m_2 are coefficients for linear and quadratic nonspecific mortality terms, C_{INTR} is the

230   concentration of intracellular carbohydrate carbon in excess of biosynthetic requirements, and

231   k_{XU} is a rate coefficient for its exudation to the environment. The nonspecific mortality terms are

232   set to 0 below 0.01 mmol C m^{-3}, to prevent biomass from being driven to excessively low levels

233   in the high latitudes in winter; linear mortality terms can result in biomass declining to levels

234   from which recovery would take much longer than the brief Arctic summer (Hayashida, 2018).

235   The full equation for phytoplankton N, Fe and chlorophyll are

236
$$\frac{dN_p}{dt} = \frac{V^N}{Q_N} - (G + m_1 C_p + m_2 C_p^2) R_{NC} - N_{XS} \tag{8}$$

237
$$\frac{dFe_p}{dt} = \frac{V^{Fe}}{Q_{Fe}} - (G + m_1 C_P + m_2 C_p^2) R_{FeC} - Fe_{XS} \tag{9}$$

238
$$\frac{dM}{dt} = \frac{\rho_{chl} V^N}{\theta_C} M - (G + m_1 C_p + m_2 C_p^2)\theta_C - k_{dgr} M \tag{10}$$

239   where k_{dgr} is a rate coefficient for nonspecific losses of chlorophyll e.g., by photooxidation, in

240   addition to losses to grazing and other processes that also affect C_p, N_p, and Fe_p. N_{XS} and Fe_{XS} are

241   remineralization of "excess" (relative to grazer or detritus ratios) N or Fe and are defined below

242   (equation 16).

243

244   **2.2 Grazing and Food Web Interactions**

245

246   Grazing rate depends on the phytoplankton carbon concentration, which most closely represents

247   the food concentration available to the grazer (Elser and Urabe 1999; Loladze et al. 2000).

248   Zooplankton biomass is also in carbon units. State equations for small and large zooplankton are

249    $\frac{dZ_s}{dt} = \lambda G_s - (R + G_Z + m_{1s}Z_s + m_{2s}Z_s^2)$ (11a)

250    $\frac{dZ_L}{dt} = \lambda G_L - (R + m_{1L}Z_L + m_{2L}Z_L^2)$ (11b)

251   where

252    $G_s = G_{so}(1 - e^{-a_s C_{ps}})Z_s$ (12a)

253    $G_L = G_{L0}(1 - e^{-a_l(C_{pl}+Z_s)})Z_L$ (12b)

254   for small and large zooplankton respectively, G_Z is grazing of small zooplankton by large

255   zooplankton, R is respiration, and m_1 and m_2 are nongrazing mortality rates. Large zooplankton

256   grazing is divided into grazing on large phytoplankton and small zooplankton in proportion to

257   the relative abundance of each

258    $G_P = G_L \frac{P_l}{P_l+Z_s}$ (13a)

259    $G_Z = G_{+L} \frac{Z_s}{P_l+Z_s}$ (13b)

260   Zooplankton biomass loss to respiration is given by

261    $R = max\{r_z T_f Z - C_{XS}. 0\}$ (14)

262   and uses the same activation energy as photosynthesis. Respiration (R) is assumed to consume

263   only carbon and not result in catabolism of existing biomass when "excess" carbon is available

264   in the prey. In addition, conservation of mass must be maintained by recycling to the dissolved

265   pool grazer consumption of elements in excess of biosynthetic requirements when grazer and

266   prey elemental ratios differ. In the case where the nutrient quota (relative to carbon) exceeds the

267   grazer fixed ratio, the excess nutrient is remineralized to the dissolved inorganic pool. In the case

268   where the nutrient quota is less than the grazer ratio, the grazer intake is reduced to what can be

269   supported by the least abundant nutrient (relative to the grazer biomass ratio) and excess carbon

270   is remineralized. For the case of two nutrients (in this case N and Fe) it is necessary to define

271      $$G' = G\min\left\{\frac{N_P}{C_P}R_{\text{CN}}, \frac{\text{Fe}_P}{C_P}R_{\text{CFe}}, 1\right\}$$ (15)

272   where G is equal to G_S (equation 12a) for small zooplankton and G_P (equation 13a) for large

273   zooplankton, and R_{XY} indicates the fixed ratio of element X to element Y in grazer biomass. The

274   'excess' carbon available for respiration is

275      $$C_{\text{XS}} = G'\left\{\frac{C_P}{N_P}R_{\text{NC}} - 1, \frac{C_P}{\text{Fe}_P}R_{\text{FeC}} - 1, 0\right\}$$ (16a)

276   and the excess nutrients remineralized to their inorganic pools are

277      $$N_{\text{XS}} = G'\max\left\{\frac{N_P}{C_P} - R_{\text{NC}}.0\right\}\varepsilon + G'\max\left\{R_{\text{NC}}\left(\frac{N_P}{\text{Fe}_P}R_{\text{FeN}} - 1\right).0\right\}(1 - \varepsilon)$$ (16b)

278      $$\text{Fe}_{\text{XS}} = G'\max\left\{\frac{\text{Fe}_P}{C_P} - R_{\text{FeC}}.0\right\}\varepsilon + G'\max\left\{R_{\text{FeC}}(\frac{\text{Fe}_P}{N_P}R_{\text{NFe}} - 1).0\right\}(1 - \varepsilon)$$ (16c)

279   where

280      $$\varepsilon = \frac{\max\{C_{\text{xs}}, 0\}}{C_{\text{xs}} + \Delta}$$

281   is a switch to prevent double-counting in cases where one of the terms is redundant (the excess

282   relative to the least abundant element is included in the other term), but would otherwise be

283   nonzero (Δ is a constant equal to 10^{-15}, to prevent divide-by-zero). For three elements, there are

284      $3! = 6$ possible cases: for N greater or less than $C_P R_{\text{NC}}$, Fe may be either in excess relative to

285   both C and N, deficient relative to both, or in excess relative to one but not the other (Table 2).

286

2.3 Organic and Inorganic Pools

There are two pools of detritus with different sinking rates but the same fixed elemental ratios. Detrital C/N/Fe ratios are the same as zooplankton, so zooplankton mortality or grazing of small zooplankton by large zooplankton produce no 'excess'. Phytoplankton mortality, and defecation by zooplankton grazing on phytoplankton, produces excess nutrient or excess C that needs to be recycled into the inorganic pool in a similar fashion as outlined above for the assimilated fraction of grazing on phytoplankton.

The conservation equations for detrital C are

$$\frac{dD_s}{dt} = m_1(C_{ps} + Z_s) + m_2(C_{ps}^2 + Z_S^2) - r_1 D_s T_g - w_s \frac{dD_s}{dz} \tag{17a}$$

$$\frac{dD_l}{dt} = m_1(C_{pl} + Z_L) + m_2(C_{pl}^2 + Z_L^2) - r_2 D_l T_g - w_l \frac{dD_l}{dz} \tag{17b}$$

where T_g is an Arrhenius function for temperature dependence of remineralization and w is the sinking speed. The conservation equations for inorganic C, N, and Fe are

$$\frac{dC_i}{dt} = (\zeta V^N - P_{phot}^C)C_p + R + C_{XS} + (r_1 D_s + r_2 D_l)T_g \tag{18a}$$

$$\frac{dN_i}{dt} = -\frac{V^N}{Q^N} N_p \left(\frac{L_{NO3}}{L_{NO3} + L_{NH4}}\right) + N_{ox} - N_{dentr}(1 - A_f) \tag{18b}$$

$$\frac{dN_a}{dt} = -\frac{V^N}{Q^N} N_p \left(\frac{L_{NH4}}{L_{NO3} + L_{NH4}}\right) + \frac{R}{R_{CN}} + N_{XS} + (r_1 D_s + r_2 D_l)R_{NC}T_g - N_{ox} + N_{dnf} - N_{dentr}A_f \tag{18c}$$

$$\frac{dFe}{dt} = \frac{V^{Fe}}{Q^{Fe}} Fe_p + \frac{R}{R_{CFe}} + Fe_{XS} + (r_1 D_s + r_2 D_l)R_{FeC}T_g \tag{18d}$$

where N_{ox} is microbial oxidation of ammonium to nitrate (nitrification), N_{dnf} and N_{dentr} are sources and sinks associated with dinitrogen fixation and denitrification, and A_f is the ammonium fraction of denitrification losses, associated with anaerobic ammonium oxidation ("anammox").

307   The oxygen equation is essentially the inverse of equation 18a, with additional terms for

308   oxidation and reduction of N, i.e.,

309    $$\frac{dO_2}{dt} = -\frac{dC_i}{dt} + 2\frac{V^N}{Q^N}N_p\left(\frac{L_{NO3}}{L_{NO3}+L_{NH4}}\right) - 2N_{ox}$$ (19)

310   Nitrification is given by

311    $$N_{ox} = k_{NH4ox}N_a\frac{K_E}{K_E+E(z)}$$ (20)

312   where $E(z)$ is the layer mean irradiance at depth z. Dinitrogen fixation is parameterized as an

313   external input of ammonium dependent on light, temperature and Fe availability, and inhibited

314   by high ambient concentrations of inorganic N,

315    $$N_{dnf} = k_{dnf}T_{dnf}(1 - e^{-aE})\left(\frac{Fe}{K_{Fe}+Fe}\right)\left(\frac{K_{NO3}}{K_{NO3}+N_i+N_a}\right)$$ (21)

316   where $T_{dnf}=\max(0, 1.962(T_f - 0.773))$, i.e., a linear multiple of equation (1) that is 0 at T<20°C

317   and unity at T=30°C.

318

319   Denitrification is parameterized as a fraction of total remineralization that increases as a linear

320   function of oxygen concentration for concentrations less than a threshold concentration O_{mxd}

321    $$N_{frxn} = 1 - \frac{\min(O_2,O_{mxd})}{O_{mxd}}$$ (22)

322   Remineralization is then divided among oxygen (1-N_{frxn}), nitrate (0.875N_{frxn}), and ammonium

323   (0.125N_{frxn}) assuming an average anammox contribution of 25% (Babbin et al., 2014). We use

324   this average ratio of anammox to classical denitrification to partition fixed N losses between

325   NO_3^- and NH_4^+; the DIC sink and organic matter source associated with anammox are small and

326   are neglected here.

327

2.4 Calcification, Calcite Dissolution, and Alkalinity

In CanOE, calcification is represented by a prognostic detrital calcite pool with its own sinking rate (distinct from that of organic detritus), and calcite burial or dissolution in the sediments depends on the saturation state (100% burial when $\Omega_C \geq 1$, 100% dissolution when $\Omega_C < 1$). Calcification is represented by a detrital calcium carbonate ($CaCO_3$) state variable, but no explicit calcifier groups. Detrital $CaCO_3$ sinks in the same fashion as detrital particulate organic carbon (POC), with a sinking rate independent of those for large and small organic detritus. Calcite production is represented as a fixed fraction of detritus production from small phytoplankton and small zooplankton mortality:

$$\frac{dCa}{dt} = m_1\left(C_{ps} + Z_s\right)P_{Ca} + m_2\left(C_{ps}^2 + Z_s^2\right)P_{Ca} - k_{Ca}Ca - w_{Ca}\frac{dCa}{dz} \tag{23}$$

Calcite dissolution occurs throughout the water column as a first order process (i.e., no dependence on temperature or saturation state). Approximately 80% of calcite produced is exported from the euphotic zone. Burial in the sediments is represented as a simple 'on/off' switch dependent on the calcite saturation state (zero when $\Omega_C < 1$ and 1 when $\Omega_C \geq 1$). In CanESM5, calcification is parameterized by a temperature dependent "rain ratio" (Zahariev et al., 2008) and 100% burial of calcite that reaches the seafloor is assumed. Calcite burial in both models is balanced by an equivalent source of DIC and alkalinity at the ocean surface (in the same vertical column) as a crude parameterization of fluvial sources.

For each mole of calcite production two moles of alkalinity equivalent are lost from the dissolved phase; the reverse occurs during calcite dissolution. There are additional sources and sinks for alkalinity associated with phytoplankton nutrient (NH_4^+, NO_3^-) uptake, organic matter

351 remineralization, nitrification, denitrification and dinitrogen fixation (Wolf-Gladrow et al., 2007,

352 see Supplementary Table S2). The anammox reaction does not in itself contribute to alkalinity

353 (Jetten at al., 2001), but there is a sink associated with ammonium oxidation to nitrite (the model

354 does not distinguish between nitrite and nitrate). Autotrophic production of organic matter by

355 anammox bacteria is a net source of alkalinity (Strous et al., 1998) but this source is extremely

356 small (~0.03 mol/molN) and is neglected here. Globally, the sources and sinks of alkalinity from

357 the N cycle offset each other such that there is no net gain or loss as long as the global fixed N

358 pool is conserved (see below Sect. 2.5). If dinitrogen fixation and denitrification are allowed to

359 vary freely, there will generally be a net gain or loss of fixed N and, therefore, of alkalinity.

360

361 **2.5 External Nutrient Sources and Sinks**

362

363 External sources and sinks consist of river inputs, aeolian deposition, biological N_2 fixation,

364 denitrification, mobilization of Fe from reducing sediments, loss of Fe to scavenging, and burial

365 of calcium carbonate in the sediments. Aeolian deposition of Fe is calculated from a climatology

366 of mineral dust deposition generated from offline (atmosphere-only) simulations with CanAM4

367 (von Salzen et al., 2013), with an Fe mass fraction of 5% and a fractional solubility of 1.4% in

368 the surface layer. Subsurface dissolution is parameterized based on PISCESv2 (Aumont et al.,

369 2015); the total dissolution is 6.35%, with 22% of soluble Fe input into the first vertical layer

370 (see Supplementary material). Iron from reducing sediments is also based on PISCES, with a

371 constant areal flux of 1000 nmol m^{-2} d^{-1} in the first model level, declining exponentially with

372 increasing seafloor depth (i.e., assuming that shelf sediments are the strongest source and the

373 sediments become progressively more oxygenated with increasing seafloor depth) with an e-

374 folding length scale of about 200 m. Scavenging of dissolved iron is first-order with a high rate

375 (2.5 d^{-1}) for concentrations in excess of 0.6 nM (Johnson et al., 1997). For concentrations below

376 this threshold, the rate is much lower (0.001 d^{-1}) and is weighted by the concentration of organic

377 detritus (Christian et al., 2002b), i.e.,

378 $$\frac{dFe}{dt} = -FeS_{Fe1}min\{(D_S + D_L)P_{Fe}, 1\} \tag{24}$$

379 where Fe is the dissolved iron concentration, D_S and D_L are the small and large detritus

380 concentrations, S_{Fe1} is the first-order scavenging rate in surface waters with abundant

381 particulates, and P_{Fe} is an empirical parameter to determine the dependence on particle

382 concentration (Table 1). The basis for this parameterization is that the rate of scavenging must

383 depend not only on the concentration of iron but on the concentration of particles available for it

384 to precipitate onto, and assumes that POC is strongly positively correlated with total particulate

385 matter. Scavenging is treated as irreversible, i.e., scavenged Fe is not tracked and does not

386 reenter the dissolved phase.

387

388 Unlike in CMOC, N_2 fixation and denitrification vary independently in CanOE, so the global

389 total N pool can change. Conservation is imposed by adjusting the global total N pool according

390 to the difference between the gain from N_2 fixation and the loss to denitrification. A slight

391 adjustment is applied to the nitrate concentration at every grid point, while preserving the overall

392 spatial structure of the nitrate field. Adjustments are multiplicative rather than additive to avoid

393 producing negative concentrations. This adjustment does not maintain (to machine precision) a

394 constant global N inventory but is intended to minimize long term drift, keeping it much smaller

395 than the free surface error (see below). This adjustment is applied every 10 days and has a

396 magnitude of approximately 7×10^{-8} of the total N.

397

398   When the total fixed N adjustment is applied, ~~One~~ one mole of alkalinity is removed per mole of

399   N added or removed, to account for the ~~since there are~~ alkalinity sources ~~of 1 mol/molN~~

400   associated with ~~both~~ N_2 fixation (creation of new NH_4^+) and denitrification (removal of NO_3^-)

401   (Wolf-Gladrow et al., 2007, see Supplementary Table S2). As there is~~, offset by~~ a 2 mol/molN

402   sink associated with nitrification, this formulation is globally conservative. As noted above, in

403   CanOE $CaCO_3$ can dissolve or be buried in the sediments depending on the calcite saturation

404   state. DIC and alkalinity lost to burial are reintroduced at the ocean surface, at the same grid

405   point as burial occurs, providing a crude parameterization of river inputs so that global

406   conservation is maintained (fresh water runoff contains no DIC or alkalinity). However, the OPA

407   free surface formulation is inherently imperfect with regard to tracer conservation. Drift in total

408   ocean alkalinity and nitrogen over time is on the order of 0.01% and 0.03% per thousand years,

409   respectively (losses due to the free surface are generally larger for tracers with less homogeneous

410   distributions).

411

412   **2.6 Ancillary data**

413

414   For first-order model validation we have relied largely on global gridded data products rather

415   than individual profile data. Global gridded data from World Ocean Atlas ~~2013~~2018

416   (WOA~~2013~~2018) (Locarnini et al., ~~2013~~2018; Zweng et al., ~~2013~~2018; Garcia et al.,

417   ~~2014a~~2018a; ~~2014b~~2018b) were used for temperature, salinity, and oxygen and nitrate

418   concentration. DIC and alkalinity were taken from the GLODAPv2.2016b gridded data product

419   (Key et al., 2015; Lauvset et al., 2016). Offline carbon chemistry calculations were done

420   following the Best Practices Guide (Dickson et al., 2007) and the OMIP-BGC protocols (Orr et

421   al., 2017), ~~which~~ and are identical to those used in the models except that constant reference

422   concentrations were used for phosphate (1 μM) and silicate (10 μM).

423

424   There is no global gridded data product for Fe, but we have made use of the GEOTRACES

425   Intermediate Data Product 2017 (Schlitzer et al., 2018), and the data compilations from MBARI

426   (Johnson et al., 1997; 2003) and PICES Working Group 22 (Takeda et al., 2013). The latter two

427   are concentrated in the Pacific, while GEOTRACES is more global. The combined data sets

428   provide more than 10000 bottle samples from more than 1000 different locations (Supplementary

429   Figure ~~S4a~~S9a) (excluding some surface transect data that involve frequent sampling of closely

430   spaced locations along the ship track). More detail about model comparison to these data

431   compilations and the list of original references are given in the Supplementary information.

432

433   Satellite ocean colour estimates of surface chlorophyll were taken from the combined

434   SeaWiFS/MODIS climatology described by Tesdal et al. (2016). Climatological satellite POC

435   was downloaded from the NASA ocean colour web site and is based on the algorithm of

436   Stramski et al. (2008) using MODIS-Aqua data. This climatology differs slightly from the

437   chlorophyll one in terms of years included and sensors utilized, but as only climatological

438   concentrations are considered and each climatology covers ~15 years, these differences will have

439   negligible effect on the results presented. Satellite chlorophyll concentrations greater than 1 mg

440   m^{-3} were excluded as these are mostly associated with coastal regions not resolved by coarse-

441   resolution global ocean models.

442

443   CMIP6 model data were regridded by distance-weighted averaging using the Climate Data

444   Operators (https://code.mpimet.mpg.de/projects/cdo/) to a common grid (2x2°, 33 levels) to

445   facilitate ensemble averaging. The vertical levels used are those used in GLODAP and in earlier

446   (through 2009) versions of the World Ocean Atlas (e.g., Locarnini et al., 2010). following the

447   GLODAP levels) to facilitate ensemble averaging. For large scale tracer distributions, using a 1°

448   or 2° grid makes little difference (for example, the spatial pattern correlation between CanESM5

449   and observed oxygen concentration at specific depths differs by an average of 0.0011). The years

450   1986-2005 of the Historical experiment were averaged into climatologies or annual means, for

451   meaningful comparison with observation-based data products. The CMIP6 Historical experiment

452   runs from 1850-2014 with atmospheric CO_2 concentration (and other atmospheric forcings)

453   based on historical observed values. A single realization was used in each case (see Table S3); as

454   20 year averages are used, internal variability is assumed to have little effect (e.g., Arguez and

455   Vose, 2011, see Table S4). Where time series are shown, 5-year means are used.

456

457   Sampling among CMIP6 models was somewhat opportunistic and the exact suite of models

458   varies among the analyses presented. When we conducted a search for a particular data field, we

459   included in the search parameters all models that published that field, and repeated the search at

460   least once for models that were unavailable the first time the search was executed. In some cases,

461   model ensemble means excluded all but one model from a particular 'family' (e.g., there are

462   three different MPI-ESM models for which ocean biogeochemistry fields were published), as the

463   solutions were found to be similar and would bias the ensemble mean towards their particular

464   climate. The models used are ACCESS-ESM1-5, CESM2, CESM2-WACCM, CNRM-ESM2-1,

465   GFDL-CM4, GFDL-ESM4, IPSL-CM6A-LR, MIROC-ES2L, MPI-ESM-1-2-HAM, MPI-

466 ESM1-2-LR, MPI-ESM1-2-HR, MRI-ESM2-0, NorESM2-LM, NorESM2-MM, and UKESM1-

467 0-LL. More dDetails of which variables and realizations are used for which models areis given in

468 Supplementary Table S3.

469

470 **3. Results**

471

472 We first describe here the large-scale distribution of oxygen, DIC, alkalinity, and the saturation

473 state with respect to $CaCO_3$ that derives from these large-scale tracer distributions. Tracer

474 distributions result partly from ocean circulation and partly from biogeochemical processes. An

475 overall evaluation of the ocean circulation model is given in Swart et al. (2019a). Analyzing

476 CanESM5 and CanESM5-CanOE (with identical circulation) as well as CanESM2 where

477 possible (same biogeochemistry as CanESM5 but different circulation) allows us to separate the

478 effects of physical circulation and biogeochemistry on evolving model skill with respect to large-

479 scale tracer distributions. In subsequent sections we address the main areas where CanESM5 and

480 CanESM5-CanOE differ, such as the interaction of the iron and nitrogen cycles and plankton

481 community structure. Finally, we present some temporal trends over the course of the historical

482 experiment (1850-2014).

483

484 **3.1 Distribution of oxygen**

485

486 The spatial distribution of oxygen concentration ($[O_2]$) at selected intermediate depths (400, 900,

487 and 1400 m) is shown in Figure 2 for gridded data from WOA2018 and differences of

488 CanESM5, CanESM5-CanOE, a model ensemble mean (MEM) of CMIP6 models (excluding

489   CanESM5 and CanESM5-CanOE) from the observational data product., and gridded data from

490   WOA2013 is shown in Figure 2. The depths were chosen to span the depth range where low

491   oxygen concentrations exist; these low-oxygen environments are of substantial scientific and

492   societal interest and are sensitive to model formulation. The major features are consistent across

493   the models. Both CanESM models as well as the MEM All three cases show elevated oxygen

494   concentrations relative to observations, particularly in the North Pacific, the North Atlantic and

495   the Southern Ocean. In the Indian Ocean, both CanESM models show high oxygen

496   concentrations in the Arabian Sea and deeper layers of the Bay of Bengal relative to observations

497   and other CMIP6 modelsthe MEM; these biases are somewhat smaller in CanESM5-CanOE than

498   in CanESM5 (Figure 2).

499

500   The ocean's oxygen minimum zones (OMZs) are mostly located in the eastern Pacific Ocean, the

501   northern North Pacific, and the northern Indian Ocean; the spatial pattern changes with

502   increasing depth (Figure 2), but the OMZs are mostly located between 200 and 2000 m depth.

503   Biases in the eastern boundary currentEBC regions are depth and model specific. CanESM5

504   shows particularly strong oxygen depletion at 1400 m in the eastern tropical Pacific. In the

505   southeastern Atlantic, models tend to be biased low at the shallower depths, and show somewhat

506   more variation at greater depths (Figure 2). Overall, [O_2] biases tend to be positive over large

507   areas of ocean with the exception of some eastern boundary currentEBC regions, implying that

508   models exaggerate the extent to which remineralization is concentrated in these regions. An

509   alternate version of Figure 2 that shows the modelled concentrations model errors relative to the

510   observational data product is given in Supplementary Figure S3S2.

511

512 The zonal mean oxygen concentration, saturation concentration, and apparent oxygen utilization

513 (AOU) are shown in Figure 3 for the same four cases. Again, the models generally show a

514 positive bias in [O_2], particularly in high-latitude deep waters. The major ocean circulation

515 features are reproduced fairly well in all cases (e.g., weaker ventilation of low-latitude

516 subsurface waters, greater vertical extent of well-ventilated surface waters in the subtropics). The

517 saturation concentration (a function of temperature and salinity) generally shows relatively little

518 bias, implying that the bias in [O_2] arises mainly from remineralization and/or ventilation. AOU

519 is lower than observed over much of the subsurface ocean. Regional biases are quite consistent

520 across models, but are slightly greater in CanESM5 than in CanESM5-CanOE or the ensemble

521 meanMEM, except in the Arctic Ocean. Again, Supplementary Figure S23 includes a version of

522 this plot that shows the modelled differences from the observationsconcentration fields.

523

524 The skill of each model with respect to the distribution of O_2 at different depths is represented by

525 Taylor diagrams (Taylor, 2001) in Figure 4. These diagrams allow us to assess how well the

526 model reproduces the spatial distribution at a range of depths, because different physical and

527 biogeochemical processes determine the distribution in different depth ranges, in which. all All

528 of the CMIP6 models that were shown as an ensemble mean in Figures 2 and 3 are shown

529 individually. The large blue dots represent CanESM5, red CanESM5-CanOE, and grey the

530 ensemble mean of all CMIP6 models except CanESM5 and CanESM5-CanOEMEM; the smaller

531 grey dots represent the individual models. CanESM5-CanOE shows slightly higher pattern

532 correlation than CanESM5 at all depths. Both models compare favourably with the full suite of

533 CMIP6 models, with r>0.85 for CanESM5 and r>0.9 for CanESM5-CanOE at all depths

534 examined, and a normalized standard deviation within ±25% of unity.

535

536   The total volume of ocean with $[O_2]$ less than 6 mmol m^{-3} (the threshold for denitrification

537   (Devol, 2008)) and 60 mmol m^{-3} (a commonly used index of hypoxia) is shown in Figure 5. The

538   total volume is highly variable among models (note, however, that there are several clusters of

539   related models with quite similar totals). CanESM5 and CanESM5-CanOE have among the

540   lowest total volumes (i.e., the interior ocean is relatively well ventilated) and are among the

541   nearest to the observed total. For $[O_2]$ <60 mmol m^{-3} the bias is, nonetheless, quite large (i.e., the

542   observed volume is underestimated by almost 50% in both models). The volume of water with

543    $[O_2]$ below the denitrification threshold is overestimated in both CanESM5 and CanESM5-

544   CanOE; CanESM5-CanOE has a much smaller total that is closer to the observed value. The bias

545   in the spatial pattern of hypoxia (not shown) is generally similar to the bias in dissolved oxygen

546   distribution (Figure 2). The low-oxygen regions are generally more concentrated in the eastern

547   tropical Pacific in the models than in observations, and the low-oxygen region in the northwest

548   Pacific is not well reproduced in CanESM models.

549

550   **3.2 Distribution of DIC, alkalinity, and CaCO$_3$ saturation**

551

552   The spatial distribution of aragonite saturation state (Ω_A) at selected depths is shown in Figure 6.

553   (tThe first two depths are the same as in Figure 2, but a much greater depth is also included, as

554   the length scale for CaCO$_3$ dissolution is greater than for organic matter remineralization the

555   third is much deeper). In this case the observations are a combination of GLODAPv2 (Key et al.,

556   2015; Lauvset et al., 2016) for DIC and alkalinity, and WOA20132018 for temperature and

557   salinity. CanESM5 and CanESM5-CanOE generally compare well with other models and

558   observations. The low saturation bias in the eastern tropical Pacific is substantially reduced in

559   CanESM5-CanOE compared to CanESM5. On the other hand CanESM5 generally does better

560   than CanESM5-CanOE, or the ~~ensemble mean~~MEM, at reproducing the low saturation states in

561   the northwestern Pacific and the Bering Sea. Both CanESM models show a high saturation state

562   bias in the North Atlantic and the well-ventilated regions of the north Pacific subtropical gyre~~;~~.

563   ~~These~~ these biases are ~~reduced~~ slightly smaller in CanESM5-CanOE~~, probably due to the smaller~~

564   ~~average remineralization length scale for organic detritus~~. Maps of the calcite and aragonite

565   saturation horizon (Ω=1) depth are shown in Supplementary Figure S3; these generally confirm

566   the same biases noted in Figure 6.

567

568   Zonal mean distributions of aragonite saturation state (Ω_A), calcite saturation state (Ω_C), and

569   carbonate ion concentration ([CO_3^{--}]) and the differences of the models from the observations are

570   shown in Figure 7 (Supplementary Figure ~~S3~~ S2 includes versions of Figures 6 and 7 that

571   ~~explicitly~~ show the model~~led fields~~ ~~differences from the observations~~). The models generally

572   compare well with the observations in the representation of the latitude/depth distribution of high

573   and low saturation waters. CanESM5 has a high saturation bias in low-latitude surface waters

574   that is somewhat reduced in CanESM5-CanOE. Both CanESM5 models show a high saturation

575   bias in Northern Hemisphere intermediate (e.g., 200-1000 m) depth waters that is larger than in

576   the ~~ensemble mean~~MEM. This is primarily a result of low Ω in the North Atlantic Ocean (Figure

577   6).

578

579   Taylor diagrams for a range of depths are shown for DIC in Figure 8 and for Ω_A in Figure 9 (for

580   alkalinity, see Supplementary Figure S~~4~~2). As expected, the MEM generally compares

581 favourably with the individual models (e.g., Lambert and Boer, 2001). CanESM5 and

582 CanESM5-CanOE compare favourably with the full suite of CMIP6 models. CanESM5-CanOE

583 shows a gain in skill relative to CanESM5, and both show improvement relative to CanESM2. At

584 400 m, CanESM2 stands out as having extremely high variance, which is mostly due to

585 extremely high DIC concentrations occurring over a limited area in the eastern equatorial Pacific

586 (not shown). This bias is present in CanESM5 and in CMIP6 models generally (Figure 6) but

587 involves much lower concentrations spread over a larger area.

588

589 **3.3 N and Fe cycles**

590

591 An important difference between ~~CMOC~~ CanESM5 and CanESM5-CanOE is the inclusion of a

592 prognostic Fe cycle. The CMOC iron mask (Zahariev et al., 2008) was a pragmatic solution in

593 the face of resource limitations but is inherently compromised as it can not evolve with a

594 changing climate. ~~Other centres that introduced a prognostic Fe cycle between CMIP5 and~~

595 ~~CMIP6 include JAMSTEC (MIROC-ESM, MIROC-ES2L) and the UK Met Office (HadGEM2-~~

596 ~~ES, UKESM1-0-LL).~~ The first order test of a model with prognostic, interacting Fe and N cycles

597 is whether it can reproduce the distribution of ~~High-Nutrient, Low-Chlorophyll (~~HNLC~~)~~ regions

598 and the approximate surface macronutrient concentrations within these. CanESM5-CanOE

599 succeeded by this standard, although the surface nitrate concentrations are biased low in the

600 subarctic Pacific and equatorial Pacific and high in the Southern Ocean and in the global mean

601 (Figure 10).

602

603 The seasonal cycle of the zonal mean surface nitrate concentration for a selection of CMIP6

604 models is shown in Figure 11. CanESM5, CanESM5-CanOE, and CNRM-ESM2-1 reproduce the

605 equatorial enrichment and the low concentrations in the tropical-subtropical latitudes fairly well.

606 Some models either have very weak equatorial enrichment (MPI-ESM1-2-LR) or too high a

607 concentration in the off-equatorial regions (UKESM1-0-LL, NorESM2-LM). UKESM1-0-LL

608 has very high concentrations throughout the low-latitude Pacific, which biases the ensemble

609 mean (Figure 11a). Supplementary Figure S611b shows the same data as Figure 11a but for a

610 more limited latitude range to better illustrate model behaviour in the tropics. CanESM5,

611 CanESM5-CanOE, and CNRM-ESM2-1 reproduce the seasonal cycle of tropical upwelling (e.g.,

612 Philander and Chao, 1991), with highest concentrations in summer.

613

614 The surface distribution of dissolved iron (dFe) in various CMIP6 models is shown in Figure 12.

615 For Fe there is no observation-based global climatology with which to compare the model

616 solutions (some comparisons to available profile data are shown in Supplementary Figures

617 S4bS9b-gh). CanESM5-CanOE shows a similar overall spatial pattern to other models, and

618 generally falls in the middle of the spread, particularly regarding concentrations in the Southern

619 Ocean. Several models show extremely high concentrations in the tropical-subtropical North

620 Atlantic (Sahara outflow region). CanESM5-CanOE, along with CNRM-ESM2-1 and CESM2,

621 has much less elevated concentrations in this region, due to lower deposition or greater

622 scavenging or both. CanESM5-CanOE has its lowest concentration in the eastern subtropical

623 South Pacific, which is common to many models (Figure 12). The area of strong surface

624 depletion is generally more spatially restricted in CanESM5-CanOE than in other models, and

625 surface dFe concentrations are greater over large areas of the Pacific. Both the north-south and

626 east-west asymmetry of distribution in the Pacific is greater in CanESM5-CanOE than in most

627 other models, some of which show the South Pacific minimum extending westward across the

628 entire basin, and others into the Northern Hemisphere. Only in CESM2 is this minimum

629 similarly limited to the southeast Pacific.

630

631 The mean depth profiles of dFe are shown in Figure 13. Some models show more of a "nutrient-

632 type" (increasing with depth due to strong near-surface biological uptake and subsequent

633 remineralization) profile, some a more "scavenged-type" (maximal at the surface, declining with

634 depth) profile (cf. Li, 1991; Nozaki, 2001), and others a hybrid profile (increasing downward but

635 with a surface enrichment). CanESM5-CanOE is at the "nutrient-type" end of spectrum with a

636 generally monotonic increase with depth to a near-constant deep-water concentration of 0.6 nM

637 and a very slight near-surface enrichment (see also Supplementary Figures S9S4b,c). In

638 CanESM5-CanOE the scavenging model is very simple, with distinct regimes for concentrations

639 greater or less than 0.6 nM; scavenging rates are very high above this threshold which causes

640 deep-water concentrations to converge on this value. The generally nutrient-like profile suggest

641 that in CanOE the scavenging rate is quite low for concentrations below 0.6 nM. CanOE

642 considers particulate organic matter (POM) as an index of all particulate matter available for

643 scavenging onto, and model POM concentrations fall off rapidly below the euphotic zone (not

644 shown).

645

646 Mean surface nitrate and dFe concentrations for selected ocean regions are shown in Figure 14.

647 CanESM5-CanOE shows concentrations that are within the range of CMIP6 models, although in

648 some cases at the higher or lower end. Surface nitrate concentrations generally compare

649 favourably with the observation-based climatology, but are biased low in HNLC regions other

650 than the Southern Ocean. These biases are not necessarily a consequence of having too much or

651 too little iron. For example, in the Southern Ocean CanESM5-CanOE has among the highest

652 surface nitrate concentrations, but it also has some of the highest dFe concentrations, and the

653 high nitrate bias is present in CanESM5 as well. Comparisons with the limited GEOTRACES

654 data available suggest that near surface dFe concentrations in the Southern Ocean are biased high

655 rather than low in CanESM5-CanOE (not shown). One region where there does seem to be a

656 strong correlation between surface nitrate and dFe concentrations is the western subarctic

657 Pacific. All but two models (CNRM-ESM2-1, NorESM2-LM) fall along a spectrum from high

658 Fe / low nitrate to low Fe / high nitrate. CanESM5-CanOE falls near the high Fe / low nitrate end

659 of the range.

660

661 Surface nitrate concentrations along the Pacific equator during the upwelling season (June-

662 October) for CanESM5 and CanESM5-CanOE are shown in Figure 15. The range of other

663 CMIP6 models is not shown here because it is large and therefore adds little information (see

664 Figure 11 and Supplementary Figure S6). CanESM5-CanOE better represents the east-west

665 gradient, while CanESM5 has slightly higher concentrations in the core upwelling region. Both

666 models underestimate the highest concentrations around 100°W. ~~Some~~ Although some localized

667 maxima in this data product are due to undersampling, equatorial upwelling is strong at this

668 location (e.g., Lukas, 2001) and the spatial coherence of the data strongly suggests that this

669 maximum accurately reflects reality~~; however, examination of ancillary data sets such as satellite~~

670 ~~sea surface temperature suggests that the enrichment at 100°W accurately reflects ocean~~

671 ~~upwelling (not shown). Although~~ It should be noted that CanESM5 iron limitation is calculated

672 from a~~n earlier~~ version of the same data product~~,~~; however, the Fe mask is based on the

673 minimum nitrate concentration over the annual cycle, whereas the data shown here are for the

674 upwelling season. ~~In CanESM5-CanOE, the distribution of surface nitrate is an emergent~~

675 ~~property of the model, and the fidelity to the observed distribution is generally good.~~

676

677 **3.4 Plankton biomass, detritus, and particle flux**

678

679 The relative abundance of the four living plankton groups are shown in Figure 16 for a range of

680 ocean regions. Both CanESM models mostly compare favourably with observation-based

681 estimates of phytoplankton biomass, except in the tropics where CanESM5-CanOE has very high

682 biomass. Both CanESM models have low phytoplankton biomass in the North Atlantic. In the

683 North Pacific and the Southern Ocean, CanESM5-CanOE reproduces the observation-based

684 estimates well, and CanESM5 slightly less well. The general pattern is that large and small

685 phytoplankton have similar abundance, and are substantially more abundant than zooplankton.

686

687 Part of the rationale for multiple food chains is that they better represent the way that actual

688 plankton communities adapt to different physical ocean regimes and therefore are better able to

689 simulate distinct ocean regions with a single parameter set (e.g., Chisholm, 1992; Armstrong,

690 1994; Landry et al., 1997; Friedrichs et al., 2007). The expectation is that small phytoplankton

691 will be more temporally stable and large phytoplankton will fluctuate more strongly between

692 high and low abundances. The mean annual cycles of surface chlorophyll largely conform to

693 this~~ese expected~~ pattern, e.g., in the North Atlantic and the western subarctic Pacific large

694 phytoplankton are dominant in summer and much more variable over the seasons (Figure 17).

695 Compared to observations, CanESM5 models underestimate the amplitude of the seasonal cycle

696   in the North Atlantic and overestimate it in the North Pacific. CanESM5 shows a stronger and

697   earlier North Atlantic spring bloom compared to CanESM5-CanOE; the observations are in

698   between the two in terms of timing, and both models underestimate the amplitude (Figure 17). In

699   the tropics, the seasonal cycle is weak. CanESM5-CanOE in Tthe tropical Atlantic shows the

700   expected seasonal cycle but not the expected dominance of large phytoplankton in summer.

701   ~~These size-fractionation patterns are difficult to validate against observations.~~ CanESM5-CanOE

702   generally overestimates the total near surface chlorophyll in both the tropical Pacific and the

703   tropical Atlantic.

704

705   Zooplankton biomass (especially microzooplankton) is also somewhat difficult to test against

706   observations, but our model concentrations appear to be biased low. ~~White et al. (1995), for~~

707   ~~example, show a transect of vertically resolved mesozooplankton abundance along 140°W in the~~

708   ~~tropical Pacific, with biomass ranging from about 0.1-0.7 mmolC m³; CanESM5-CanOE~~

709   ~~concentrations in this region are much lower.~~ Stock et al. (2014) estimated depth-integrated

710   biomass of phytoplankton, mesozooplankton, and microzooplankton for a range of oceanic

711   locations in which intensive field campaigns have occurred (estimates of microzooplankton

712   biomass are relatively sparse). They found that in most locations phytoplankton and (combined)

713   zooplankton biomass are of comparable magnitude, whereas in CanESM5-CanOE zooplankton

714   biomass is consistently lower (Figure 16). The global integral biomass of mesozooplankton is

715   about an order of magnitude less than the 0.19 PgC estimated by Moriarty and O'Brien (2013).

716   The CanESM5 total of 0.14 Pg is relatively close to the Moriarty estimate but implicitly includes

717   microzooplankton.

718

719   Surface chlorophyll and POC for CanESM5-CanOE and for ocean colour observational data are

720   shown in Figure 18 (POC in the model is the sum of phytoplankton, microzooplankton, and

721   detrital carbon). The observations have a lower limit for POC that is not present in the model

722   (~17 mgC m^{-3}), which is unsurprising given the processes neglected in the model, i.e., in regions

723   of very low chlorophyll there is still substantial dissolved organic carbon, bacteria that consume

724   it, and microzooplankton that consume the bacteria and produce particulate detritus. The

725   observational data show a fairly linear relationship at low concentrations, but with a curvature

726   that implies a greater phytoplankton fraction in more eutrophic environments (cf. Chisholm,

727   1992). The model, by contrast, shows a fairly linear relationship over the whole range of

728   concentrations. In other words, the phytoplankton share of POC is higher and more constant in

729   the model than in the observations. The living biomass (phytoplankton + microzooplankton)

730   fraction of total POC in CanOE is generally in excess of 50% (not shown), which is implausible

731   for a real-world oceanic microbial community (e.g., Christian and Karl, 1994) but consistent

732   with the relatively low rates of export from the euphotic zone.

733

734   Export production for a range of CMIP6 models is shown in Figure 19a. CanESM5-CanOE is at

735   the low end of the range. Observations are not shown because the range of observational

736   estimates covers ~~almost~~ the entire range of ~~the plot~~model estimates (e.g., Siegel et al., 2016).

737   Note also that CanESM5 export is quite a bit lower than in CanESM2, which is relatively high

738   for CMIP5 models (not shown). The difference between CanESM2 and CanESM5 is attributable

739   primarily to different circulation, although the different initialization fields for nitrate might also

740   play a small role. The lower rate in CanESM5-CanOE is consistent with the above results

741   regarding plankton community structure (e.g., the concentration of detritus is generally low

742 compared to living biomass), as well as the lower sinking rate for small detritus. The latitudinal

743 distribution of export is shown in Figure 19b. CanESM5 shows very high export in the mid-

744 latitudes of the Southern Ocean, similar to CanESM2 (not shown). Both CanESM5 and

745 CanESM5-CanOEmodels show latitudinal patterns consistent with the range of other CMIP6

746 models. CanESM5 has slightly greater export in the equatorial zone; in both CanESM5 and

747 CanESM5-CanOE models the equatorial enrichment attenuates very rapidly with latitude and the

748 rates are low in the subtropics.

749

750 **3.5 Historical trends**

751

752 Cumulative ocean uptake of CO_2 is shown in Figure 20 for the historical experiment (1850-

753 2014). CanESM models are biased low relative to observation based estimates (~145 PgC, see

754 Friedlingstein et al., 2020) and the ensemble mean of other modelsMEM (144 PgC, Figure 20),

755 but fall well within the spread of CMIP6 models. CanESM5-CanOE has lower cumulative

756 uptake than CanESM5 by ~10 PgC. As the models were not fully equilibrated when the

757 historical run was launched, this difference does not necessarily arise from the biogeochemical

758 model structure; part of the difference can be attributed to differences in the spinup protocol (cf.

759 Séférian et al., 2016). The drift in the piControl experiment over the 165 years from the

760 branching off of the historical experiment is -10.5 0 PgC in CanESM5-CanOE and -5.9 1 PgC in

761 CanESM5 (see Supplementary Table S6), so drift accounts for about half (48 4%) of the

762 difference in net ocean CO_2 uptake. The vertical spatial distribution of anthropogenic DIC is

763 very similar between CanESM5 and CanESM5-CanOE (not shownSupplementary Figure S7).

764

765   The long-term trend in global total export production is shown in Figure 21. The model values

766   must be normalized in order to compare trends, since the differences among means are large

767   compared to the changes over the historical period (Figure 19). Such trends are difficult or

768   impossible to meaningfully constrain with observations, but the general expectation has been that

769   export will decline somewhat due to increasing stratification (e.g., Steinacher et al., 2010).

770   CanESM5 shows a greater decline than most other CMIP6 models, while CanESM5-CanOE is

771   more similar to non-CanESM models. Such trends are difficult or impossible to meaningfully

772   constrain with observations, but the general expectation has been that export will decline

773   somewhat due to increasing stratification (e.g., Steinacher et al., 2010). The change in CanESM5

774   is geographically widespread and not concentrated in a specific region or regions: export is

775   maximal in the tropics and the northern and southern mid-latitudes (Figure 19b) and declines

776   over the historical period in all of these regions (Supplementary Figure S8$\underline{8}$1). In CanESM5-

777   CanOE, export declines in the same regions, but the magnitude of the change is smaller, and in

778   the Southern Ocean increases and decreases in different latitude bands largely offset each other.

779

780   The trend in the volume of ocean water with O_2 concentration less than 6 or 60 mmol m^{-3} is

781   shown in Figure 22. Again, the totals are normalized to a value close to the preindustrial, as the

782   differences among models are large (Figure 5). For the volume with <60 mmol m^{-3}, CanESM

783   models show relatively little change; in CanESM5 the volume actually declines slightly, while in

784   CanESM5-CanOE it increases, but the total change is <1% in each case. As with the baseline

785   volumes, the range among models is large, with one model showing an increase approaching

786   10% of the total volume estimated for WOA2013$\underline{2018}$ (Figures 5b and 22b). For the volume

787   with <6 mmol m^{-3} (Figure 22a), CanESM models are among the most stable over time. In

788   CanESM5, the volume again declines, although this is within the range of internal variability.

789   Again some models show fairly large excursions, but in this case none shows a strong secular

790   trend over the last half-century.

791

792   **4. Discussion**

793

794   CanESM5 and CanESM5-CanOE are new coupled ocean-atmosphere climate models with

795   prognostic ocean biogeochemistry. The two have the same physical climate (in experiments with

796   specified atmospheric CO_2) and differ only in their ocean biogeochemistry components.

797   CanESM5-CanOE has a much more complex biogeochemistry model including a prognostic iron

798   cycle. We have presented results that assess how these two models simulate the overall

799   distribution of major tracers like DIC, alkalinity, nitrate and oxygen, as well as analyses of the

800   interaction of the iron and nitrogen cycles, plankton community structure, export of organic

801   matter from the euphotic zone, and historical trends over 1850-2014.

802

803   The overall distribution of major tracers indicates that both models do a reasonable job of

804   simulating both biogeochemical (e.g., export and remineralization of organic matter) and

805   physical (e.g., deep and intermediate ocean ventilation) processes. The volume of ocean with

806   oxygen concentration below 6 or 60 µM compares favourably with other CMIP6 models (Figure

807   5), and is among the most stable over historical time (Figure 22). CanESM5-CanOE has a

808   substantially lower volume of water with $[O_2]<6$ µM than CanESM5 and much closer to

809   observation-based estimates (Figure 22a). Both models are biased slightly low in terms of

810   historical uptake of anthropogenic CO_2, which may indicate weak Southern Ocean upwelling or

811   too shallow remineralization of DIC or both (Figure 20). The spatial distribution of

812   anthropogenic DIC is very similar between the two models (Supplementary Figure S7), which is

813   expected as it is mainly a function of the physical ocean model circulation. However, CanESM5

814   has higher concentrations in the main areas of accumulation, particularly the North Atlantic and

815   the Southern Ocean. This probably indicates more efficient removal and export of 'natural' DIC

816   by the plankton, particularly in the Southern Ocean upwelling zone (Figure 19), and deeper

817   average remineralization, with the caveat that the preindustrial control simulations had different

818   degrees of equilibration when the historical experiment was launched (cf. Séférian et al., 2016,

819   Supplementary Table S6).

820

821   Analysis of phytoplankton and zooplankton biomass concentrations show that CanESM5 and

822   CanESM5-CanOE compare somewhat favourably with available observational data but do have

823   distinct biases. In particular, both zooplankton biomass and detrital organic matter concentration

824   tend to be very low in CanESM5-CanOE; the total biomass of the plankton community and the

825   standing crop of particulate organic matter are dominated by phytoplankton (e.g., Figure 17).

826   Regional biases differ between the two models, with CanESM5-CanOE showing excessively

827   large phytoplankton biomass in the tropics. We note, however, that the seasonal cycle of

828   equatorial upwelling and the formation of the equatorial Pacific HNLC are reproduced rather

829   well by our models (e.g., Figures 11, 15 and S6), and that CanESM5-CanOE is the first CanESM

830   model to have genuinely simulated this as an emergent property (see section 3.3). In CanESM5-

831   CanOE, decoupling of large and small phytoplankton populations associated with seasonal

832   upwelling or convection (see below) is observed in some regions but not others.

833

834 Global export production is biased low, particularly in CanESM5-CanOE. This is due in part to

835 the biogeochemical model and in part to ocean circulation. CanESM5 has the same ocean

836 biology as CanESM2 but a different physical ocean model, and global ocean export production is

837 substantially lower in CanESM5. It is lower still in CanESM5-CanOE (Figure 19). We note that

838 CanESM5 performs better than CanESM2 on most metrics of physical ocean model evaluation

839 (Swart et al., 2019a), and shows a more realistic distribution of major tracers like DIC (Figure 8).

840 While the range of observation-based estimates of global ocean export production is large, and

841 encompasses the full range of CMIP5 and CMIP6 models, the change between CanESM2 and

842 CanESM5 is large. Changes in the physical ocean are not entirely independent of the

843 biogeochemistry model even when the latter is ostensibly identical. In CanESM2 and CanESM5,

844 iron limitation is specified as a spatially static 'mask' based on the observed distribution of

845 surface nitrate, and it is possible that in these two models ocean upwelling occurs in different

846 places relative to the specified boundary of the region of Southern Ocean iron imitation (Figure 3

847 of Zahariev et al., 2008). It is also possible that the lower export production in CanESM5-CanOE

848 is due to low iron supply to the surface waters of the Southern Ocean, but comparison with

849 available observations do not suggest that this is the case. Several biases are common to

850 CanESM5 and CanESM5-CanOE that relate to Southern Ocean upwelling (high Southern Ocean

851 surface nitrate concentration, low export production, weak anthropogenic CO_2 uptake) and so are

852 probably more attributable to the physical ocean model than to the Fe submodel. The difference

853 between CanESM2 and CanESM5 bears this out.

854

855 The development of CanOE was undertaken in response to some of the most severe limitations

856 of CanESM1/2, and in light of our collective experience. In addition to CMOC (Zahariev et al.,

857   2008), previous models developed by members of our group include Denman and Peña (1999;

858   2002), Christian et al. (2002a; 2002b), Christian (2005), and Denman et al. (2006). Christian et

859   al. (2002a) had a prognostic Fe cycle and multiple phytoplankton and zooplankton species, but

860   had fixed elemental ratios. Christian (2005) incorporated a cellular-regulation model, but only

861   for a single species and without Fe limitation. Christian (2005) had prognostic chlorophyll

862   whereas Denman and Peña (1999; 2002) and Christian et al. (2002a) used an irradiance-

863   dependent diagnostic formulation. Christian et al. (2002a) used multiplicative (Franks et al.,

864   1986) grazing, which creates stability in predator-prey interactions but severely limits

865   phytoplankton biomass accumulation under nutrient-replete conditions.

866

867   One of the most important lessons from Christian et al. (2002a; 2002b) was that when a fixed

868   Fe/N ratio is employed, sensitivity to this parameter is extreme. Because Fe cell quotas are far

869   more variable than N, P, or Si quotas, treating this parameter as constant results in the specified

870   value influencing the overall solution far more than any other parameter. CanESM5-CanOE

871   largely succeeded in creating a prognostic Fe-N limitation model that produces HNLC conditions

872   in the expected regions (Figures 10, 11, 14, 15, S6), although surface nitrate concentration is low

873   relative to observation-based estimates in some cases. External Fe sources and scavenging

874   parameterizations will be revisited and refined in future versions. In CanESM5-CanOE the

875   scavenging model is very simple, with distinct regimes for concentrations greater or less than 0.6

876   nM; scavenging rates are very high above this threshold which causes deep-water concentrations

877   to converge on this value. The generally nutrient-like profile suggest that in CanOE the

878   scavenging rate is quite low for concentrations below 0.6 nM (Figure 13; see also Supplementary

879   Figure S9h). We note that the aeolian mineral dust deposition field employed here is derived

880   from the CanESM atmosphere model; these processes are not presently interactive but could be

881   made so in the future.

882

883   A particular issue with CanESM2 was that extremely high concentrations of nitrate occurred

884   under the ~~Eastern Boundary Current (~~EBC~~)~~ upwelling regions. This error resulted from

885   spreading denitrification out over the ocean basin so that introduction of new fixed N from N_2

886   fixation would balance denitrification losses within each vertical column, whereas in the real

887   world denitrification is highly localized in the low oxygen environments under the EBCs.

888   CanESM2 did not include oxygen, but CanESM5 ~~CMOC~~ incorporates oxygen as a 'downstream'

889   tracer that does not feed back on other biogeochemical processes. The incorporation of a more

890   process-based denitrification parameterization in CanESM5-CanOE is independent of the many

891   other processes that are present in CanESM5-CanOE but not in CanESM5~~CMOC~~: a CMOC-like

892   model with prognostic denitrification is clearly an option. We chose not to include explicit,

893   oxygen-dependent denitrification in CanESM5 because we wanted to maintain a CMOC-based

894   model as close to the CanESM2 version as possible, and because oxygen would not then be a

895   downstream tracer that does not affect other processes.

896

897   ~~CanOE for the most part successfully reproduces the overall distribution of major tracers such as~~

898   ~~nitrate, oxygen, DIC and alkalinity (and dFe, to the extent that its distribution is known). One~~

899   ~~could argue that the gains made relative to CMOC are incremental. However, it is also important~~

900   ~~to note that CanOE explicitly simulates important processes that are highly parameterized or~~

901   ~~specified in CMOC. For example, the maintenance of HNLC regions is hardwired into CMOC~~

902   ~~by specifying iron limitation as a function of the present-day observed distribution of surface~~

903 ~~nitrate. Both models show substantial gains in skill relative to CanESM2. These gains are similar~~

904 ~~in the two models and in coupled or ocean only (with CanESM2 forcing) mode (not shown) and~~

905 ~~are, therefore, attributable primarily to improvements in the ocean circulation model, although~~

906 ~~differences in initialization and spinup may also play a small role (e.g., Séférian et al., 2016).~~

907

908 Plankton community structure in CanESM5-CanOE is somewhat biased toward high

909 concentrations of phytoplankton, low concentrations of zooplankton and detritus, and low export

910 (Sect. 3.4). In the development phase, a fair number of experiments were conducted with various

911 values of the grazing rates and detritus sinking speeds. A wide range of values of these

912 parameters was tested, with no resulting improvement in the overall results. Possibly the detrital

913 remineralization rates are too high, although primary production is also on the low end of the

914 CMIP6 range (not shown), and would probably decline further if these rates were decreased. The

915 model was designed around the Armstrong (1994) hypothesis of 'supplementation' vs

916 'replacement', i.e., small phytoplankton and their grazers do not become much more abundant in

917 more nutrient-rich environments but rather stay at about the same level and are joined by larger

918 species that are absent in more oligotrophic conditions (see also Chisholm, 1992; Landry et al.,

919 1997; Friedrichs et al., 2007). The results presented here suggest that this was partially achieved

920 but further improvement is possible (Figure 17).

921

922 As to whether the gains in skill with CanESM5-CanOE justify the extra computational cost,

923 Taylor diagrams (Figures 4, 8, 9, and Supplementary Figure ~~S2~~S4) show a modest but consistent

924 gain in skill at simulating the major biogeochemical species (O_2, DIC, alkalinity) across

925 variables and depths, especially for alkalinity at mid-depths (Supplementary Figure ~~S2~~S4), for

926 which CanESM5 displays the least skill relative to other fields or depths. Other processes that are

927 highly parameterized in CanESM5CMOC, such as calcification and $CaCO_3$ dissolution, were not

928 addressed in detail in this paper, but are an important factor in determining the subsurface

929 distribution of alkalinity. As noted above for maintenance of HNLC conditions, Again, we

930 emphasize that we are simulating as an emergent property something that is parameterized in

931 CanESM5CMOC (as previously noted for surface nitrate concentration in HNLC regions), and

932 doing at least as well in terms of model skill. As a general rule, the potential for improving skill

933 and achieving better results in novel environments (e.g., topographically complex regional

934 domains like the Arctic Ocean and the boreal marginal seas), is expected to be greater in less

935 parameterized models (e.g., Friedrichs et al., 2007; Tesdal et al., 2016).

936

937 An updated version of CanESM5 CMOC with prognostic denitrification is clearly possible.

938 However, for the reasons discussed above, a prognostic Fe cycle with a fixed phytoplankton

939 Fe/N remains problematic, and the model would still have a single detritus sinking speed and

940 remineralization length scale. We are also developing CanOE for regional downscaling

941 applications (Hayashida, 2018; Holdsworth et al., 2021), and it is likely that the simplification of

942 having a single particle sinking speed is not well suited to a domain with complex topography

943 and prominent continental shelf and slope. The number of tracers in CanOE is not particularly

944 large compared with other CMIP6 models. We expect to further refine CanOE and its

945 parameterizations, evaluate it against new and emerging ocean data sets (e.g., GEOTRACES,

946 biogeochemical Argo), and incrementally improve CMOC (which we will maintain for a wide

947 suite of physical-climate experiments for which ocean biogeochemistry is not central to the

948 purpose). For CMIP6, we chose to keep CMOC as close to the CanESM2 version as possible.

949   This strategy allows us to quantify how much of the improvement in model skill is due to the

950   physical circulation, ~~which is in fact substantial~~as is illustrated by greater skill with respect to

951   DIC (~~e.g.,~~ Figure 8) and alkalinity (Supplementary Figure S4), particularly at intermediate depths

952   (400-900 m). The CanESM terrestrial carbon model is also undergoing important new

953   developments (e.g., Asaadi and Arora, 2021) and we expect CanESM to continue to offer a

954   credible contribution to global carbon cycle studies, as well as advancing regional downscaling

955   and impacts science.

956

957   *Code availability*. The full CanESM5 source code is publicly available at

958   gitlab.com/cccma/canesm; within this tree the ~~CMOC/CanOE~~ ocean biogeochemistry code can

959   be found at gitlab.com/cccma/cannemo/-

960   /tree/v5.0.3/nemo/CONFIG/CCC_CANCPL_ORCA1_LIM_CMOC or

961   CCC_CANCPL_ORCA1_LIM_CANOE (last access: 21 September 2021). The version of the

962   code which can be used to produce all the simulations submitted to CMIP6, and described in this

963   paper, is tagged as v5.0.3 and has the associated DOI: https://doi.org/10.5281/zenodo.3251113

964   (Swart et al., 2019b).

965

966   *Data availability*. All ~~CanESM5~~ simulations conducted for CMIP6, including those described in

967   this paper, are publicly available via the Earth System Grid Federation ~~(ESGF)~~ (source_id =

968   CanESM5 or CanESM5-CanOE). All observational data and other CMIP6 model data used are

969   publicly available.

970

971 *Author contributions.* Formulation of the overall research goals and aims: JRC, KLD, NS, NCS;

972 Implementation and testing of the model code: JRC, HH, AMH, WGL, OGJR, AES, NCS;

973 Carrying out the experiments: JRC, WGL, OGJR, AES, NCS; Creation of the published work:

974 JRC, HH, AMH, AES, NS, NCS.

975

976 *Competing interests*. The authors declare that they have no conflict of interest.

977

978 *Disclaimer*. CanESM has been customized to run on the ECCC high-performance computer, and

979 a significant fraction of the software infrastructure used to run the model is specific to the

980 individual machines and architecture. While we publicly provide the code, we cannot provide

981 any support for migrating the model to different machines or architectures.

982

983 *Acknowledgments*. This work was made possible by the combined efforts of the CCCMa model

984 development team and computing support team. We thank all of the the data contributors to and

985 developers of the observational based data products, the NASA ocean colour team, and all of the

986 CMIP6 data contributors. The Python packages mocsy by Jim Orr and SkillMetrics by Peter

987 Rochford were invaluable tools in the analysis. William Merryfield and Andrew Ross made

988 useful comments on an earlier draft. Fiona Davidson helped with figure preparation. This paper

989 is dedicated to the memory of Mr. Fouad Majaess, who supported CCCMa supercomputer users

990 for many years and passed away suddenly in 2020.

991
992
993 **Literature Cited**
994
995

Arguez, A. and Vose, R.: The Definition of the Standard WMO Climate Normal The Key to Deriving Alternative Climate Normals, Bulletin of the American Meteorological Society, 92, 699-704, 10.1175/2010BAMS2955.1, 2011.

Armstrong, R.: Grazing limitation and nutrient limitation in marine ecosystems - steady-state solutions of an ecosystem model with multiple food-chains, Limnology and Oceanography, 39, 597-608, 1994.

Arora, V., Scinocca, J., Boer, G., Christian, J., Denman, K., Flato, G., Kharin, V., Lee, W., and Merryfield, W.: Carbon emission limits required to satisfy future representative concentration pathways of greenhouse gases, Geophysical Research Letters, 38, 10.1029/2010GL046270, 2011.

Arora, V., Boer, G., Christian, J., Curry, C., Denman, K., Zahariev, K., Flato, G., Scinocca, J., Merryfield, W., and Lee, W.: The Effect of Terrestrial Photosynthesis Down Regulation on the Twentieth-Century Carbon Budget Simulated with the CCCma Earth System Model, Journal of Climate, 22, 6066-6088, 10.1175/2009JCLI3037.1, 2009.

Arora, V., Katavouta, A., Williams, R., Jones, C., Brovkin, V., Friedlingstein, P., Schwinger, J., Bopp, L., Boucher, O., Cadule, P., Chamberlain, M., Christian, J., Delire, C., Fisher, R., Hajima, T., Ilyina, T., Joetzjer, E., Kawamiya, M., Koven, C., Krasting, J., Law, R., Lawrence, D., Lenton, A., Lindsay, K., Pongratz, J., Raddatz, T., Seferian, R., Tachiiri, K., Tjiputra, J., Wiltshire, A., Wu, T., and Ziehn, T.: Carbon-concentration and carbon-climate feedbacks in CMIP6 models and their comparison to CMIP5 models, Biogeosciences, 17, 4173-4222, 10.5194/bg-17-4173-2020, 2020.

Asaadi, A. and Arora, V.: Implementation of nitrogen cycle in the CLASSIC land model, Biogeosciences, 18, 669-706, 10.5194/bg-18-669-2021, 2021.

Aumont, O., Ethe, C., Tagliabue, A., Bopp, L., and Gehlen, M.: PISCES-v2: an ocean biogeochemical model for carbon and ecosystem studies, Geoscientific Model Development, 8, 2465-2513, 10.5194/gmd-8-2465-2015, 2015.

Babbin, A., Keil, R., Devol, A., and Ward, B.: Organic Matter Stoichiometry, Flux, and Oxygen Control Nitrogen Loss in the Ocean, Science, 344, 406-408, 10.1126/science.1248364, 2014.

Chisholm, S.W.: Phytoplankton size, in: Primary productivity and biogeochemical cycles in the sea, edited by: Falkowski, P.G., and Woodhead A.D., Plenum, New York, 213-237, 1992.

Christian, J.: Biogeochemical cycling in the oligotrophic ocean: Redfield and non-Redfield models, Limnology and Oceanography, 50, 646-657, 2005.

Christian, J. and Karl, D.: Microbial community structure at the United States Joint Global Ocean Flux Study Station ALOHA - inverse methods for estimating biochemical indicator ratios, Journal of Geophysical Research-Oceans, 99, 14269-14276, 10.1029/94JC00681, 1994.

Christian, J., Arora, V., Boer, G., Curry, C., Zahariev, K., Denman, K., Flato, G., Lee, W., Merryfield, W., Roulet, N., and Scinocca, J.: The global carbon cycle in the Canadian Earth system model (CanESM1): Preindustrial control simulation, Journal of Geophysical Research-Biogeosciences, 115, 10.1029/2008JG000920, 2010.

Christian, J. R., Verschell, M. A., Murtugudde, R., Busalacchi, A. J., and McClain, C. R.: Biogeochemical modelling of the tropical Pacific Ocean. I: Seasonal and interannual variability, Deep-Sea Research Part II-Topical Studies in Oceanography, 49, 509-543, 2002a.

Christian, J. R., Verschell, M. A., Murtugudde, R., Busalacchi, A. J., and McClain, C. R.: Biogeochemical modelling of the tropical Pacific Ocean. II: Iron biogeochemistry, Deep-Sea Research Part II-Topical Studies in Oceanography, 49, 545-565, 0.1016/s0967-0645(01)00111-4, 2002b.

Denman, K. and Pena, M.: A coupled 1-D biological/physical model of the northeast subarctic Pacific Ocean with iron limitation, Deep-Sea Research Part II-Topical Studies in Oceanography, 46, 2877-2908, 10.1016/S0967-0645(99)00087-9, 1999.

Denman, K. and Pena, M.: The response of two coupled one-dimensional mixed layer/planktonic ecosystem models to climate change in the NE subarctic Pacific Ocean, Deep-Sea Research Part II-Topical Studies in Oceanography, 49, 5739-5757, 10.1016/S0967-0645(02)00212-6, 2002.

Denman, K., Voelker, C., Pena, M., and Rivkin, R.: Modelling the ecosystem response to iron fertilization in the subarctic NE Pacific: The influence of grazing, and Si and N cycling on CO_2 drawdown, Deep-Sea Research Part II-Topical Studies in Oceanography, 53, 2327-2352, 10.1016/j.dsr2.2006.05.026, 2006.

Devol, A.H.: Denitrification including anammox, in: Nitrogen in the Marine Environment, 2nd edition, edited by: Capone, D.G., Bronk D.A., Mulholland M.R., and Carpenter E.J., Elsevier, Amsterdam, 263-301, 2008.

Dickson, A.G., Sabine, C.L., and Christian, J.R., eds.: Guide to best practices for ocean CO_2 measurements, PICES Special Publication #3, 2007

Elser, J. and Urabe, J.: The stoichiometry of consumer-driven nutrient recycling: Theory, observations, and consequences, Ecology, 80, 735-751, 10.1890/0012-9658(1999)080[0735:TSOCDN]2.0.CO;2, 1999.

Franks, P., Wroblewski, J., and Flierl, G.: Behavior of a simple plankton model with food-level acclimation by herbivores, Marine Biology, 91, 121-129, 10.1007/BF00397577, 1986.

Friedlingstein, P., O'Sullivan, M., Jones, M., Andrew, R., Hauck, J., Olsen, A., Peters, G., Peters, W., Pongratz, J., Sitch, S., Le Quere, C., Canadell, J., Ciais, P., Jackson, R., Alin, S., Aragao, L., Arneth, A., Arora, V., Bates, N., Becker, M., Benoit-Cattin, A., Bittig, H., Bopp, L., Bultan, S., Chandra, N., Chevallier, F., Chini, L., Evans, W., Florentie, L., Forster, P., Gasser, T., Gehlen, M., Gilfillan, D., Gkritzalis, T., Gregor, L., Gruber, N., Harris, I., Hartung, K., Haverd, V.,

Houghton, R., Ilyina, T., Jain, A., Joetzjer, E., Kadono, K., Kato, E., Kitidis, V., Korsbakken, J., Landschutzer, P., Lefevre, N., Lenton, A., Lienert, S., Liu, Z., Lombardozzi, D., Marland, G., Metzl, N., Munro, D., Nabel, J., Nakaoka, S., Niwa, Y., O'Brien, K., Ono, T., Palmer, P., Pierrot, D., Poulter, B., Resplandy, L., Robertson, E., Rodenbeck, C., Schwinger, J., Seferian, R., Skjelvan, I., Smith, A., Sutton, A., Tanhua, T., Tans, P., Tian, H., Tilbrook, B., Van der Werf, G., Vuichard, N., Walker, A., Wanninkhof, R., Watson, A., Willis, D., Wiltshire, A., Yuan, W., Yue, X., and Zaehle, S.: Global Carbon Budget 2020, Earth System Science Data, 12, 3269-3340, 10.5194/essd-12-3269-2020, 2020.

Friedrichs, M., Dusenberry, J., Anderson, L., Armstrong, R., Chai, F., Christian, J., Doney, S., Dunne, J., Fujii, M., Hood, R., McGillicuddy, D., Moore, J., Schartau, M., Spitz, Y., and Wiggert, J.: Assessment of skill and portability in regional marine biogeochemical models: Role of multiple planktonic groups, Journal of Geophysical Research-Oceans, 112, 10.1029/2006JC003852, 2007.

Garcia, H.E., Locarnini, R.A., Boyer, T.P., Antonov, J.I., Baranova, O.K., Zweng, M.M., Reagan, J.R., and Johnson, D.R.: World Ocean Atlas 2013, Volume 3: Dissolved Oxygen, Apparent Oxygen Utilization, and Oxygen Saturation. S. Levitus, Ed., A. Mishonov Technical Ed., NOAA Atlas NESDIS 75, 27 pp., 2014.

Garcia, H.E., Weathers, K., Paver, C.R., Smolyar, I., Boyer, T.P., Locarnini, R.A., Zweng, M.M., Mishonov, A.V., Baranova, O.K., Seidov, D. and Reagan, J.R.: World Ocean Atlas 2018, Volume 3: Dissolved Oxygen, Apparent Oxygen Utilization, and Oxygen Saturation. A. Mishonov Technical Ed., NOAA Atlas NESDIS 83, 38pp., 2018a.

Garcia, H.E., Locarnini, R.A., Boyer, T.P., Antonov, J.I., Baranova, O.K., Zweng, M.M., Reagan, J.R., and Johnson, D.R.: World Ocean Atlas 2013, Volume 4: Dissolved Inorganic Nutrients (phosphate, nitrate, silicate). S. Levitus, Ed., A. Mishonov Technical Ed.; NOAA Atlas NESDIS 76, 25 pp., 2014.

Garcia, H.E., Weathers, K., Paver, C.R., Smolyar, I., Boyer, T.P., Locarnini, R.A., Zweng, M.M., Mishonov, A.V., Baranova, O.K., Seidov, D. and Reagan, J.R.: World Ocean Atlas 2018, Volume 4: Dissolved Inorganic Nutrients (phosphate, nitrate and nitrate+nitrite, silicate). A. Mishonov Technical Ed., NOAA Atlas NESDIS 84, 35pp., 2018b.

Geider, R., MacIntyre, H., and Kana, T.: A dynamic regulatory model of phytoplanktonic acclimation to light, nutrients, and temperature, Limnology and Oceanography, 43, 679-694, 1998.

Guilyardi, E. and Madec, G.: Performance of the OPA/ARPEGE-T21 global ocean-atmosphere coupled model, Climate Dynamics, 13, 149-165, 10.1007/s003820050157, 1997.

Hayashida, H.: Modelling sea-ice and oceanic dimethylsulde production and emissions in the Arctic, PhD thesis, University of Victoria, 2018.

Holdsworth, A.M., Zhai, L., Lu, Y., and Christian, J.R.: Future changes in oceanography and

biogeochemistry along the Canadian Pacific continental margin, Frontiers in Marine Science, 10.3389/fmars.2021.602991, 2021.

Jetten, M., Wagner, M., Fuerst, J., van Loosdrecht, M., Kuenen, G., and Strous, M.: Microbiology and application of the anaerobic ammonium oxidation ('anammox') process, Current Opinion in Biotechnology, 12, 283-288, 10.1016/S0958-1669(00)00211-1, 2001.

Johnson, K. S., Gordon, R. M., and Coale, K. H.: What controls dissolved iron concentrations in the world ocean?, Marine Chemistry, 57, 137-161, 10.1016/s0304-4203(97)00043-1, 1997.

Johnson, K., Elrod, V., Fitzwater, S., Plant, J., Chavez, F., Tanner, S., Gordon, R., Westphal, D., Perry, K., Wu, J., and Karl, D.: Surface ocean-lower atmosphere interactions in the Northeast Pacific Ocean Gyre: Aerosols, iron, and the ecosystem response, Global Biogeochemical Cycles, 17, 10.1029/2002GB002004, 2003.

Key, R.M., Olsen, A., van Heuven, S., Lauvset, S. K., Velo, A., Lin, X., Schirnick, C., Kozyr, A., Tanhua, T., Hoppema, M., Jutterström, S., Steinfeldt, R., Jeansson, E., Ishii, M., Perez, F. F., and Suzuki, T.: 2015. Global Ocean Data Analysis Project, Version 2 (GLODAPv2), ORNL/CDIAC-162, NDP-093. Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, US Department of Energy, Oak Ridge, Tennessee. doi:10.3334/CDIAC/OTG.NDP093_GLODAPv2

Lambert, S. and Boer, G.: CMIP1 evaluation and intercomparison of coupled climate models, Climate Dynamics, 17, 83-106, 10.1007/PL00013736, 2001.

Landry, M., Barber, R., Bidigare, R., Chai, F., Coale, K., Dam, H., Lewis, M., Lindley, S., McCarthy, J., Roman, M., Stoecker, D., Verity, P., and White, J.: Iron and grazing constraints on primary production in the central equatorial Pacific: An EqPac synthesis, Limnology and Oceanography, 42, 405-418, 10.4319/lo.1997.42.3.0405, 1997.

Lauvset, S., Key, R., Olsen, A., van Heuven, S., Velo, A., Lin, X., Schirnick, C., Kozyr, A., Tanhua, T., Hoppema, M., Jutterstrom, S., Steinfeldt, R., Jeansson, E., Ishii, M., Perez, F., Suzuki, T., and Watelet, S.: A new global interior ocean mapped climatology: the 1 degrees x 1 degrees GLODAP version 2, Earth System Science Data, 8, 325-340, 10.5194/essd-8-325-2016, 2016.

Li, Y.: Distribution patterns of the elements in the ocean - a synthesis, Geochimica et Cosmochimica Acta, 55, 3223-3240, 1991.

Locarnini, R.A., Mishonov, A.V., Antonov, J.I., Boyer, T.P., Garcia, H.E., Baranova, O.K., Zweng, M.M., Paver, C.R., Reagan, J.R., Johnson, D.R., Hamilton, M., and Seidov, D.: World Ocean Atlas 2013, Volume 1: Temperature. S. Levitus, Ed., A. Mishonov Technical Ed., NOAA Atlas NESDIS 73, 40 pp., 2013.

Locarnini, R.A., Mishonov, A.V., Antonov, J.I., Boyer, T.P., Garcia, H.E., Baranova, O.K., Zweng, M.M., and Johnson, D.R.: World Ocean Atlas 2009, Volume 1: Temperature. Levitus,

S., Ed., NOAA Atlas NESDIS 68, 184 pp., 2010.

Locarnini, R.A., Mishonov, A.V., Baranova, O.K., Boyer, T.P., Zweng, M.M., Garcia, H.E., Reagan, J.R., Seidov, D., Weathers, K., Paver, C.R., and Smolyar, I.: World Ocean Atlas 2018, Volume 1: Temperature. A. Mishonov Technical Ed., NOAA Atlas NESDIS 81, 52pp., 2018.

Loladze, I., Kuang, Y., and Elser, J.: Stoichiometry in producer-grazer systems: Linking energy flow with element cycling, Bulletin of Mathematical Biology, 62, 1137-1162, 10.1006/bulm.2000.0201, 2000.

Lukas, R.: Pacific ocean equatorial currents, in: Encyclopedia of Ocean Sciences, edited by: Steele, J.H., Thorpe, S.A., and Turekian, K.K., Academic, San Diego, 2069-2076, 10.1006/rwos.2001.0354, 2001.

Madec, G. and Imbard, M.: A global ocean mesh to overcome the North Pole singularity, Climate Dynamics, 12, 381-388, 10.1007/s003820050115, 1996.

Moriarty, R. and O'Brien, T.: Distribution of mesozooplankton biomass in the global ocean, Earth System Science Data, 5, 45-55, 10.5194/essd-5-45-2013, 2013.

Nozaki, Y.: Elemental Distribution, in: Encyclopedia of Ocean Sciences, edited by: Steele, J.H., Thorpe, S.A., and Turekian, K.K., Academic, San Diego, 840-845, 10.1006/rwos.2001.0402, 2001.

Orr, J. C., Najjar, R. G., Aumont, O., Bopp, L., Bullister, J. L., Danabasoglu, G., Doney, S. C., Dunne, J. P., Dutay, J. C., Graven, H., Griffies, S. M., John, J. G., Joos, F., Levin, I., Lindsay, K., Matear, R. J., McKinley, G. A., Mouchet, A., Oschlies, A., Romanou, A., Schlitzer, R., Tagliabue, A., Tanhua, T., and Yool, A.: Biogeochemical protocols and diagnostics for the CMIP6 Ocean Model Intercomparison Project (OMIP), Geoscientific Model Development, 10, 2169-2199, 10.5194/gmd-10-2169-2017, 2017.

Philander, S.G. and Chao, Y.: On the contrast between the seasonal cycles of the equatorial Atlantic and Pacific oceans, Journal of Physical Oceanography, 21, 1399-1406, 1991.

Raven, J. and Geider, R.: Temperature and algal growth, New Phytologist, 110, 441-461, 10.1111/j.1469-8137.1988.tb00282.x, 1988.

Riche, O. and Christian, J.: Ocean dinitrogen fixation and its potential effects on ocean primary production in Earth system model simulations of anthropogenic warming, Elementa-Science of the Anthropocene, 6, 10.1525/elementa.277, 2018.

Schlitzer, R., Anderson, R., Dodas, E., Lohan, M., Geibere, W., Tagliabue, A., Bowie, A., Jeandel, C., Maldonado, M., Landing, W., Cockwell, D., Abadie, C., Abouchami, W., Achterberg, E., Agather, A., Aguliar-Islas, A., van Aken, H., Andersen, M., Archer, C., Auro, M., de Baar, H., Baars, O., Baker, A., Bakker, K., Basak, C., Baskaran, M., Bates, N., Bauch, D., van Beek, P., Behrens, M., Black, E., Bluhm, K., Bopp, L., Bouman, H., Bowman, K., Bown, J.,

Boyd, P., Boye, M., Boyle, E., Branellec, P., Bridgestock, L., Brissebrat, G., Browning, T., Bruland, K., Brumsack, H., Brzezinski, M., Buck, C., Buck, K., Buesseler, K., Bull, A., Butler, E., Cai, P., Mor, P., Cardinal, D., Carlson, C., Carrasco, G., Casacuberta, N., Casciotti, K., Castrillejo, M., Chamizo, E., Chance, R., Charette, M., Chaves, J., Cheng, H., Chever, F., Christl, M., Church, T., Closset, I., Colman, A., Conway, T., Cossa, D., Croot, P., Cullen, J., Cutter, G., Daniels, C., Dehairs, F., Deng, F., Dieu, H., Duggan, B., Dulaquais, G., Dumousseaud, C., Echegoyen-Sanz, Y., Edwards, R., Ellwood, M., Fahrbach, E., Fitzsimmons, J., Flegal, A., Fleisher, M., van de Flierdt, T., Frank, M., Friedrich, J., Fripiat, F., Frollje, H., Galer, S., Gamo, T., Ganeshram, R., Garcia-Orellana, J., Garcia-Solsona, E., Gault-Ringold, M., George, E., Gerringa, L., Gilbert, M., Godoy, J., Goldstein, S., Gonzalez, S., Grissom, K., Hammerschmidt, C., Hartman, A., Hassler, C., Hathorne, E., Hatta, M., Hawco, N., Hayes, C., Heimburger, L., Helgoe, J., Heller, M., Henderson, G., Henderson, P., van Heuven, S., Ho, P., Horner, T., Hsieh, Y., Huang, K., Humphreys, M., Isshiki, K., Jacquot, J., Janssen, D., Jenkins, W., John, S., Jones, E., Jones, J., Kadko, D., Kayser, R., Kenna, T., Khondoker, R., Kim, T., Kipp, L., Klar, J., Klunder, M., Kretschmer, S., Kumamoto, Y., Laan, P., Labatut, M., Lacan, F., Lam, P., Lambelet, M., Lamborg, C., Le Moigne, F., Le Roy, E., Lechtenfeld, O., Lee, J., Lherminier, P., Little, S., Lopez-Lora, M., Lu, Y., Masque, P., Mawji, E., Mcclain, C., Measures, C., Mehic, S., Menzel Barraqueta, J., van der Merwe, P., Middag, R., Mieruch, S., Milne, A., Minami, T., Moffett, J., Moncoiffe, G., Moore, W., Morris, P., Morton, P., Nakaguchi, Y., Nakayama, N., Niedermiller, J., Nishioka, J., Nishiuchi, A., Noble, A., Obata, H., Ober, S., Ohnemus, D., van Ooijen, J., O'Sullivan, J., Owens, S., Pahnke, K., Paul, M., Pavia, F., Pena, L., Petersh, B., Planchon, F., Planquette, H., Pradoux, C., Puigcorbe, V., Quay, P., Queroue, F., Radic, A., Rauschenberg, S., Rehkamper, M., Rember, R., Remenyi, T., Resing, J., Rickli, J., Rigaud, S., Rijkenberg, M., Rintoul, S., Robinson, L., Roca-Marti, M., Rodellas, V., Roeske, T., Rolison, J., Rosenberg, M., Roshan, S., van der Loaff, M., Ryabenko, E., Saito, M., Salt, L., Sanial, V., Sarthou, G., Schallenberg, C., Schauer, U., Scher, H., Schlosser, C., Schnetger, B., Scott, P., Sedwick, P., Semiletov, I., Shelley, R., Sherrell, R., Shiller, A., Sigman, D., Singh, S., Slagter, H., Slater, E., Smethie, W., Snaith, H., Sohrin, Y., Sohst, B., Sonke, J., Speich, S., Steinfeldt, R., Stewart, G., Stichel, T., Stirling, C., Stutsman, J., Swarr, G., Swift, J., Thomas, A., Thorne, K., Till, C., Till, R., Townsend, A., Townsend, E., Tuerena, R., Twining, B., Vance, D., Velazquez, S., Venchiarutti, C., Villa-Alfageme, M., Vivancos, S., Voelker, A., Wake, B., Warner, M., Watson, R., van Weerlee, E., Weigand, M., Weinstein, Y., Weiss, D., Wisotzki, A., Woodward, E., Wu, J., Wu, Y., Wuttig, K., Wyatt, N., Xiang, Y., Xie, R., Xue, Z., Yoshikawa, H., Zhang, J., Zhang, P., Zhao, Y., Zheng, L., Zheng, X., Zieringer, M., Zimmer, L., Ziveri, P., Zunino, P., and Zurbrick, C.: The GEOTRACES Intermediate Data Product 2017, Chemical Geology, 493, 210-223, 10.1016/j.chemgeo.2018.05.040, 2018.

Séférian, R., Gehlen, M., Bopp, L., Resplandy, L., Orr, J., Marti, O., Dunne, J., Christian, J., Doney, S., Ilyina, T., Lindsay, K., Halloran, P., Heinze, C., Segschneider, J., Tjiputra, J., Aumont, O., and Romanou, A.: Inconsistent strategies to spin up models in CMIP5: implications for ocean biogeochemical model performance assessment, Geoscientific Model Development, 9, 1827-1851, 10.5194/gmd-9-1827-2016, 2016.

Siegel, D., Buesseler, K., Behrenfeld, M., Benitez-Nelson, C., Boss, E., Brzezinski, M., Burd, A., Carlson, C., D'Asaro, E., Doney, S., Perry, M., Stanley, R., and Steinberg, D.: Prediction of the Export and Fate of Global Ocean Net Primary Production: The EXPORTS Science Plan,

Frontiers in Marine Science, 3, 10.3389/fmars.2016.00022, 2016.

Steinacher, M., Joos, F., Frolicher, T., Bopp, L., Cadule, P., Cocco, V., Doney, S., Gehlen, M., Lindsay, K., Moore, J., Schneider, B., and Segschneider, J.: Projected 21st century decrease in marine productivity: a multi-model analysis, Biogeosciences, 7, 979-1005, 10.5194/bg-7-979-2010, 2010.

Stramski, D., Reynolds, R., Babin, M., Kaczmarek, S., Lewis, M., Rottgers, R., Sciandra, A., Stramska, M., Twardowski, M., Franz, B., and Claustre, H.: Relationships between the surface concentration of particulate organic carbon and optical properties in the eastern South Pacific and eastern Atlantic Oceans, Biogeosciences, 5, 171-201, 10.5194/bg-5-171-2008, 2008.

Strous, M., Heijnen, J., Kuenen, J., and Jetten, M.: The sequencing batch reactor as a powerful tool for the study of slowly growing anaerobic ammonium-oxidizing microorganisms, Applied Microbiology and Biotechnology, 50, 589-596, 10.1007/s002530051340, 1998.

Swart, N., Cole, J., Kharin, V., Lazare, M., Scinocca, J., Gillett, N., Anstey, J., Arora, V., Christian, J., Hanna, S., Jiao, Y., Lee, W., Majaess, F., Saenko, O., Seiler, C., Seinen, C., Shao, A., Sigmond, M., Solheim, L., von Salzen, K., Yang, D., and Winter, B.: The Canadian Earth System Model version 5 (CanESM5.0.3), Geoscientific Model Development, 12, 4823-4873, 10.5194/gmd-12-4823-2019, 2019a.

Swart, N. C., Cole, J., Kharin, S., Lazare, M., Scinocca, J., Gillett, N., Anstey, J., Arora, V., Christian, J., Hanna, S., Jiao, Y., Lee, W., Majaess, F., Saenko, O., Seiler, C., Seinen, C., Shao, A., Solheim, L., von Salzen, K., Yang, D., and Winter, B.: The Canadian Earth System Model (CanESM), v5.0.3., https://doi.org/10.5281/zenodo.3251114, 2019b.

Takeda, S., Chai, F., and Nishioka, J., Eds.: Report of Working Group 22 on Iron Supply and its Impact on Biogeochemistry and Ecosystems in the North Pacific Ocean. PICES Sci. Rep. No. 42, 60 pp., 2013.

Taylor, K.: Summarizing multiple aspects of model performance in a single diagram., Journal of Geophysical Research-Atmospheres, 106, 7183-7192, 10.1029/2000JD900719, 2001.

Tesdal, J., Christian, J., Monahan, A., and von Salzen, K.: Evaluation of diverse approaches for estimating sea-surface DMS concentration and air-sea exchange at global scale, Environmental Chemistry, 13, 390-412, 10.1071/EN14255, 2016.

von Salzen, K., Scinocca, J. F., McFarlane, N. A., Li, J. N., Cole, J. N. S., Plummer, D., Verseghy, D., Reader, M. C., Ma, X. Y., Lazare, M., and Solheim, L.: The Canadian Fourth Generation Atmospheric Global Climate Model (CanAM4). Part I: Representation of Physical Processes, Atmosphere-Ocean, 51, 104-125, 10.1080/07055900.2012.755610, 2013.

White, J., Zhang, X., Welling, L., Roman, M., And Dam, H.: Latitudinal gradients in zooplankton biomass in the tropical Pacific At 140°W During The JGOFS Eqpac study - effects of El Nino, Deep-Sea Research Part II-Topical Studies in Oceanography, 42, 715-733,

10.1016/0967-0645(95)00033-M, 1995.

Wolf-Gladrow, D., Zeebe, R., Klaas, C., Kortzinger, A., and Dickson, A.: Total alkalinity: The explicit conservative expression and its application to biogeochemical processes, Marine Chemistry, 106, 287-300, 10.1016/j.marchem.2007.01.006, 2007.

Zahariev, K., Christian, J., and Denman, K.: Preindustrial, historical, and fertilization simulations using a global ocean carbon model with new parameterizations of iron limitation, calcification, and N_2 fixation, Progress in Oceanography, 77, 56-82, 10.1016/j.pocean.2008.01.007, 2008.

Zweng, M.M., Reagan, J.R., Antonov, J.I., Locarnini, R.A., Mishonov, A.V., Boyer, T.P., Garcia, H.E., Baranova, O.K., Johnson, D.R., Seidov, D., and Biddle, M.M.: World Ocean Atlas 2013, Volume 2: Salinity. S. Levitus, Ed.; A. Mishonov, Technical Ed., NOAA Atlas NESDIS 74, 39 pp., 2013.

Zweng, M.M., Reagan, J.R., Seidov, D., Boyer, T.P., Locarnini, R.A., Garcia, H. E., Mishonov, A.V., Baranova, O.K., Weathers, K., Paver, C.R., and Smolyar, I.: World Ocean Atlas 2018, Volume 2: Salinity. A. Mishonov Technical Ed., NOAA Atlas NESDIS 82, 50pp., 2018.

Table 1 – Ecosystem model parameters.

Symbol	Description	Unit	
T_{ref}	Reference temperature	K	298.15
E_{ap}	Activation energy for photosynthesis	kJ mol^{-1}	37.4
Q^N_{mins}	Small phytoplankton minimum N quota	g N g C^{-1}	0.04
Q^N_{maxs}	Small phytoplankton maximum N quota	g N g C^{-1}	0.172
Q^N_{minl}	Large phytoplankton minimum N quota	g N g C^{-1}	0.04
Q^N_{maxl}	Large phytoplankton maximum N quota	g N g C^{-1}	0.172
Q^{Fe}_{mins}	Small phytoplankton minimum Fe quota	µg Fe g C^{-1}	4.65
Q^{Fe}_{maxs}	Small phytoplankton maximum Fe quota	µg Fe g C^{-1}	93.
Q^{Fe}_{minl}	Large phytoplankton minimum Fe quota	µg Fe g C^{-1}	6.5
Q^{Fe}_{maxl}	Large phytoplankton maximum Fe quota	µg Fe g C^{-1}	70.
V^N_{ref}	Reference rate of N uptake	g N g C^{-1} d^{-1}	0.6
V^{Fe}_{ref}	Reference rate of Fe uptake	µg Fe g C^{-1} d^{-1}	79.
P^C_{ref}	Reference rate of photosynthesis	g C g C^{-1} d^{-1}	3
k_{XU}	Rate coefficient for exhudation	d^{-1}	1.7
k_{dgr}	Rate coefficient for chlorophyll degradation	d^{-1}	0.02
ζ	Respiratory cost of biosynthesis	g C g N^{-1}	2
α_{chl}	Initial slope of P-E curve	g C g CHL^{-1} h^{-1} (µmol m^{-2} s^{-1})$^{-1}$	1.08
Θ_{max}^N	Maximum chlorophyll-nitrogen ratio	g g^{-1}	0.18
K_{NiS}	Half-saturation for small phytoplankton nitrate uptake	mmol^{-1} m^3	0.1
K_{NaS}	Half-saturation for small phytoplankton ammonium uptake	mmol^{-1} m^3	0.05
K_{FeS}	Half-saturation for small phytoplankton iron uptake	nmol^{-1} m^3	100
K_{NiL}	Half-saturation for large phytoplankton nitrate uptake	mmol^{-1} m^3	1.0
K_{NaL}	Half-saturation for large phytoplankton ammonium uptake	mmol^{-1} m^3	0.05
K_{FeL}	Half-saturation for large phytoplankton iron uptake	nmol^{-1} m^3	200
m_{1S}	Small phytoplankton/zooplankton mortality rate (linear)	d^{-1}	0.05
m_{2S}	Small phytoplankton/zooplankton mortality coefficient	(mmol C m^{-3})$^{-1}$ d^{-1}	0.06
m_{1L}	Large phytoplankton/zooplankton mortality rate (linear)	d^{-1}	0.1

m_{2L}	Large phytoplankton/zooplankton mortality coefficient	(mmol C m^{-3})$^{-1}$ d^{-1}	0.06
X_{minp}	Minimum phytoplankton concentration for linear mortality	mmol C m^{-3}	0.01
a_L	Large zooplankton grazing parameter	(mmol C m^{-3})$^{-1}$	0.25
G_{L0}	Large zooplankton maximum grazing rate	d^{-1}	0.85
a_S	Small zooplankton grazing parameter	(mmol C m^{-3})$^{-1}$	0.25
G_{S0}	Small zooplankton maximum grazing rate	d^{-1}	1.7
λ	Assimilation efficiency	n.d.	0.8
r_{zs}	Microzooplankton specific respiration rate at T_{ref}	d^{-1}	0.3
r_{zl}	Mesozooplankton specific respiration rate at T_{ref}	d^{-1}	0.1
r_1	Small detritus remineralization rate at T_{ref}	d^{-1}	0.25
r_2	Large detritus remineralization rate at T_{ref}	d^{-1}	0.25
E_{ar}	Activation energy for detritus remineralization	kJ mol^{-1}	54.0
w_s	Small detritus sinking speed	m d^{-1}	2.
w_l	Large detritus sinking speed	m d^{-1}	30.
w_{Ca}	CaCO$_3$ sinking speed	m d^{-1}	20.
P_{Ca}	CaCO$_3$ production as fraction of mortality	mol CaCO$_3$ molC^{-1}	0.05
k_{Ca}	CaCO$_3$ dissolution rate	d^{-1}	0.0074
S_{Fe1}	Dissolved iron scavenging loss rate (Fe≤L$_{Fe}$)	d^{-1}	0.001
S_{Fe2}	Dissolved iron scavenging loss rate (Fe>L$_{Fe}$)	d^{-1}	2.5
L_{Fe}	Ligand concentration	nmol Fe m^{-3}	600.
P_{Fe}	POC-dependence parameter for Fe scavenging	(mmolC m^{-3})$^{-1}$	0.66
~~k_{NH4ox}~~k_{NH4ox}	Nitrification rate ~~in darkness~~constant	d^{-1}	0.05
K_E	Half-saturation for irradiance inhibition of nitrification	W m^{-2}	1.
k_{dnf}	Light and nutrient saturated rate of N$_2$ fixation at 30°C	mmol m^{-3} d^{-1}	0.0225
a	Initial slope for irradiance-dependence of N$_2$ fixation	(W m^{-2})$^{-1}$	0.02
K_{Fe}	Half-saturation for Fe dependence of N$_2$ fixation	nmol m^{-3}	100.
K_{NO3}	Half-saturation for DIN inhibition of N$_2$ fixation	mmol m^{-3}	0.1
O_{mxd}	O$_2$ concentration threshold for denitrification	mmol m^{-3}	6.
A_f	Anammox fraction of N loss to denitrification	n.d.	0.25

Figure 1 - Schematic of the CanOE biology model. Model currencies including chlorophyll (Chl) are indicated by coloured boxes except oxygen (O_2) and carbonate ($CaCO_3$). Arrows indicate flows of carbon (C), nitrogen (N) and iron (Fe) between compartments containing small (S) and large (L) phytoplankton (P), zooplankton (Z), and detritus (D) components; counterflows of oxygen are not shown.

Figure 2 - Global distribution of oxygen (O_2) concentration in mmol m^{-3} at 400, 900, and 1400 m (rows) for CanESM5-CanOE, CanESM5, the mean for other (non-CanESM) CMIP6 models, and World Ocean Atlas ~~2013~~2018 (WOA~~2013~~2018) observations (columns). Difference from the observation-based fields are shown in Supplementary Figure S3.

Figure 3 - Latitude-depth distribution (surface to 1750 m) of zonal mean oxygen concentration (O_2), oxygen concentration at saturation (O_2(sat)), and apparent oxygen utilization (AOU) in mmol m^{-3} for CanESM5-CanOE, CanESM5, the mean for other CMIP6 models, and observations (WOA~~2013~~2018). Note different colour scales for different rows. Difference from the observation-based fields are shown in Supplementary Figure S3.

Figure 4 - Taylor diagrams (Taylor, 2001) comparing modelled and observed distributions of oxygen at specific depths from 100 to 3500 m. Angle from the vertical indicates spatial pattern correlation. Distance from the origin indicates ratio of standard deviation in modelled vs. observed (WOA~~2013~~2018) fields. Red dots represent CanESM5-CanOE, blue dots CanESM5, small grey dots other CMIP6 models, and large grey dots the model ensemble mean for all CMIP6 models except CanESM5 and CanESM5-CanOE.

Figure 5 - Total volume of ocean with oxygen (O_2) concentration less than (a) 6 mmol m^{-3} (mean for last 30 years of the historical experiment) and (b) 60 mmol m^{-3}. Observation are from WOA~~2013~~2018.

Figure 6 - Global distribution of aragonite saturation (Ω_A) at 400, 900, and 3500 m for CanESM5-CanOE, CanESM5, the mean for other CMIP6 models, and observations (GLODAP~~v2~~ + WOA~~2013~~2018). Note different colour scales for different depths. Difference from the observation-based fields are shown in Supplementary Figure S3.

Figure 7 - Latitude-depth distribution of zonal mean (surface to 1150 m) aragonite saturation state (Ω_A), calcite saturation state (Ω_C), and carbonate ion concentration ($[CO_3^{--}]$) in mmol m^{-3} for CanESM5-CanOE, CanESM5, the mean for other CMIP6 models, and observations (GLODAP~~v2~~ + WOA~~2013~~2018). Difference from the observation-based fields are shown in Supplementary Figure S3.

Figure 8 - Taylor diagrams comparing modelled and observed distributions of DIC at specific depths from 100 to 3500 m. Observations are from GLODAP~~v2~~ (Lauvset et al., 2016). Red dots represent CanESM5-CanOE, blue dots CanESM5, magenta dots CanESM2, small grey dots other CMIP6 models, and large grey dots the model ensemble mean for all CMIP6 models except CanESM5 and CanESM5-CanOE.

Figure 9 - Taylor diagrams comparing modelled and observed (GLODAP~~v2~~ + WOA~~2013~~2018) distributions of Ω_A at specific depths from 100 to 3500 m. Symbol colours as in Figure 8.

Figure 10 - Climatological seasonal cycle of surface nitrate concentration averaged for selected ocean regions. Thick red line represents CanESM5-CanOE, thick blue line CanESM5, thick black line observations (WOA~~2013~~2018), thin grey lines individual CMIP6 models, and thick grey line the model ensemble mean (excluding CanESM5 and CanESM5-CanOE). Regional boundaries are given in Supplementary Table ~~S4~~S5 and Supplementary Figure S5.

Figure 11 - Climatological seasonal cycle of zonal mean surface nitrate concentration for a selection of CMIP6 models, a model ensemble mean (MEM) excluding CanESM5 and CanESM5-CanOE, and an observation-based data product (WOA~~2013~~2018). An alternate version showing only ~~for (a) all latitudes and (b)~~ latitudes <20° is given in Supplementary Figure S6.

Figure 12 - Global distribution of dissolved iron (dFe) concentration (log10 of concentration in nmol m^{-3}) at the ocean surface for CanESM5-CanOE and other CMIP6 models that published this field. Concentrations exceeding 1000 nmol m^{-3} are masked white. CanESM5 is not included because it does not have prognostic iron.

Figure 13 - Global mean depth profiles of dissolved iron concentration for CanESM5-CanOE and other CMIP6 models that published this field. GFDL-CM4 is excluded because it has very high concentrations (>2000 nmol m^{-3}) near the surface. Thick red line represents CanESM5-CanOE, thin grey lines individual CMIP6 models, and the thick grey line the model ensemble mean (excluding CanESM5-CanOE and GFDL-CM4).

Figure 14 - Mean surface nitrate (NO_3) vs. dissolved iron (dFe) concentrations in different oceans, including the major high nutrient / low chlorophyll (HNLC) regions. CanESM5-CanOE is shown as a red dot and other CMIP5 models as grey dots (CanESM5 is not included because it does not have iron). Observed NO_3 is shown as a vertical black line as there are no observational estimates of dFe concentration. For GFDL-CM4, nitrate is estimated as phosphate x 16. Region definitions are given in Supplementary Table ~~S4~~S5 and Supplementary Figure S5.

Figure 15 - Surface nitrate (NO_3) concentrations along the Pacific equator (mean from 2°S-2°N) during the upwelling season (June-October) for CanESM5-CanOE (red), CanESM5 (blue), and WOA~~2013~~2018 observations (black).

Figure 16 - Annual mean surface ocean concentration of large and small phytoplankton and zooplankton in CanESM5-CanOE (red) and of phytoplankton and zooplankton in CanESM5 (blue) for the representative ocean regions shown in Figure 14. Observational estimates (black) are for phytoplankton biomass calculated from satellite ocean colour estimates of surface chlorophyll (SeaWiFS/MODIS; Tesdal et al. 2016), assuming a carbon-to-chlorophyll ratio of 50 g/g. Region definitions are given in Supplementary Table S5~~4~~ and Supplementary Figure S5.

Figure 17 - Mean annual cycle of surface chlorophyll for the representative ocean regions shown in Figures 14 and 16. CanESM5-CanOE large and small phytoplankton concentrations are shown separately and combined (red) along with CanESM5 (blue) and observational estimates (black). Region definitions are shown in Supplementary Table S5~~4~~ and Supplementary Figure S5.

Figure 18 - Climatological surface particulate organic carbon (POC) vs. chlorophyll for CanESM5-CanOE (red) and observations (black). Data are for all ocean grid points (2x2° uniform global grid) for all months of the year where observational data are available. Model POC is offset 17 mg m^{-3} for illustrative purposes. Observed chlorophyll concentrations >1 mg m-3 are excluded as they largely represent coastal areas poorly resolved by coarse resolution global ocean models.

Figure 19 - (a) Global total export production (epc100) in PgC y^{-1} (b) and zonal mean export production in molC m^{-2} y^{-1} according to selected CMIP6 models (mean for 1985-2014 of historical experiment). Thick red line represents CanESM5-CanOE, thick blue line CanESM5, thin grey lines individual CMIP6 models, and thick grey line the model ensemble mean (excluding CanESM5 and CanESM5-CanOE).

Figure 20 - Cumulative ocean uptake of carbon dioxide (CO_2) as anthropogenic dissolved inorganic carbon (AnthDIC) in PgC over the course of the historical experiment (1850-2014). Data are shown as successive five-year means. CMIP6 mean (thick grey line) indicates ensemble mean for CMIP6 models (thin grey lines) excluding CanESM5 (blue) and CanESM5-CanOE (red). An observation-based estimate of 145±20 PgC (Friedlingstein et al., 2020) is shown for nominal year 2014 (black).

Figure 21 - Change in export production (epc100) over the course of the historical experiment (1850-2014), normalized to the 1850-1900 mean. Data are shown as successive five-year means. Thick red line represents CanESM5-CanOE, thick blue line CanESM5, thin grey lines other CMIP6 models, and thick grey line the ensemble mean of non-CanESM models.

Figure 22 - (a) Change in total ocean volume with oxygen (O_2) concentration less than (a) 6 mmol m^{-3} and (b) 60 mmol m^{-3} over the course of the historical experiment (1850-2014), normalized to the 1850-1870 mean. Data are shown as successive five-year means. Thick red line represents CanESM5-CanOE, thick blue line CanESM5, and thin grey lines other CMIP6 models..

Dissolved iron model comparison to observations

As there is no gridded global data product for dissolved iron (dFe) we present here some comparisons with individual bottle samples from the GEOTRACES Intermediate Data Product 2017 (Schlitzer et al., 2018), the MBARI data compilation (www3.mbari.org/chemsensor/Data/), and the Pacific data set compiled by PICES WG22 (meetings.pices.int/members/working-groups/disbanded/wg22). The MBARI data include both profile (the "Global Iron Data") and surface transect ("MBARI SOLAS") data (the SOLAS data also include some profiles but are primarily underway surface measurements). The GEOTRACES data are primarily profile data (393 profiles total, only ~10% of 7519 data are <50 m). Most of the MBARI data are also in the PICES data set; these were preprocessed to remove redundant data. Data from the underway surface transects were excluded to avoid overweighting these regions due to autocorrelation, leaving a total of 3575 data points. Concentrations > 2.5 nM, which comprise 1.3% of GEOTRACES data and 3.3% of PICES/MBARI data, were excluded from the model/data comparisons below, except for the individual GA-02 profiles.

Evaluating the CanOE model beyond what has been discussed in the main text identified several key points, most of which have already been made to some degree. CanESM5-CanOE compares favourably with other CMIP6 models but is quite biased towards a 'nutrient-type' rather than 'scavenged-type' profile (Figures S4S9b, c, e-hg). In CanOE rates of Fe scavenging are very high above 0.6 nM dFe and very low below this concentration, resulting in an almost constant deep water concentration of 0.6 nM. The collected data (of which there are >10000 vs only a few hundred in 1997) show that the basic hypothesis articulated by Johnson et al. (1997) still holds: deep-water concentrations are generally close to 0.6 nM, although much higher concentrations are sometimes observed at depths of thousands of metres (Figure S4S9b, see e.g. Resing et al., 2014). High concentrations are also sometimes observed in near-surface waters (Figure S4S9b) and CanESM5-CanOE is among the models least able to reproduce these. CanESM5-CanOE consistently overestimates the lowest observed concentrations and underestimates the higher ones (Figure S4S9d). One might argue that other models do little better in a statistical sense. However, in several models the deviations from the 1:1 line are less systematic than in CanESM5-CanOE; the residuals may be larger but they are more homogeneous. This comparison also largely confirms the results of Séférian et al. (2020, their Figure 5), although only two of the four models shown here were included in their analysis. CNRM-ESM2-1 shows a capacity to simulate the full range of dFe concentrations, as was the case in Séférian et al., although in no case is the spatial pattern correlation very large (in Séférian et al., CNRM-ESM2-1 was the second highest at 0.21). Nonetheless, this analysis indicates that Therefore purely statistical comparisons as in Figures S4S9e-g can be misleading. CanESM5-CanOE shows fairly good skill by these metrics but has systematic biases. These biases can clearly be attributed to the rather simplistic scavenging model employed, and will be addressed in future versions.

We made direct comparisons of our modelled dissolved iron concentrations with GEOTRACES transect data for GA-02 in the Atlantic (e.g., Middag et al. 2015), which was the most spatially extensive transect available (Figure S9a). We show depth profiles from 47°S to 47°N. Mostly this confirms what we already knew from the other analyses presented: our model has a very low scavenging rate below 0.6 nM and a very high rate above, so that deep water concentrations are quite uniform and near-surface concentrations are biased low in high-deposition regions like the northern tropical Atlantic. For the most part, the model reproduces the observed concentrations quite well, given these known biases. What we learn from including this additional analysis is (a) the model is biased high in the Antarctic Bottom Water (Middag et al. 2015 give an excellent presentation of the location and biogeochemical properties of this water mass), and (b) the seasonal biological drawdown in the mid-latitude North Atlantic is weak. The former is probably due mainly to the low scavenging rate at

concentrations <0.6 nM, although it may also indicate a high bias in surface waters of the source region (see section 3.3). The latter is probably related to the generally low rate of export production (Figure 19) and the weak North Atlantic spring/summer bloom (Figures 16 and 17). Model annual mean data are shown, but seasonal minima are not much less, compared to the strong summer drawdown in the observations.

Figure ~~S4~~S9a – Global distribution of dissolved iron (dFe) measurements in the GEOTRACES (black/green) and PICES/MBARI (blue) data compilations. GEOTRACES Pacific data north of 25ºS are indicated in green. Blue dots outside the Pacific are MBARI data. Large symbols indicate stations along GA-02 that are compared in Figure S9h.

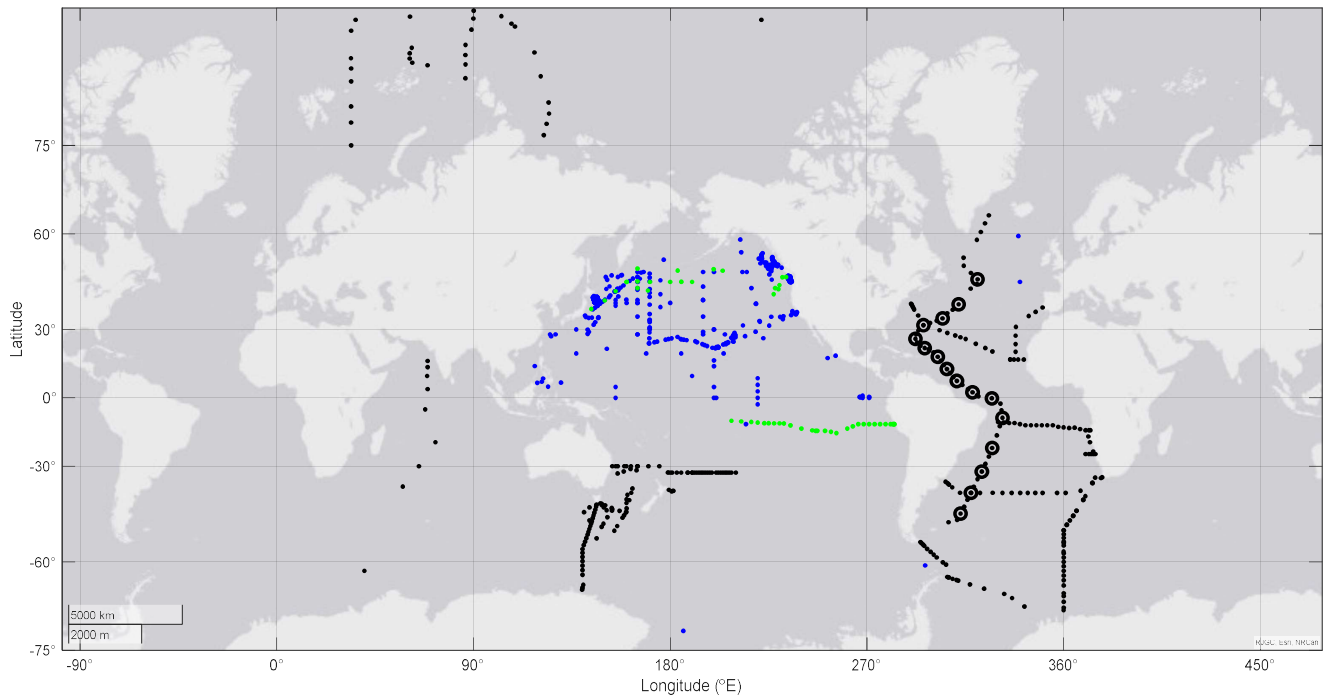



Figure S4S9b - All observations of dissolved iron (dFe) concentration (<2 nM) plotted against sampling depth. GEOTRACES data are in black and PICES/MBARI data are blue. Model data are global mean profiles of annual mean data as in Figure 13 in the main text.

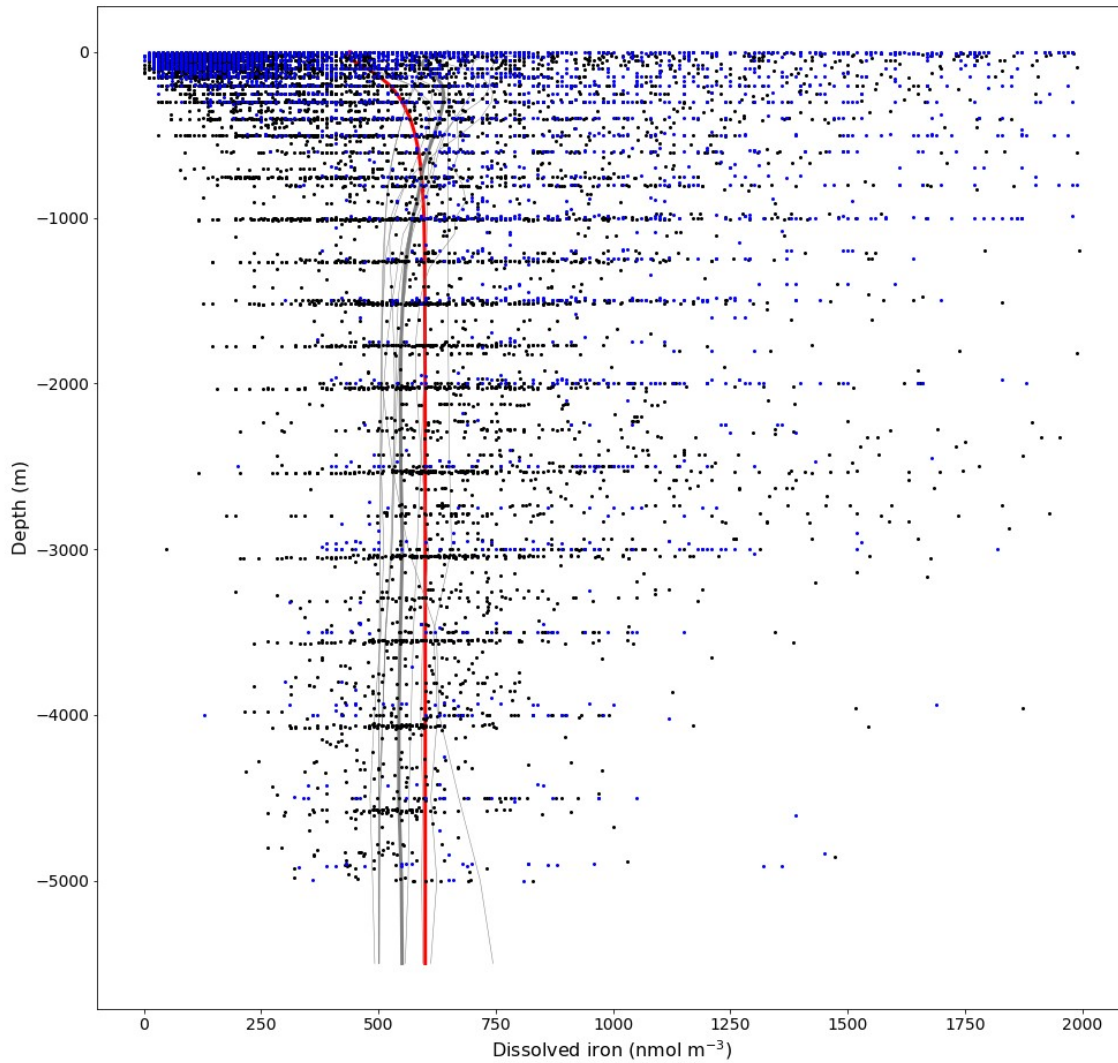



Figure S4S9c - Mean of observations of dissolved iron (dFe) concentration (<2 nM) for various depth strata, including both GEOTRACES and PICES/MBARI data. Means of all data available in that depth range; spatial coverage is very incomplete and inconsistent among the depth strata. Averaging layers are 50 m (0-1000 m), 100 m (1000-3000 m) or 300 m (>3000 m) thick. Model data are global mean profiles of annual mean data as in Figure 13 in the main text.

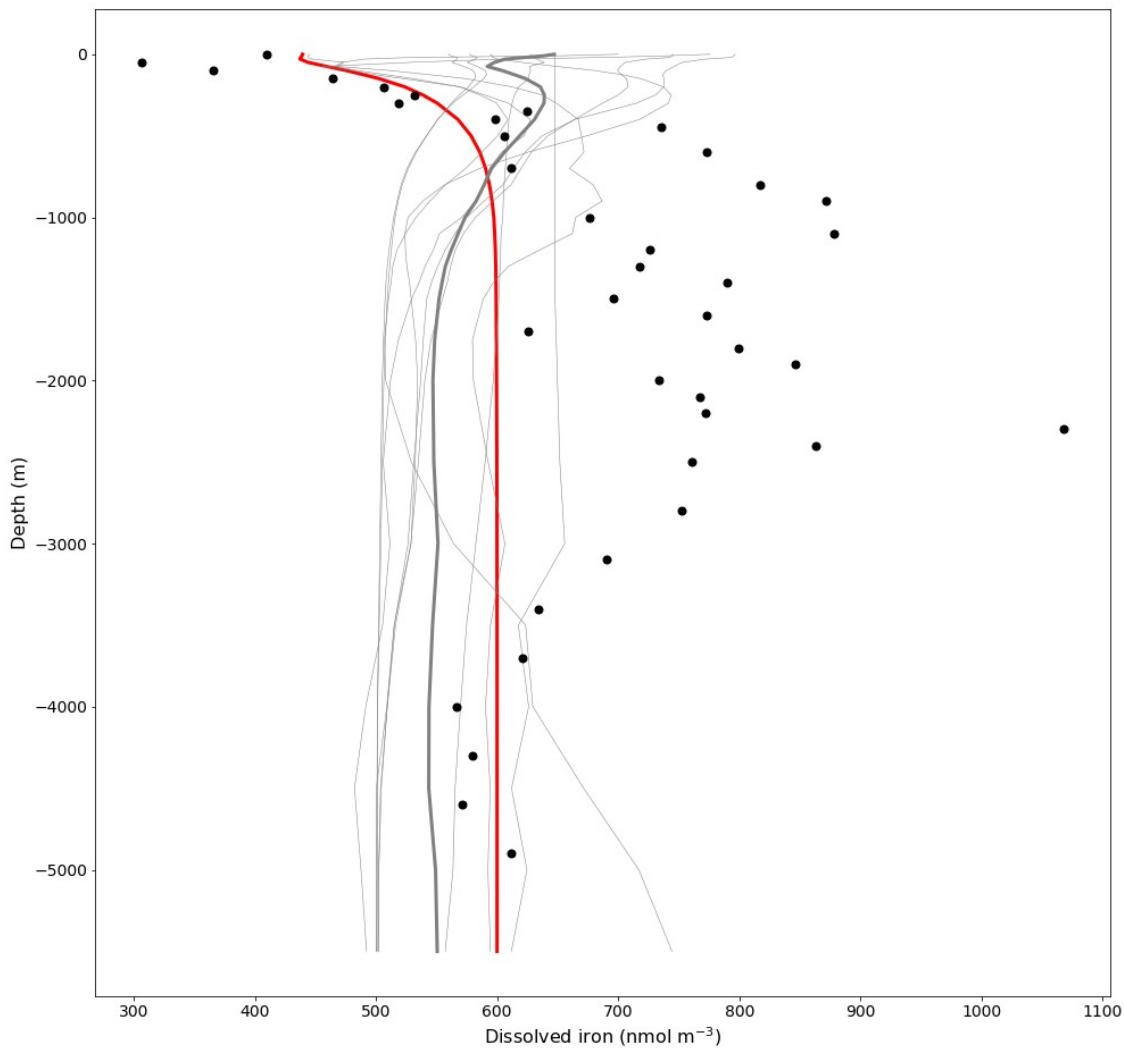



Figure S4S9d - Modelled and observed mean dissolved iron (dFe) concentrations (excluding those >2.5 nM) in the upper 50 m at locations where a depth profile was taken (N=1122). Observed data are means of all measurements made at depths <50 m within a given profile. Model data are climatological surface values for the month in which the observed data were collected (red: CanESM5-CanOE; black CNRM-ESM2-1; blue: GFDL-CM4; magenta: MPI-ESM1-2-LR). Selection of models is somewhat arbitrary but includes the models that show the highest overall skill according to the metrics shown in Figures S4S9e-g. Thin lines are linear regressions for individual models; thick black line is 1:1.

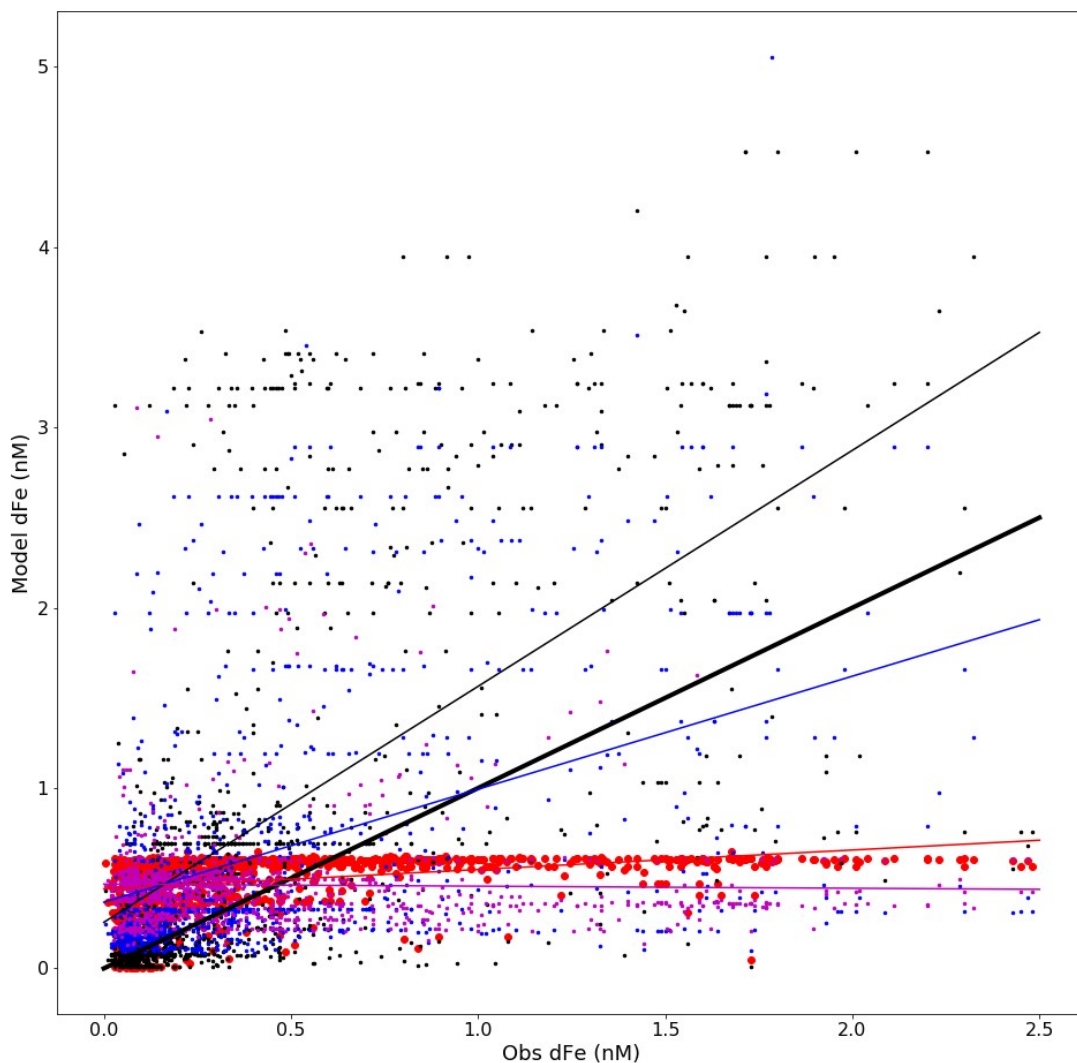



Figure S4S9e - Root mean square error and correlation coefficient for surface dissolved iron data (upper 50 m mean with maximum of 2.5 nM, as in Figure S4S9d) for CMIP6 models for which seasonal data were available.

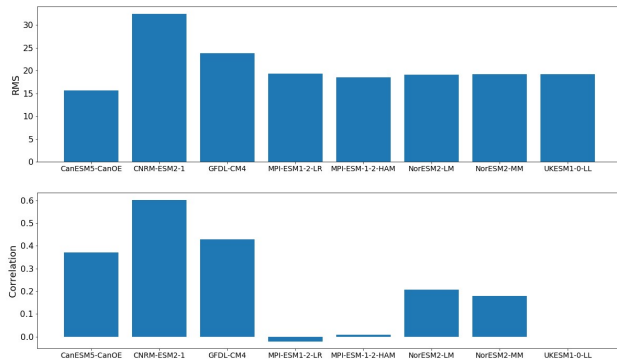



Figure S4S9f - As S4S9e but for the Pacific only (north of 25ºS).

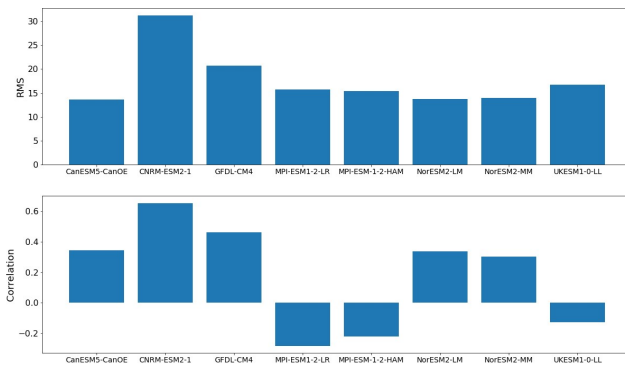
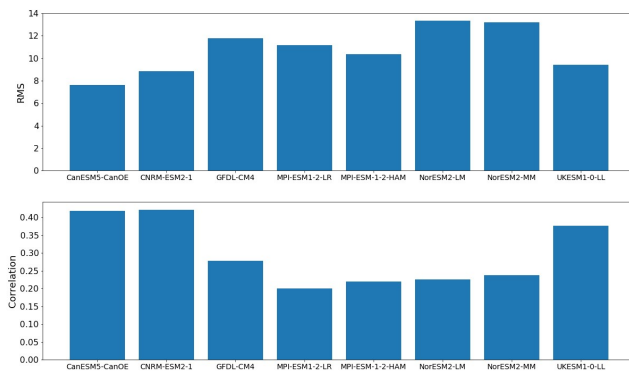
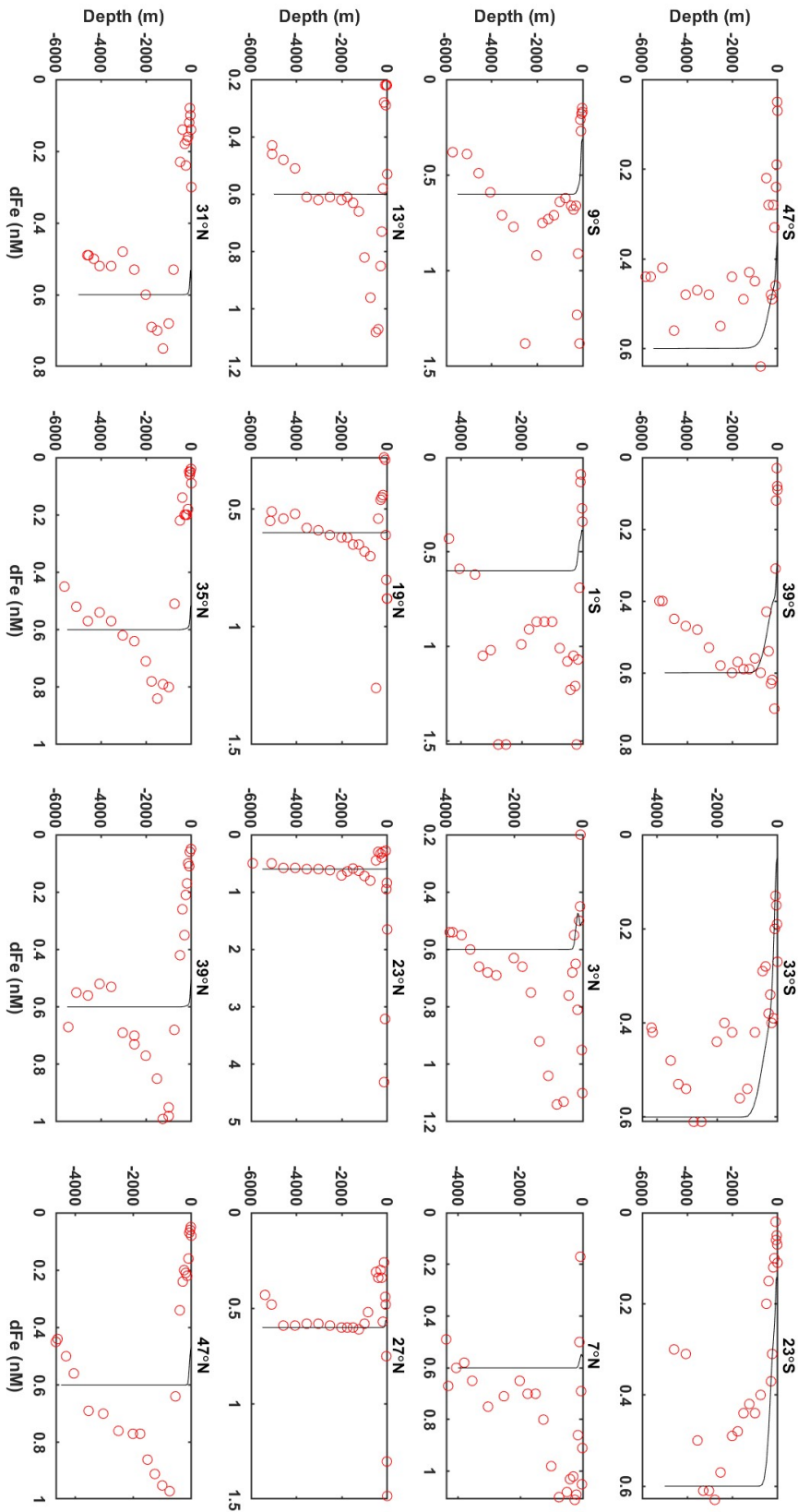



Figure S4S9g - As S4S9e but excluding the Pacific north of 25ºS.

References

Middag, R., M.M.P. van Hulten, H.M. Van Aken, M.J.A. Rijkenberg, L.J.A. Gerringa, P. Laan, and H.J.W. de Baar, 2015. Dissolved aluminium in the ocean conveyor of theWest Atlantic Ocean: Effects of the biological cycle, scavenging, sediment resuspension and hydrography. Marine Chemistry 177: 69–86.

Resing, J.A., P.N. Sedwick, C.R. German, W.J. Jenkins, J.W. Moffett, B.M. Sohst and A. Tagliabue, 2015. Basin-scale transport of hydrothermal dissolved metals across the South Pacific Ocean. Nature 523: 200–203.

Séférian, R., et al., 2020. Tracking improvement in simulated marine biogeochemistry between CMIP5 and CMIP6, Current Climate Change Reports, 6: 95-119.

Schlitzer, R., et al., 2018. The GEOTRACES Intermediate Data Product 2017. Chemical Geology 493: 210-223.

The data file for the PICES WG22 data compilation requests that data users cite the original publications. As all of the data are used except for the SOLAS underway data, all of the original references are listed here:

Boyle, E.A., B.A. Bergquist, R.A. Kayser and N. Mahowald, 2005. Iron, manganese and lead at Hawaii Ocean Time-Series station ALOHA: Temporal variability and an intermediate water hydrothermal plume. Geochimca et Cosmochimica Acta, 69: 933-952.

Brown, M. T., W. M. Landing, and C. I. Measures, 2005. Dissolved and particulate Fe in the western and central North Pacific: Results from the 2002 IOC cruise. Geochem, Geophys. Geosyst., 6, Q10001, doi: 10.1029/2004GC000893.

Bruland, K., Orians, K. and Cowen, P. ,1994. Reactive trace metals in the stratified central North Pacific. Geochem. Cosmochim. Acta. 58, 3171-3182.

Coale, K. H., Fitzwater, S. E., Gordon, R. M., Johnson, K. S. and Barber, R. T., 1996a.   Control of community growth and export production by upwelled iron the equatorial Pacific Ocean. Nature, 379:621-624.

Ezoe M , Ishita T , Kinugasa M , Lai X, Norisuye K, Sohrin Y, 2004, Distributions of dissolved and acid-dissolvable bioactive trace metals in the North Pacific Ocean. Geochemical Journal, 38(6), 535-550.

Fitzwater, S.E., K.H. Coale, R.M. Gordon, K.S. Johnson, M.E. Ondrusek, 1996, Iron deficiency and phytoplankton growth in the equatorial Pacific, Deep-Sea Res. II, 43, 995-1015

Fujishima, Y., K. Ueda, M. Maruo, E. Nakayama, C.Tokutome, H.Hasegawa M. Matsui and Y. Sohrin. Distribution of Trace Bioelements in the Subarctic North Pacific Ocean and the Bering Sea (the R/V Hakuhou Maru Cruise KH-97-2), Journal of Oceanography, Vol. 57, pp.261 to 273, 2001

Gordon, R. M., Johnson, K. S. and Coale, K. H., 1998. The behavior of iron and other trace elements during the IronEx I and PlumEx experiments in the equatorial Pacific. Deep-Sea Res. II 45: 995-1041.

Gordon, R.M. , J.H. Martin, G.A. Knauer, 1982, Iron in north-east Pacific waters, Nature, 299, 611-612

Johnson - MLML World Iron Data Set: Appendix A to Johnson, K. S., Gordon, R. M. and Coale, K. H., 1997, What controls dissolved iron concentrations in the world ocean? Mar. Chem. 57: 137-161.

Johnson, K.S., Elrod, V., Fitzwater, S., Plant, J.N., Chavez, F.P., Tanner, S.J., Gordon, M., Westphal, D.L., Perry, K.D., Wu, J., and Karl, D.M., 2003. Surface ocean-lower atmosphere interactions in the

Northeast Pacific Ocean Gyre: Aerosols, iron, and the ecosystem response, GBC(17), 10.1029/2002GB002004

Johnson, K.S., F.P. Chavez, V.A. Elrod, S.E. Fitzwater, J.T. Pennington, K.R. Buck, and P.M. Walz, 2001. The annual cycle of iron and the biological response in central California coastal waters. Geophys. Res. Letters 28: 1247-1250.

Johnson, WK, Miller, LA, Sutherland, ND, Wong, CS, 2005. Iron transport by mesoscale Haida eddies in the Gulf of Alaska, Deep-Sea Res. II, 52: 933-953.

Kitayama, S., K. Kuma, E. Manabe, K. Sugie, H. Takata, Y. Isoda,K.Toda,S.Saitoh,S.Takagi,Y.Kamei,and K.Sakaoka,2009.Controls on iron distributions in the deep water column of the North Paciffic Ocean:iron(?) hydroxide solubility and humictype fluorescent dissolved organic matter.J.Geophys.Res.,114:C08019,doi:10.1029/2008JC004754

Kondo, Y., 2007. Dynamics of the organic Fe complexing ligands and phytoplankton in the Pacific Ocean.The University of Tokyo. Ph.D.Thesis, 256pp.

Kondo, Y., S. Takeda and K. Furuya, 2012. Distinct trends in dissolved Fe speciation between shallow and deep waters in the Pacific Ocean. Marine Chemistry, 134–135, 18–28

Kuma, K, J. Nishioka, K. Matsunaga, 1996, Controls on iron (III) hydroxide solubility in seawater: The influence of pH and natural organic chelators, Limnol. Oceanogr., 41, 396-407

Kuma, K., A. Katsumoto, H. Kawakami, F. Takatori, K. Matsunaga,1998, Spatial variability of Fe(III) hydroxide solubility in the water column of the northern North Pacific Ocean, Deep-Sea Res. I, 45, 91-113

Kuma, K., Isoda, Y. and Nakabayashi, S., 2003, Control on dissolved iron concentrations in deep waters in the western North Pacific: Iron(III) hydroxide solubility. J. Geophys. Res., 108 (C9): 3289, doi:10.1029/2002JC001481 (2003).

Landing, W.M., K.W. Bruland, 1987, The contrasting biogeochemistry of iron and manganese in the Pacific Ocean, Geochim. et Cosmo. Acta, 51, 29-43

Mackey, D.J., J.E. O'Sullivan, R.J. Watson, 2002, Iron in the western Pacific: A riverine or hydrothermal source for iron in the Equatorial Undercurrnet?, Deep-Sea Res. I, 49, 877-893

Martin, J.H. and R.M. Gordon, 1988, Northeast Pacific iron distributions in relation to phytoplankton productivity, Deep-Sea Res. I, 35, 177-196

Martin, J.H., R.M. Gordon, S. Fitzwater, W.W. Broenkow, 1989, VERTEX: phytoplankton/iron studies in the Gulf of Alaska, Deep-Sea Res., 36, 649-680

Nakabayashi, S., K. Kuma, K. Sasaoka, S. Saitoh, M. Mochizuki,, N. Shiga, and M. Kusakabe, 2002. Variation in iron(III) solubility and iron concentration in the northwestern North Pacific Ocean. Limnol. Oceanogr. 47: 885-892

Nakabayashi, S., M. Kusakabe, K. Kuma, and I. Kudo, 2001. Vertical distributions of Iron(III) hydroxide solubility and dissolved iron in the northwestern North Pacific Ocean. Geophys. Res. Letters 28: 4611-4614.

Nishioka, J., S. Takeda, I. Kudo, D. Tsumune, T. Yoshimura, K. Kuma, A. Tsuda, 2003, Size-fractionated iron distributions and iron-limitation processes in the subarctic NW Pacific, GRL, 30, NO. 14, 1730, doi:10.1029/2002GL016853.

Nishioka, J., T. Ono, H. Saito, K. Sakaoka, and T. Yoshimura (2011),Oceanic iron supply mechanisms which support the spring diatom bloom in the Oyasio region,western subarctic Pacific,J.Geophys.Res.,116,C02021,doi:10.1029/2010JC006321

Nishioka, J., T. Ono, H. Saito, T. Nakatsuka, S. Takeda, T. Yoshimura, K. Suzuki, K. Kuma, S. Nakabayashi, D. Tsumune, H. Mitsudera, W. K. Johnson, A. Tsuda, (2007) Iron input into the western subarctic Pacific, importance of iron export from the Sea of Okhotsk, Journal of Geophysical Research, 112, C10012, doi:10.1029/2006JC004055.

Nishioka, J., Takeda, S., Wong, C., Johnson, W. 2001. Size-fractionated iron concentrations in the northeast Pacific Ocean: Distribution of soluble and small colloidal iron. Mar. Chem. 74, 157-179

Obata H., 1997. Development of an automated analytical method of iron in seawater and studies on the behavior of iron in the ocean.Kyoto University.Ph.D.Thesis,109pp.

Obata, H., H. Karatani and E. Nakayama, 1993. Automated determination of iron in seawater by chelating resin concentration and chemiluminescence detection.Analytical Chemistry,65:1524-1528

Obata, H., H. Karatani, M. Matsui and E. Nakayama, 1997. Fundamental studies for chemical speciation of iron in seawater with an improved analytical method.Marine Chemistry,56:97-106

Roy, E., 2009. The detection and biogeochemistry of trace metals in natural waters. Ph.D. Thesis, University of Maine.

Rue, E.L. and K.W. Bruland, 1995. Complexation of iron(III) by natural organic ligands in the central North Pacific as determined by a new competitive ligand equilibration/adsorptive cathodic stripping voltammetric method. Mar. Chem. 50: 117-138.

Takata, H, Kuma, K, Saitoh, Y, Chikira, M, Saitoh, S, Isoda, Y, Takagi, S, Sakaoka, K, 2006. Comparing the vertical distribution of iron in the eastern and western North Pacific Ocean. Geophys. Res. Lett., 33, L02613, doi:10.1029/2005GL024538, 2006.

Takata, H., K. Kuma, S. Iwade, Y. Yamajyoh, A. Yamaguchi, S. Takagi, K. Sakaoka, Y. Yamashita, E. Tanoue, T. Midorikawa, K. Kimura, J. Nishioka, 2004, Spatial variability of iron in the surface water of the northwestern North Pacific Ocean, Mar. Chem., 86, 139-157

Takeda, S., H. Obata, 1995, Response of equatorial Pacific phytoplankton to subnanomolar Fe enrichment, Mar. Chem, 50, 219-227

Tsuda, A., S. Takeda, H. Saito, J. Nishioka, Y. Nojiri, I. Kudo and others, 2003, A mesoscale iron enrichment in the western Subarctic Pacific induces a large centric diatom bloom, Science, 300, 958-961

Wu, J., E. Boyle, W. Sunda, and L.-S. Wen, 2001. Soluble and colloidal iron in the oligotrophic North Atlantic and North Pacific. Science 293: 847-849.