Perturbed biology and physics signatures in a 1-D ocean biogeochemical model ensemble

Article

Published Version

Open Access

It is advisable to refer to the publisher's version if you intend to cite from the work. See [Guidance on citing](#).

www.reading.ac.uk/centaur

Perturbed Biology and Physics signatures in a 1-D ocean biogeochemical model ensemble

Prima Anugerahanti,[1,*]**, Shovonlal Roy,**[1*] **and Keith Haines,**[2]

[1]*Department of Geography and Environmental Science, University of Reading, Whiteknights, Reading, RG6 6AB,UK*

[2]*Department of Meteorology and National Centre for Earth Observation, University of Reading, Whiteknights campus Early Gate, Reading, RG6 6BB, UK*

Correspondence*:
Prima Anugerahanti
p.anugerahanti@pgr.reading.ac.uk

Shovonlal Roy
shovonlal.roy@reading.ac.uk

2 ABSTRACT

Sources of uncertainty in a marine biogeochemical model include input from physical processes and the choice of functional forms representing the strength and dependencies of biogeochemical processes. This study explores characteristic signatures from these uncertainties by generating ensembles from perturbing the biogeochemistry equations and perturbing physical input using a 1-D intermediately-complex model run at five oceanographic stations. Perturbed biogeochemistry ensemble (PBE) produces larger spreads than perturbed physics ensemble (PPE), and distinctly different ensemble variations. Fractions of nitrogen in phytoplankton pool from observations show a larger variability than in any single model-ensemble member, but the PBE spread generally captures this variability, whereas the PPE spread does not. The results show that the PBE method gives a more realistic representation of uncertainty than PPE in our 1D-model setup. Our method needs to be tested in more complex models in order to understand its significance on larger scales.

Keywords: Perturbed biogeochemistry ensemble, Ocean biogeochemical model, Ensemble modelling, Structural uncertainty, perturbed physics ensemble

1 INTRODUCTION

Ocean biogeochemical (OBGC) models have been developed to understand how the ocean ecosystem responds to the changes in both the physics and the biogeochemistry (Doney et al., 2012; Yool et al., 2013; Butenschon et al., 2016). Key uncertainties that affect OBGC models include physical processes, with vertical mixing and upwelling of nutrients often poorly known (Doney, 1999; Sinha et al., 2010; Friedrichs et al., 2006), and the various choices for formulating the biological processes such as nutrient uptake, zooplankton grazing, and plankton mortality (Gentleman et al., 2003; Anderson et al., 2010; Adamson and Morozov, 2013). These biological processes are described by functional forms relating them to concentrations of plankton and nutrients, as well as ambient temperature and light availability. Different physical environments can strongly affect simulations of chlorophyll distribution through the water column (Friedrichs et al., 2006), as well as regional distributions of phytoplankton functional types at

27 the surface ocean (Sinha et al., 2010). Spurious vertical velocities that can occur when assimilating physical
28 ocean data into models can also raise nutrient concentrations in the upper water column (Subramanian and
29 Palmer, 2017). Furthermore, when using different physical models, anthropogenic CO_2 uptake can vary
30 between 25%-30% (Doney et al., 2004). The structure of an OBGC model, especially the choice of the
31 functional representation of biogeochemical processes, strongly determine the model dynamics (Edwards
32 and Yool, 2000; Fussmann and Blasius, 2005). For example, when the grazing function alone is altered
33 from hyperbolic to sigmoidal (both of which are common in the literature) three times higher phytoplankton
34 concentrations can been produced (Anderson et al., 2010). Impacts of altering mortality are shown in both
35 uncoupled Nutrient-Phytoplankton-Zooplankton (NPZ) models (Steele and Henderson, 1992; Edwards and
36 Yool, 2000) and coupled OBGC models (Yool et al., 2011). Choosing a linear mortality, can double the
37 diatom biomass at high latitudes, compared to using other functions (Yool et al., 2011). So the uncertainties
38 arising from both physical and biogeochemical formulations may contribute to discrepancies between the
39 models and observations (Anderson, 2010; Allen et al., 2010).

40   One way of accounting for these multiple sources of uncertainty is to move away from deterministic
41 simulations towards ensemble results which can be designed to deliver a probability distribution of
42 outcomes. Perturbed physics ensembles have, for example been used to estimate the uncertainties of climate
43 projections (Tinker et al., 2015; Subramanian and Palmer, 2017) or to forecast the climate probabilistically
44 (Tebaldi and Knutti, 2007; Murphy et al., 2007). Ensembles are also regularly used to quantify uncertainties
45 in data assimilation applications (Anderson, 2001; Moradkhani and Meskele, 2010; Roy et al., 2012) to
46 allow weighting of model results compared with new observations.

47   Recently, Anugerahanti et al. (2018) has introduced an approach for generating an ensemble of an an
48 OBGC model by perturbing its core biogeochemistry processes. Here we extend the study of Anugerahanti
49 et al. (2018), to decouple and compare the variability that may arise in an intermediately complex 1-D
50 OBGC model from both biology and physics uncertainties, by generating three sets of ensembles perturbing:
51 (i) the biogeochemistry, by altering the choice of functional forms (perturbed biogeochemistry ensemble,
52 PBE), (ii) the physics, by adding noise to the vertical velocity, mixed layer depth (MLD) and therefore
53 the vertical diffusivity coefficient, supplying nutrients to the surface layers (perturbed physics ensemble,
54 PPE), and (iii) both the biogeochemistry and physics together (perturbed biogeochemistry and physics
55 ensemble, PBPE). Since the OBGC model behaviour varies across different biogeographical provinces
56 (Kriest et al., 2012), the ensemble is run at five monitored ocean sites ranging from coastal to oligotrophic
57 regions. We quantify the variability generated by the perturbed ensembles, identifying and distinguishing
58 the characteristics from the different biological and physical perturbations based on several biogeochemical
59 property metrics. From these characteristics we can explore how the different perturbations may affect the
60 model dynamics.

61   This paper is organised as follows: Brief description of the 1-D OBGC model, generating the ensembles,
62 and the description of metrics are explained in section 2. The basic diagnostics of the ensembles which relate
63 to the bulk properties of the model states, followed by the effect of perturbations in vertical distribution of
64 chlorophyll are discussed in section 3.1. The different characteristic signatures of the PBE and PPE are
65 described and discussed in section 3.2. Finally the conclusions of the study are in section 4.

2 METHODS

66 We use the Model of Ecosystem Dynamics, nutrient Utilisation, Sequestration, and Acidification (MEDUSA
67 1.0) (Yool et al., 2011). MEDUSA is an intermediately complex biogeochemical model that has two

68 phytoplankton types (diatoms and non-diatoms), two zooplankton types (mesozooplankton and microzoo-
69 plankton), and three nutrients (dissolved inorganic nitrogen, silica, and iron), and uses nitrogen as the model
70 currency. The 1-D version of this model is run in the Marine Model Optimisation Testbed (MarMOT-1.1)
71 (Hemmings et al., 2015). The physical forcings, such as vertical velocity and solar radiation, are taken
72 from the NEMO-FOAM output (Storkey et al., 2010), with output frequency every 5-days for all of the
73 stations. NEMO-FOAM is a data assimilation product and therefore biases in well observed quantities are
74 small, however for temperature and mixed layer depth (MLD) we introduce an additional bias correction to
75 match the mean seasonal physical conditions observed at the stations. The vertical diffusivity coefficient
76 is matched to the bias corrected MLD. Bias correction is done for all of the stations apart from station
77 PAP where observational data are insufficient, so at PAP we use unadjusted NEMO-FOAM output. The
78 MEDUSA ensembles are run from 1 January 1998 to 31 December 2007, with output produced everyday,
79 at five different oceanographic stations; oligotrophic (represented by stations BATS (32°50'N, 64°10'W)
80 and ALOHA (22°45'N, 158°00'W)), coastal (represented by stations Cariaco (10°30'N, 64°40'W) and L4
81 (50°15'N, 4°12.3'W)), and abyssal plain (represented by station PAP (49°N, 16.5°W)). Further information
82 about running MEDUSA and a map of the station locations can be found in the Supplementary section 1.
83
84 ## 2.1 Generating the ensembles
86 We generate the PBE by altering the equivalent functional forms for key biogeochemical processes.
87 In the previous study (Anugerahanti et al., 2018) we used all possible functional form combinations,
88 generally used in literature to describe four key processes; nutrient uptake, phytoplankton and zooplankton
89 mortalities, and zooplankton grazing. The functional forms for phytoplankton nutrient uptake are Monod
90 (U_h), which is the default function, exponential (U_e), sigmoidal (U_s), and trigonometric (U_t). For plankton
91 mortalities, the default function is hyperbolic (denoted ζ_h for zooplankton and ρ_h for phytoplankton). Other
92 functions available in MEDUSA are: linear (ζ_l, ρ_l), quadratic (ζ_q, ρ_q), and sigmoidal (ζ_s, ρ_s). Finally, for
93 zooplankton grazing, we use Holling type III (G_1), which is the default function, and Holling type II (G_2).
94 The shape defining parameters for these functional forms are tuned to each other so that over a wide range
95 of conditions the key processes remains similar, see (Anugerahanti et al., 2018). Rate maxima are also
96 similar to the original MEDUSA-1.0 run, apart from linear and quadratic mortalities, as these functions
97 have no shape defining parameters. These process formulations with respective alternative functions made
98 128 combinations, which was the size of the original ensemble reported in Anugerahanti et al. (2018). But
99 to reduce the computational cost while keeping the ensemble properties mostly unchanged, here we we
100 limit the biogeochemical ensemble to 12 members chosen using principal component analysis (PCA) and
101 k-means cluster, to span a similar range of variability for measurable metrics of chlorophyll and nutrients
102 as the larger ensemble (see Supplementary section 2, for further details).

103 At each of the stations the PPE is generated by adding "noise" to the vertical velocity, temperature, MLD
104 and vertical diffusivity, in a regionally dependent and covarying way (as these fields are related) in order to
105 increase variability, see supplementary section 3 for details. The vertical diffusivity profile is then matched
106 to the perturbed MLD. The perturbations for vertical velocity at all stations are done by first subtracting the
107 monthly average vertical velocity. The anomalies are multiplied by a random number between -2 and 2 and
108 added to each five-day average field. These anomalies are generated randomly for each ensemble member.
109 For station PAP, the perturbations to MLD are similar to perturbing the vertical velocity, and the vertical
110 diffusivity profile is matched with the perturbed MLD. Further explanation of the PPE generation is in
111 the Supplementary section 3, Figures S4 and S5. We use a PPE ensemble size of 12 members to match

112   the PBE ensembles discussed above. Finally the combination of perturbing physics and biogeochemistry
113   together is generated by running the PBE using the physical inputs from the PPE, to produce a PBPE.
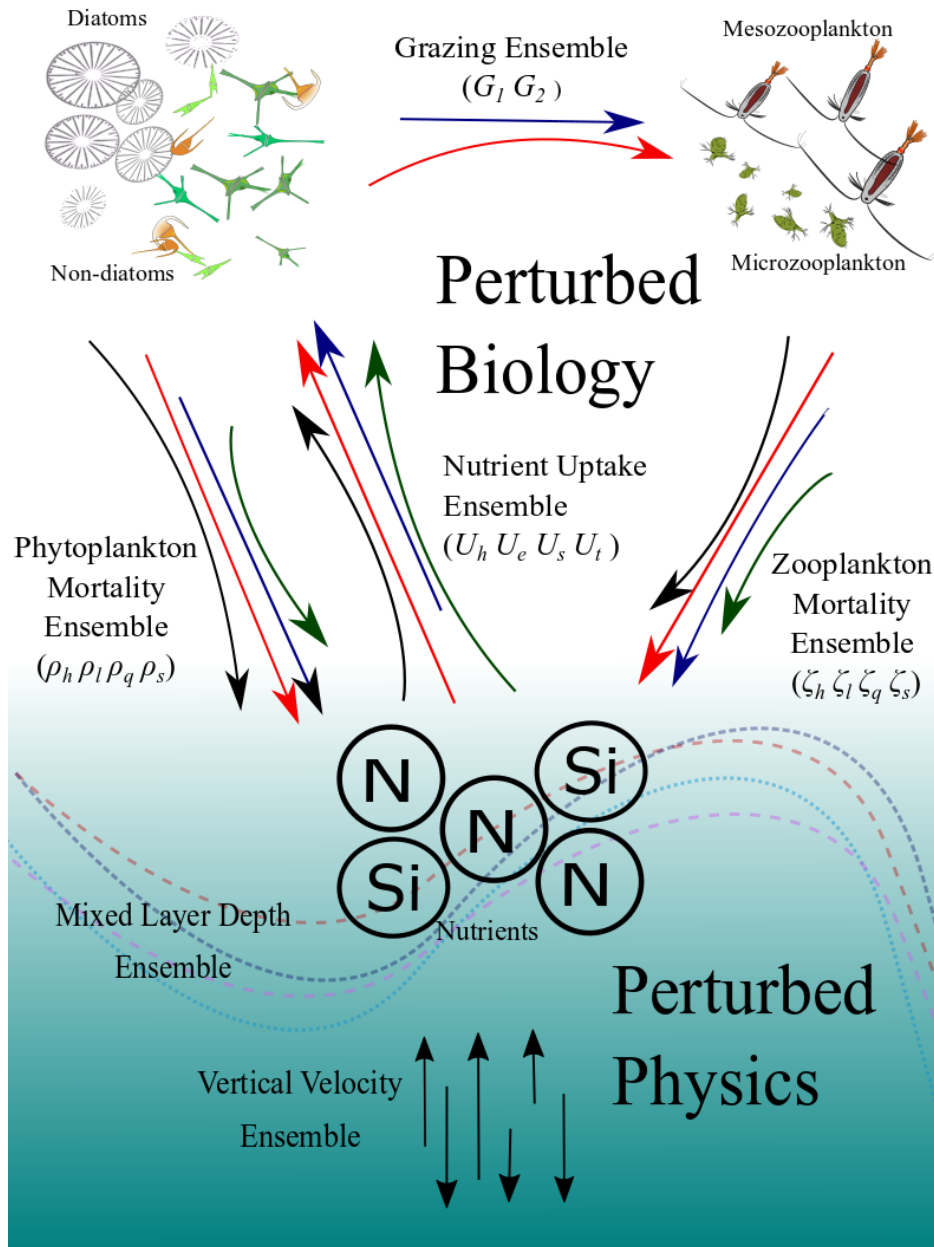

Figure 1. Schematic diagram showing how the ensembles are generated. The coloured and curved arrows in the top part represent the different functional forms which describe the key biogeochemical processes which generate the PBE. The straight vertical arrows at the bottom represent varying vertical velocities and the curved lines represent climatology of mixed-layer depths which generate the PPE. The PBPE is the combination of the two.

2.2   Ensemble metrics

115   We are interested in key properties of the model ensembles which we use to compare with observations
116   at the five oceanographic stations. The spread of the annual means of dissolved inorganic nitrogen (DIN

117 mmol m^{-3}), chlorophyll (mg m^{-3}), and zooplankton (mmol m^{-3}) concentrations are the basic diagnostics
118 throughout the water column. At the oligotrophic stations a deep chlorophyll maximum (DCM) is a common
119 feature that occurs below the mixed layer when surface chlorophyll concentration is low (Fennel and Boss,
120 2003; Letelier et al., 2004). The DCM evolution is explored phenologically by its maximum depth and
121 concentration over the winter (December-January-February), spring (March-April-May), summer (June-
122 July-August), and fall (September-October-November). The range of DCM depth, timing of maximum
123 depth, and concentration are examined for both the PPE, PBE, and the observational data.

124 We also examine the fractions of total nitrogen in the phytoplankton pool to reveal a signature of the
125 processes which have been varied within the ensembles, in particular this distinguishes PPE from PBE
126 induced variations. This fraction is calculated by using the chlorophyll to nitrogen ratios, taken from Yool
127 et al. (2011), for both the in situ and model ensemble. This metric can give an indication of the processes
128 involved in the temporal changes seen from the in situ observations, suggesting it may be possible to infer
129 which processes (physical, biological, or both) may be responsible for model-observational discrepancies
130 at different times.

3 RESULTS

3.1 Chlorophyll Range and Distributions

132 Perturbations to the vertical velocity and MLD used for the PPE, produce relatively little spread in the
133 bulk properties, especially for phytoplankton and zooplankton Figure 2. The PPE DIN concentrations vary
134 little in the top 75m (in all stations except PAP), however the DIN range increases below, suggesting that
135 the physical variations impact more below the euphotic layers. These deeper variations however do not
136 have much impact on bulk properties near the surface such as the total DIN, chlorophyll (phytoplankton) or
137 zooplankton concentrations (as seen in Figure 2). At the oligotrophic stations, the PPE range is clearly
138 insufficient to cover the in situ concentrations. However at all five stations, from surface to deep water, the
139 observed chlorophyll values mostly lie within the much larger PBE range (Figure 2a-d), suggesting that the
140 full range of biological production through a strong nutrient gradient can be obtained by perturbing the
141 biological processes. Only at the oligotrophic stations, below ∼100m, are in situ chlorophyll concentrations
142 still outside the PBE range. The combined PBPE ensemble has a slightly wider range than PBE but is
143 otherwise similar.

144 The PBE and PPE members also differ in DCM generation at the oligotrophic stations. Figure 3 shows
145 chlorophyll distributions from four different members at BATS and ALOHA, (see supplementary for
146 monthly profiles of PPE section 5, Figures S8 to S11). The DCM is always present for part of each year
147 but with considerable variability in maximum chlorophyll concentration and depth. In observations the
148 deepest DCM always occurs in the summer and the shallowest in winter (Mignot et al., 2014). The range of
149 DCM depths from the PBE is larger than that from the PPE, with observed deepest DCM depths generally
150 within the PBE range (e.g. the deepest DCM depths at ALOHA, are 51-115m (PBE), 82-95m (PPE), and
151 depth=114m from observations). Similarly, for the minimum DCM depth, the PBE produces a larger
152 range, although this still underestimates that in the observations (PPE DCM range= 21-37m, PBE= 3-51m,
153 observation=92m). Additionally all PPE members have the deepest DCM later in the autumn, instead of
154 in summer, but not all PBE members show this discrepancy. There are some differences in chlorophyll
155 distributions between PPE members and the default run, especially the thickness of the chlorophyll layer
156 during winter/spring at BATS, although differences are not as distinct as for the PBE, as seen in Figure 3.
157 The PBPE follows the pattern and timings of PBE, although the DCM depth range is slightly wider (e.g. at
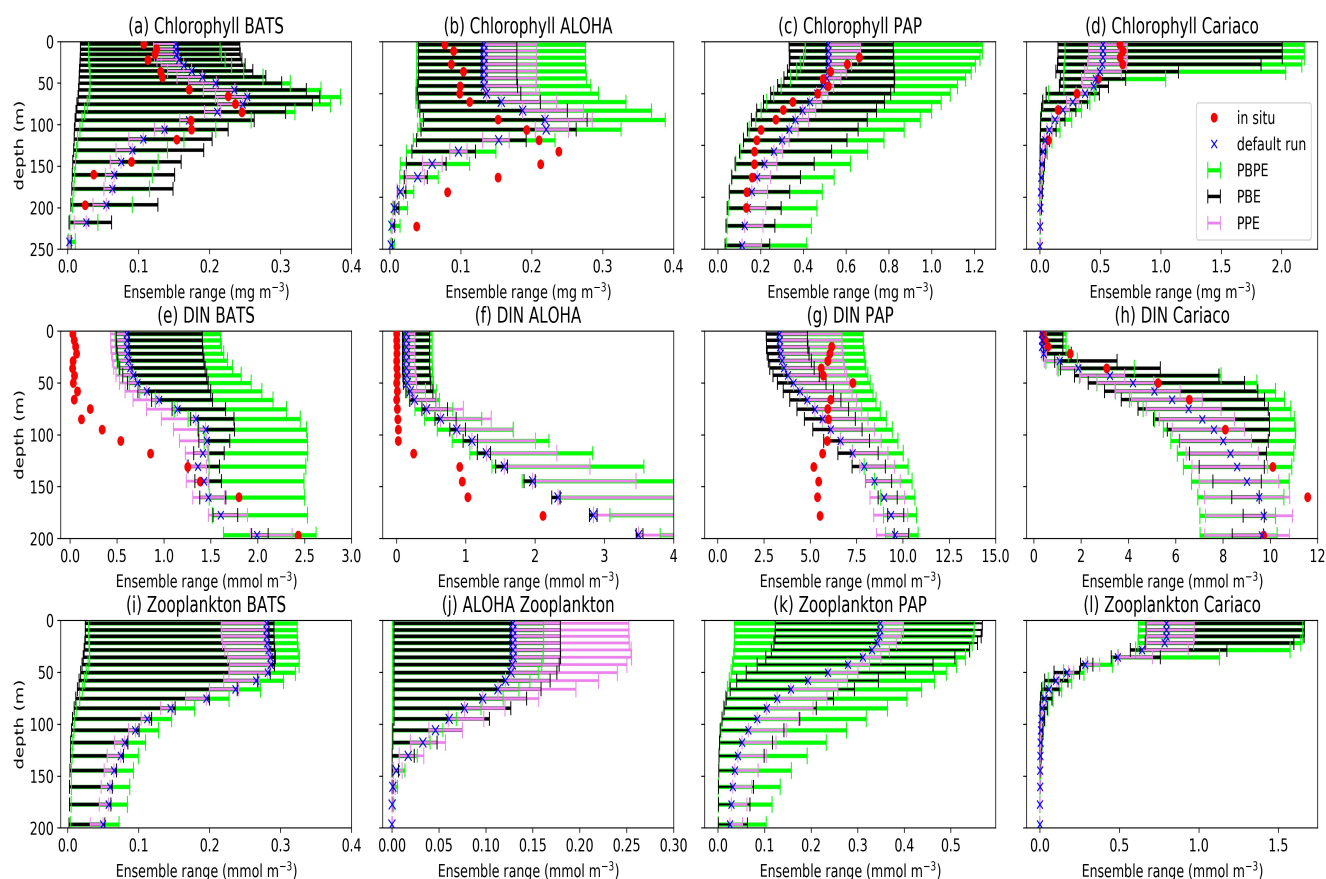158 ALOHA, PBPE range 69-118m for maximum DCM depth).

Figure 2. Ensemble range of mean chlorophyll (a to d), DIN (e to h), and zooplankton (i to l) profiles calculated from 1 January 1998- 31 December 2007 at BATS (a, e, and i), ALOHA (b, f, and j), PAP (c, g, and k), and Cariaco (d, h, and l). Blue crosses show the mean concentrations from the default run, red dots show the mean concentration from in situ, the violet bars denote the mean concentration from PPE, the green bars show the mean concentrations from the PBPE, and the black bars show the mean concentrations from the PBE. For station PAP the annual mean is taken between 2002 to 2004 for DIN and between 2003 to 2005 for chlorophyll (see Supplementary section 4, Figures S6 and S7, for the in situ monthly averages for DIN and chlorophyll at PAP). The model calculations for the annual means matched with the timing of observational sampling. Station L4 profiles are not shown because in situ data are only available at the surface.

159   These results suggest that perturbing the biogeochemistry can result in considerably greater variability in
160 the evolution of the DCM, compared to perturbing the physics alone. Furthermore, when perturbing both
161 physics and biogeochemistry, the effect of perturbing the latter predominantly determines the ensemble
162 spread and chlorophyll distribution.

163
164 **3.2 Characteristics of the different ensembles**

165   The phytoplankton nitrogen fraction shows how much nitrogen resides in the phytoplankton pool, relative
166 to the total DIN and phytoplankton nitrogen. The size of the phytoplankton nitrogen fraction can also
167 indicate the concentration of nutrients (DIN) in the water column. For example, at ALOHA and BATS,
168 the observed phytoplankton nitrogen fractions are always close to 1, indicating that most of the time, this
169 region is nutrient limited. At stations such as L4, the phytoplankton nitrogen fraction can change drastically
170 over the course of a season in both the observations and the model (Figure 4c and g).
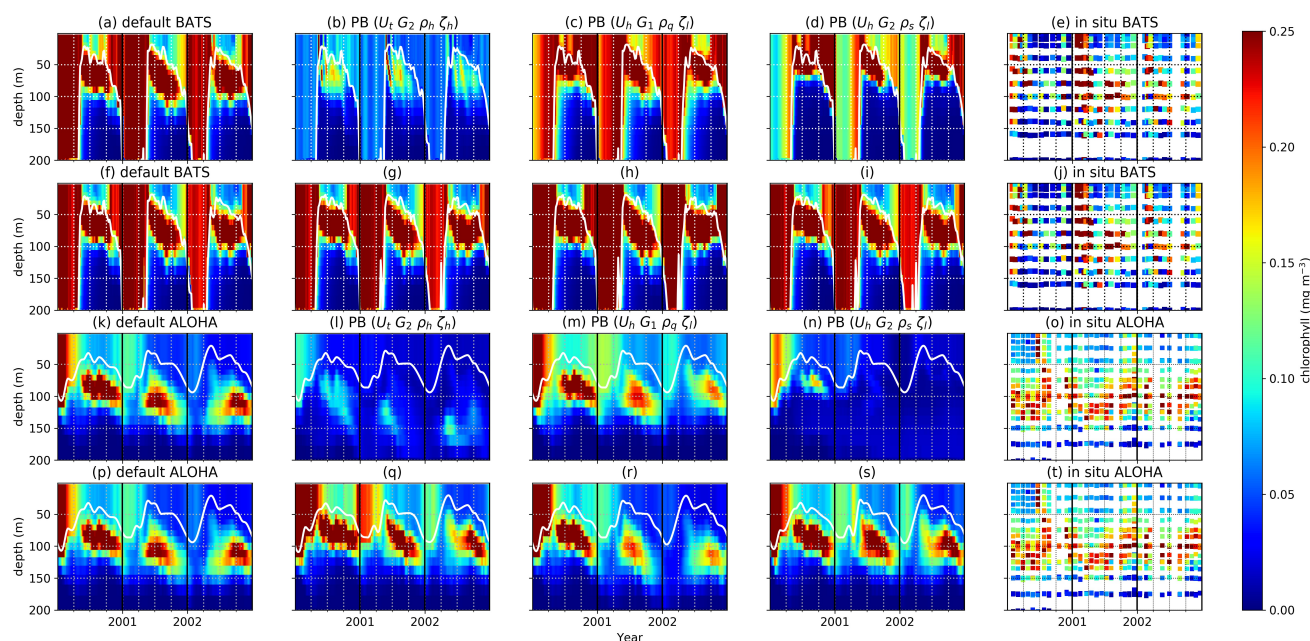
Figure 3. Chlorophyll distribution in the water column from 1st January 2000 to 31st December 2002 at station BATS (a to j) and ALOHA (k to t). White solid lines are the MLDs. Selected ensemble members, that are the most distinct from default run from PBE, with their functional form combinations are shown in (b) to (d), for BATS, and (k) to (n) for ALOHA, and for PPE are shown in (g) to (i) for BATS and (q) to (s) for ALOHA

.

171   From figure 4a and b, the proportion of nitrogen in phytoplankton is seen to vary strongly across the PBE
172   members. In contrast the PPE shows very little spread in nitrogen fractions across the whole ensemble,
173   Figure 4e and f. However, at the coastal stations L4 and Cariaco, there is more variability between PPE
174   ensemble members, Figure 4g and h, and the timing of maximum phytoplankton nitrogen fraction varies
175   across the ensemble.

176   The contrast between PBE and PPE is more distinct in the phytoplankton nitrogen fractions than in
177   the spread differences in chlorophyll, for example seen in Figure 2, where the PPE chlorophyll range is
178   seen to show more spread than the phytoplankton nitrogen fraction, especially in the oligotrophic regions.
179   The small changes in the functional representation of uptake, grazing, and mortality curves in the PBE,
180   represented by the exchange arrows in the upper part of Figure 1, can strongly alter the mean nitrogen
181   distributions because they directly alter the cycling between biological pools. In contrast the PPE variability
182   really only alters the supply of nutrients from deeper layers, represented by the lower part of Figure 1,
183   and not the fluxes between the biological compartments and biological fractional distributions, hence the
184   smaller PPE spreads in Fig 4e-h. The larger PBE spreads mostly capture the observed seasonal variations in
185   nitrogen fractions e.g. at L4, where the PPE ensemble cannot, and thus PBE provides a better representation
186   of uncertainty.

4   DISCUSSION

187   Previous studies such as Najjar et al. (2007), show that a simple biogeochemical model forced by different
188   GCMs can produce large variability in dissolved organic matter both in the surface and at depth. Another
189   study by Séférian et al. (2013) shows that atmosphere-ocean models differing in ocean subgrid physics
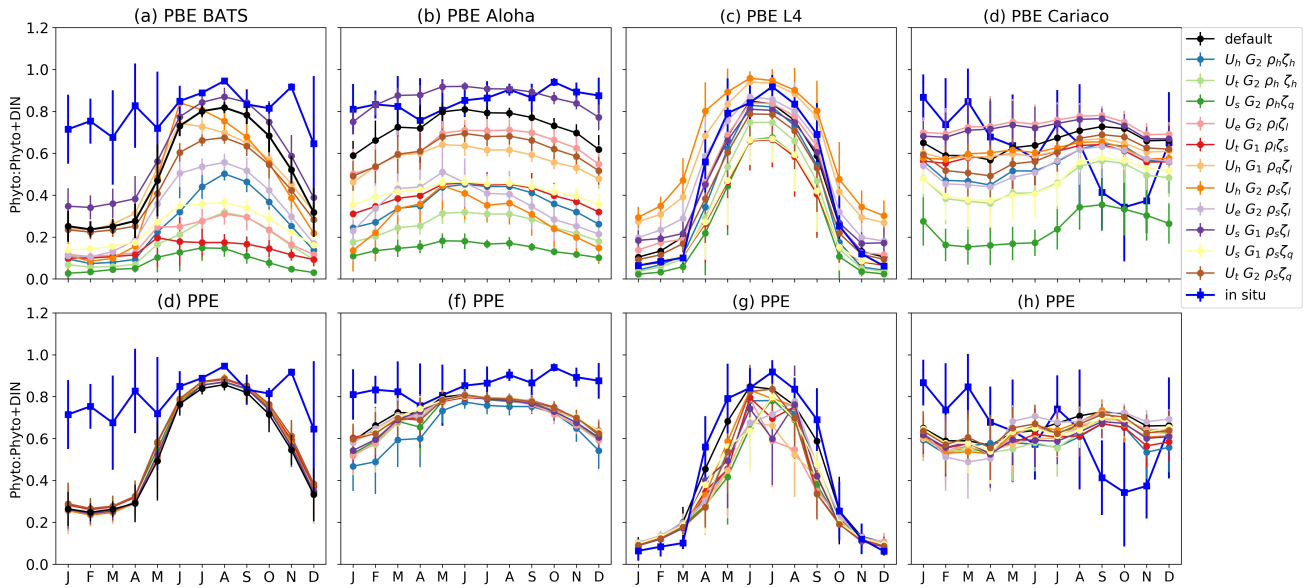
Figure 4. Monthly averaged phytoplankton fraction P/(P+D) in nitrogen units at the surface for four oceanographic stations. The top and bottom panels show the phytoplankton fractions from PBE and PPE, respectively, with different lines representing ensemble members. The observations are shown in blue. The bars are the standard deviations of the monthly P/(P+D). The nitrogen within phytoplankton is calculated using the chlorophyll to nitrogen ratio which is calculated using the C:N conversion fraction, and the calculation is described in Yool et al. (2011). These are calculated from 1 January 1998 to 31 December 2007, apart from station L4, which are calculated from January 2000, to match the in situ data.

190   and resolution can also produce varying biogeochemical tracers, such as nutrients and chlorophyll. In this
191   study, we found that the uncertainty arising from biogeochemistry processes gives a larger range, especially
192   in chlorophyll and zooplankton, as shown in Figures 2, and 4. In terms of bulk properties, the fact that a
193   PPE generates a small range, is consistent with studies where different ocean general circulation models
194   are coupled with the same OBGC model (e.g., Sinha et al. (2010)). However, below the depths of ∼75m
195   the PPE DIN shows a larger range, due to the absence of activities between nutrient phytoplankton and
196   zooplankton, and physics perturbations therefore have more effect on DIN. At PAP the larger PPE DIN
197   range, at depths of active phytoplankton growth may occur due to the restricted sampling to winter months,
198   when biological activity is low even at the surface, and the physical perturbations are the only control on
199   DIN.

200   Physically perturbing the vertical velocity, MLD, vertical diffusivity, and temperature in the PPE can alter
201   the chlorophyll distributions in the water column and the depth of DCMs because these physical variables
202   control the nutrients (vertical velocity and MLD) and light (MLD) availability (Siegel et al., 2002). The
203   variations in nutrient and light availability then alter the timing of peak phytoplankton concentrations
204   (Henson et al., 2013). Perturbing the MLD using the described method in supplementary section 3, changes
205   the magnitude of vertical diffusivity leading to an increase/decrease in nutrient concentrations at euphotic
206   depths (Huisman et al., 2006). From Figure 4 the PPE range depends on the model temperature bias; at
207   stations where the model bias is small, such as BATS and ALOHA (mean temperature bias are -0.24 and
208   -0.44, respectively), the range of phytoplankton nitrogen fraction is low, and the seasonality is similar
209   across members. However, at stations where model temperature bias is high, such as L4 and Cariaco (mean
210   temperature bias are 0.90 and -1.58, respectively), the PPE range is larger, with more variable seasonality.

211   Perturbing the biogeochemistry produces a larger range of DCM depths, as the DCM depends on nutrient
212   uptake, zooplankton grazing, and plankton mortality from surface to deep water. This makes the depth of
213   the DCM vary across all ensemble members when the grazing or mortality functions are altered (e.g Figure
214   3b and g). The DCMs occur at depths where the phytoplankton growth rate is in balance with the loss rate
215   (Fennel and Boss, 2003; Cullen, 2014). Variations in DCM depths, pattern, and continuity across the PBE
216   are therefore due to different loss and growth rates throughout euphotic depths. In oligotrophic regions,
217   the nutrient concentration is low in the top $\sim$ 150m (see Figure 2e and f). Some PBE members produce
218   higher phytoplankton loss rates compared to growth rate in the top 75m due to the nutrient scarcity (e.g.
219   members which use G_2, ρ_h, and ρ_l). At deeper depths, nutrient is plentiful allowing phytoplankton growth
220   to exceed the loss rate, giving a deeper DCM for these PBE members. When the mixed layer becomes
221   deeper, a balance cannot be achieved as light becomes a limiting factor and chlorophyll concentrations
222   reduce (see Figure3b and g). The slightly larger maximum DCM depth range in PBPE may be caused by
223   the additional net upwelling and the change in mixed layer depth from perturbing the physics, which gives
224   the maximum depth for members with more downwelling and deeper MLD, and therefore a deeper DCM.

225   PBE and PPE ranges are also shown and compared for nitrogen fractions in Figure 4, because nitrogen
226   is the model currency and we can examine its distribution to phytoplankton across different ensemble
227   members, and these variables are available from observations. Variations in phytoplankton nitrogen
228   proportions, both temporal and between the PPE members, may result from perturbing the MLD, as this can
229   also controls the timing of maximum phytoplankton concentrations, by controlling the light and nutrient
230   availability, as well as distribution of phytoplankton in the water column (Behrenfeld et al., 2013; Henson
231   et al., 2013).

232   At station BATS, only three PBE members produce a nitrogen fraction comparable to that seen in the
233   observations; the default function, $U_h G_2 \rho_s \zeta_l$, and $U_s G_1 \rho_s \zeta_l$. This is because the hyperbolic uptake function
234   has higher nutrient uptake at low nutrient concentrations, compared to other functional forms, and both
235   sigmoidal phytoplankton mortalities and linear zooplankton mortality produce lower phytoplankton loss.
236   Note that the uptake functions in the default MEDUSA and the ensemble do not permit acclimatisation in
237   nutrient uptake, such as described in Smith et al. (2009). The underestimation at the oligotrophic stations
238   may also be caused by the bias introduced when reducing the ensemble members from 128 to 12 (see
239   details in Supplementary section S2), which considers observations and model outputs at all stations
240   across different oceanographic regions, in which there are 10 other PBE members that produce higher
241   phytoplankton nitrogen fractions than the default run.

242   Apart from the possibility of inefficient uptake in the MEDUSA 1-D model, some physical parameters,
243   such as horizontal advection and eddies, are not represented at all. In the subtropical gyre 3-D advection is
244   thought to be essential in controlling primary productivity (Dave and Lozier, 2010; Palter et al., 2005),
245   which may explain the discrepancies between the in situ and ensemble phytoplankton nitrogen fractions
246   shown in Figure 4. In order to fully address the physical model bias, the impact of 3-D advection should be
247   represented, and any errors in that circulation would need to be accounted for through ensemble spread,
248   possibly by using multi-model ensembles, although even these may contain shared biases (Abramowitz
249   et al., 2019).

250   Both PBE and PPE spreads are better at capturing the nitrogen fraction at light limited stations such
251   as L4. The ensembles generally follow the observations, even when nutrients become limiting in the
252   summer, because light also controls the nutrient uptake rate. The observed phytoplankton nitrogen fraction
253   generally falls within the PBE range throughout the year, for example from October-March, the in situ
254   phytoplankton fraction generally matches ensemble members with lower phytoplankton growth rates at

255 low concentrations (such as $U_t G_2 \rho_l \zeta_s$ and $U_h G_2 \rho_h \zeta_h$), and from April to September it matches members
256 with higher phytoplankton growth rates and high zooplankton mortality (such as $U_h G_1 \rho_s \zeta_l$ and $U_h G_2 \rho_q \zeta_l$).
257 This is consistent with North Atlantic bloom studies, where the phytoplankton nitrogen proportions and
258 growth rates change over the year (Behrenfeld et al., 2013; Behrenfeld and Boss, 2014; Roy et al., 2012),
259 being controlled by nutrients, light, and mixed layer conditions. For example in the summer, the growth
260 rate of phytoplankton is in equilibrium with loss rate as nutrient is depleted and grazing rates are high
261 (Behrenfeld et al., 2013; Behrenfeld and Boss, 2014).

262   These results suggest that in a 1-D biogeochemical model the PBE generates enough spread to encompass
263 the uncertainty within the observed phytoplankton fraction even if the region is seasonally varying, and can
264 explain the variations of growth and loss rate in phytoplankton. We can also see that none of the single
265 PBE or PPE members fully capture the observations throughout the year, therefore using a single set of
266 functional forms is not sufficient to capture the observed behaviour and its uncertainty. The PBE ensemble
267 members that best match the in situ fractions vary through the year as the ensemble members behave
268 differently depending on the concentrations of nutrient, phytoplankton, and zooplankton, especially in
269 strongly seasonally varying regions.

270   We have further attempted to compare our PBE model with different biogeochemical model types used
271 previously in model intercomparison studies e.g, Kwiatkowski et al. (2014). Acknowledging that the PBE
272 model presented here was a 1D model, running only at 5 stations, a rigorous comparison with 3D models
273 would be difficult. However, when compared to all surface observations at five stations, the mean of PBE
274 surface chlorophyll produces a correlation of 0.55, with correlation range of [0.491, 0.583] produced by
275 model ensemble members. In the inter-comparison study, Kwiatkowski et al. (2014) reported the range of
276 [0.15, 0.50] across all models. Similarly, considering all observed surface DIN at five stations, the mean
277 of PBE produces a correlation of 0.41, with the ensemble correlation range of [0.333, 0.595], which are
278 lower than for the models reported by Kwiatkowski et al. (2014) which has surface DIN within the range
279 [0.94, 0.79]. However, generality of the results needs to be tested beyond the five stations, and through
280 comparison of other models with observations beyond annual average of surface fields.

281   When assessing the risks of climate change a structural ensemble may also be useful for representing
282 model uncertainty. It has been shown in earlier studies, eg. by (Hawkins and Sutton, 2012) using a CMIP3
283 multimodel ensemble, that the uncertainty in climate change predictions may be strongly dominated
284 by model uncertainty in the near term, and the detection time for anthropogenic impacts is conditioned
285 by these uncertainties. The PBE ensemble method clearly demonstrates the importance of structural
286 uncertainties, which should then be relevant in assessing climate change impacts on ecological indicators
287 such as phytoplankton phenology.

5   SUMMARY AND CONCLUSION

288 We have run three different ensembles using 1-D MEDUSA, generated by perturbing the biology (PBE),
289 the physics (PPE), and both together (PBPE). The ensemble spreads, chlorophyll distributions, and
290 characteristics of these ensembles are explored. The PBE and PBPE generally produce larger spread of
291 the chlorophyll annual means compared to PPE, and are able to encompass the in situ concentrations seen
292 at 5 different oceanographic stations. Below the active phytoplankton growth region, the PPE produces
293 larger DIN (nutrient) spread than PBE, as below this depth there is less biological activity and nutrient
294 supply is dependent on the PPE. For the chlorophyll distributions we used the time evolution of the DCM
295 as an ensemble metric at oligotrophic stations and this shows that across different ensemble members
296 the PBE and PBPE produce larger spreads of DCM depth compared to PPE, with different chlorophyll

297 patterns. This is because the PBE produces more variable loss and growth rates of phytoplankton with
298 different nutrient supply rates. This means that perturbing the biogeochemistry produces a stronger effect
299 than perturbing physics.

300   To see how nitrogen, the model currency, is distributed to the phytoplankton compartments, we used
301 phytoplankton nitrogen fraction as a metric. This metric shows that the PBE produces a much larger spread
302 than PPE in terms of the monthly variability, and nearly covers the in situ standard deviations, especially at
303 the strongly seasonally varying stations. The large spread from the PBE show that altering the steepness
304 of the uptake, mortality, and grazing curves changes the way nitrogen is distributed to the phytoplankton
305 compartments, while in PPE the perturbations only alter the nutrient supply, both in terms of distribution in
306 the water column and concentrations.

307   Our 1D-model experiments suggest that the PBE or PBPE better represent model uncertainties arising
308 from the model structural errors, as shown by their ensemble ranges, and how the model currency is
309 distributed between the different compartments. A 1D model does however contain many simplifications
310 when it comes to ocean physics. To understand the implications of model structural errors on larger scales,
311 this method should also be tested in 3D coupled physical-biogeochemical models

ACKNOWLEDGEMENTS

CONFLICT OF INTEREST STATEMENT

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

AUTHOR CONTRIBUTIONS

328 SR and KH conceived the study, and led the development of the methodology with PA. PA run the model,
329 performed the analysis, created visualisation and wrote the first draft of the manuscript. All authors
330 contributed significantly to writing the subsequent drafts and prepared the final version for publication.

REFERENCES

331 Abramowitz, G., Herger, N., Gutmann, E., Hammerling, D., Knutti, R., Leduc, M., et al. (2019).
332   ESD Reviews: Model dependence in multi-model climate ensembles: Weighting, sub-selection and
333   out-of-sample testing. doi:10.5194/esd-10-91-2019

334 Adamson, M. W. and Morozov, A. Y. (2013). When can we trust our model predictions? Unearthing
335   structural sensitivity in biological systems. *Proceedings of the Royal Society of London A: Mathematical,*
336   *Physical and Engineering Sciences* 469, 20120500. doi:10.1098/rspa.2012.0500

337 Allen, J. I., Aiken, J., Anderson, T. R., Buitenhuis, E., Cornell, S., Geider, R. J., et al. (2010). Marine
338   ecosystem models for earth systems applications: The MarQUEST experience. *Journal of Marine*
339   *Systems* 81, 19–33. doi:10.1016/j.jmarsys.2009.12.017

340 Anderson, J. L. (2001). An Ensemble Adjustment Kalman Filter for Data Assimilation. *Monthly Weather*
341   *Review* 129, 2884–2903. doi:10.1175/1520-0493(2001)129⟨2884:AEAKFF⟩2.0.CO;2

342 Anderson, T. R. (2010). Progress in marine ecosystem modelling and the "unreasonable effectiveness of
343   mathematics". *Journal of Marine Systems* 81, 4–11. doi:10.1016/j.jmarsys.2009.12.015

344 Anderson, T. R., Gentleman, W. C., and Sinha, B. (2010). Influence of grazing formulations on the
345   emergent properties of a complex ecosystem model in a global ocean general circulation model. *Progress*
346   *in Oceanography* 87, 201–213. doi:10.1016/j.pocean.2010.06.003

347 Anugerahanti, P., Roy, S., and Haines, K. (2018). A perturbed biogeochemistry model ensemble evaluated
348   against in situ and satellite observations. *Biogeosciences* 15, 6685–6711. doi:10.5194/bg-15-6685-2018

349 Behrenfeld, M. J. and Boss, E. S. (2014). Resurrecting the ecological underpinnings of ocean plankton
350   blooms. *Annual review of marine science* 6, 167–94. doi:10.1146/annurev-marine-052913-021325

351 Behrenfeld, M. J., Doney, S. C., Lima, I., Boss, E. S., and Siegel, D. A. (2013). Annual cycles of
352   ecological disturbance and recovery underlying the subarctic Atlantic spring plankton bloom. *Global*
353   *Biogeochemical Cycles* 27, 1291–1293. doi:10.1002/2013GB004681

354 Butenschon, M., Clark, J., Aldridge, J. N., Icarus Allen, J., Artioli, Y., Blackford, J., et al. (2016). ERSEM
355   15.06: A generic model for marine biogeochemistry and the ecosystem dynamics of the lower trophic
356   levels. *Geoscientific Model Development* 9, 1293–1339. doi:10.5194/gmd-9-1293-2016

357 Cullen, J. J. (2014). Subsurface Chlorophyll Maximum Layers: Enduring Enigma or Mystery Solved?
358   *Annual Review of Marine Science* 7, 207–239. doi:10.1146/annurev-marine-010213-135111

359 Dave, A. C. and Lozier, M. S. (2010). Local stratification control of marine productivity in the subtropical
360   North Pacific. *Journal of Geophysical Research: Oceans* 115, 1–16. doi:10.1029/2010JC006507

361 de Boyer Montégut, C., Madec, G., Fischer, A. S., Lazar, A., and Iudicone, D. (2004). Mixed layer
362   depth over the global ocean: An examination of profile data and a profile-based climatology. *Journal of*
363   *Geophysical Research C: Oceans* 109, 1–20. doi:10.1029/2004JC002378

364 Doney, S. C. (1999). Major challenges confronting marine biogeochemical modeling. *Global*
365   *Biogeochemical Cycles* 13, 705–714

366 Doney, S. C., Lindsay, K., Caldeira, K., Campin, J., Drange, H., Dutay, J., et al. (2004). Evaluating
367   global ocean carbon models : The importance of realistic physics. *Global Biogeochemical Cycles* 18.
368   doi:10.1029/2003GB002150

369 Doney, S. C., Ruckelshaus, M., Emmett Duffy, J., Barry, J. P., Chan, F., English, C. A., et al. (2012).
370   Climate Change Impacts on Marine Ecosystems. *Annual Review of Marine Science* 4, 11–37. doi:10.
371   1146/annurev-marine-041911-111611

372 Edwards, A. and Yool, A. (2000). The role of higher predation in plankton population models. *Journal of*
373   *Plankton Research* 22, 1085–1112. doi:10.1093/plankt/22.6.1085

374 Fennel, K. and Boss, E. (2003). Subsurface maxima of phytoplankton and chlorophyll: Steady-state
375   solutions from a simple model. *Limnology and Oceanography* 48, 1521–1534. doi:10.4319/lo.2003.48.
376   4.1521

377 Friedrichs, M. A. M., Hood, R. R., and Wiggert, J. D. (2006). Ecosystem model complexity versus physical
378   forcing: Quantification of their relative impact with assimilated Arabian Sea data. *Deep-Sea Research*

379   *Part II: Topical Studies in Oceanography* 53, 576–600. doi:10.1016/j.dsr2.2006.01.026

380 Fussmann, G. F. and Blasius, B. (2005). Community response to enrichment is highly sensitive to model
381   structure. *Biology letters* 1, 9–12. doi:10.1098/rsbl.2004.0246

382 Gentleman, W., Leising, A., Frost, B., Strom, S., and Murray, J. (2003). Functional responses for
383   zooplankton feeding on multiple resources: A review of assumptions and biological dynamics. *Deep-Sea*
384   *Research Part II: Topical Studies in Oceanography* 50, 2847–2875. doi:10.1016/j.dsr2.2003.07.001

385 Hawkins, E. and Sutton, R. (2012). Time of Emergence of climate signals. *Geophysical Research Letters*
386   38, 1–6. doi:10.1029/2011GL050087

387 Hemmings, J. C. P., Challenor, P. G., and Yool, A. (2015). Mechanistic site-based emulation of a global
388   ocean biogeochemical model (MEDUSA 1.0) for parametric analysis and calibration: An application of
389   the Marine Model Optimization Testbed (MarMOT 1.1). *Geoscientific Model Development* 8, 697–731.
390   doi:10.5194/gmd-8-697-2015

391 Henson, S., Cole, H., Beaulieu, C., and Yool, A. (2013). The impact of global warming on seasonality of
392   ocean primary production. *Biogeosciences* 10, 4357–4369. doi:10.5194/bg-10-4357-2013

393 Huisman, J., Pham Thi, N. N., Karl, D. M., and Sommeijer, B. (2006). Reduced mixing generates
394   oscillations and chaos in the oceanic deep chlorophyll maximum. *Nature* 439, 322–325. doi:10.1038/
395   nature04245

396 Kriest, I., Khatiwala, S., and Oschlies, A. (2010). Towards an assessment of simple global marine
397   biogeochemical models of different complexity. *Progress in Oceanography* 86, 337–360. doi:10.1016/j.
398   pocean.2010.05.002

399 Kriest, I., Oschlies, A., and Khatiwala, S. (2012). Sensitivity analysis of simple global marine
400   biogeochemical models. *Global Biogeochemical Cycles* 26, 1–15. doi:10.1029/2011GB004072

401 Kwiatkowski, L., Yool, A., Allen, J. I., Anderson, T. R., Barciela, R., Buitenhuis, E. T., et al. (2014).
402   IMarNet: An ocean biogeochemistry model intercomparison project within a common physical ocean
403   modelling framework. *Biogeosciences* 11, 7291–7304. doi:10.5194/bg-11-7291-2014

404 Letelier, R. M., Karl, D. M., Abbott, M. R., and Bidigare, R. R. (2004). Light driven seasonal patterns
405   of chlorophyll and nitrate in the lower euphotic zone of the North Pacific Subtropical Gyre. *Limnol.*
406   *Oceanogr* 49, 508–519. doi:10.4319/lo.2004.49.2.0508

407 Mignot, A., Claustre, H., Uitz, J., Poteau, A., Ortenzio, F. D., and Xing, X. (2014). Understanding the
408   seasonal dynamics and the deep chlorophyll maximum in oligotrophic. *AGU. global biogeochemical*
409   *cycles* , 856–876doi:10.1002/2013GB004781.Received

410 Moradkhani, H. and Meskele, T. T. (2010). Probabilistic Assessment of the Satellite Retrieval Error
411   Translation to Hydrologic Response. In *Satellite Rainfall Applications for Surface Hydrology*, eds.
412   M. Gebremichael and F. Hossain (Netherlands: Springer), chap. II. 235–241

413 Murphy, J. M., Booth, B. B. B., Collins, M., Harris, G. R., Sexton, D. M. H., and Webb, M. J. (2007). A
414   methodology for probabilistic predictions of regional climate change from perturbed physics ensembles.
415   *Philosophical transactions. Series A, Mathematical, physical, and engineering sciences* 365, 1993–2028.
416   doi:10.1098/rsta.2007.2077

417 Najjar, R. G., Jin, X., Louanchi, F., Aumont, O., Caldeira, K., Doney, S. C., et al. (2007). Impact
418   of circulation on export production, dissolved organic matter, and dissolved oxygen in the ocean:
419   Results from Phase II of the Ocean Carbon-cycle Model Intercomparison Project (OCMIP-2). *Global*
420   *Biogeochemical Cycles* 21. doi:10.1029/2006GB002857

421 Palter, J. B., Lozier, M. S., and Barber, R. T. (2005). The effect of advection on the nutrient reservoir in the
422   North Atlantic subtropical gyre. *Nature* 437, 687–692. doi:10.1038/nature03969

423 Roy, S., Broomhead, D. S., Platt, T., Sathyendranath, S., and Ciavatta, S. (2012). Sequential variations of
424   phytoplankton growth and mortality in an NPZ model: A remote-sensing-based assessment. *Journal of*
425   *Marine Systems* 92, 16–29. doi:10.1016/j.jmarsys.2011.10.001

426 Séférian, R., Bopp, L., Gehlen, M., Orr, J. C., Ethé, C., Cadule, P., et al. (2013). Skill assessment of
427   three earth system models with common marine biogeochemistry. *Climate Dynamics* 40, 2549–2573.
428   doi:10.1007/s00382-012-1362-8

429 Siegel, D. A., Doney, S. C., and Yoder, J. A. (2002). The North Atlantic spring phytoplankton bloom and
430   Sverdrup's critical depth hypothesis. *Science* 296, 730–3. doi:10.1126/science.1069174

431 Sinha, B., Buitenhuis, E. T., Le Quere, C., and Anderson, T. R. (2010). Progress in Oceanography
432   Comparison of the emergent behavior of a complex ecosystem model in two ocean general circulation
433   models. *Progress in Oceanography* 84, 204–224. doi:10.1016/j.pocean.2009.10.003

434 Smith, S., Yamanaka, Y., Pahlow, M., and Oschlies, A. (2009). Optimal uptake kinetics: physiological
435   acclimation explains the pattern of nitrate uptake by phytoplankton in the ocean. *Marine Ecology*
436   *Progress Series* 384, 1–12. doi:10.3354/meps08022

437 Steele, J. H. and Henderson, E. W. (1992). The role of predation in planckton models. *Journal of Plankton*
438   *Research* 14, 157–172. doi:10.1093/plankt/14.1.157

439 [Dataset] Storkey, D., Blockley, E., Furner, R., Guiavarc'h, C., Lea, D., Martin, M., et al. (2010).
440   Forecasting the ocean state using NEMO:The new FOAM system. doi:10.1080/1755876X.2010.
441   11020109

442 Subramanian, A. C. and Palmer, T. N. (2017). Ensemble superparameterization versus stochastic paramete-
443   rization: A comparison of model uncertainty representation in tropical weather prediction. *Journal of*
444   *Advances in Modeling Earth Systems* 9, 1231–1250. doi:10.1002/2016MS000857

445 Tebaldi, C. and Knutti, R. (2007). The use of the multi-model ensemble in probabilistic climate projections.
446   *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*
447   365, 2053–2075. doi:10.1098/rsta.2007.2076

448 Tinker, J., Lowe, J., Holt, J., Pardaens, A., and Wiltshire, A. (2015). Validation of an ensemble modelling
449   system for climate projections for the northwest European shelf seas. *Progress in Oceanography* 138,
450   211–237. doi:10.1016/j.pocean.2015.07.002

451 Yool, A., Popova, E. E., and Anderson, T. R. (2011). MEDUSA-1.0: A new intermediate complexity
452   plankton ecosystem model for the global domain. *Geoscientific Model Development* 4, 381–417.
453   doi:10.5194/gmd-4-381-2011

454 Yool, A., Popova, E. E., and Anderson, T. R. (2013). MEDUSA-2.0: An intermediate complexity
455   biogeochemical model of the marine carbon cycle for climate change and ocean acidification studies.
456   *Geoscientific Model Development* 6, 1767–1811. doi:10.5194/gmd-6-1767-2013