AGU PUBLICATIONS

JGR

Journal of Geophysical Research: Oceans

RESEARCH ARTICLE

10.1002/2015JC011080

Inferred support for disturbance-recovery hypothesis of North Atlantic phytoplankton blooms

Matthew J. Smith[1], Derek P. Tittensor[2,3], Vassily Lyutsarev[1], and Eugene Murphy[4]

[1]Computational Science Laboratory, Microsoft Research, Cambridge, UK, [2]Department of Biology, Dalhousie University, Halifax, Nova Scotia, Canada, [3]United Nations Environment Programme World Conservation Monitoring Centre, Cambridge, UK, [4]British Antarctic Survey, NERC, Cambridge, UK

Abstract Analyses of satellite-derived chlorophyll data indicate that the phase of rapid phytoplankton population growth in the North Atlantic (the "spring bloom") is actually initiated in the winter rather than the spring, contradicting Sverdrup's critical depth hypothesis. An alternative disturbance-recovery hypothesis (DRH) has been proposed to explain this discrepancy, in which the rapid deepening of the mixed layer reduces zooplankton grazing rates sufficiently to initiate the bloom. We use Bayesian parameter inference on a simple Nutrient-Phytoplankton-Zooplankton (NPZ) model to investigate the DRH and also investigate how well the model can capture the multiyear and spatial dynamics of phytoplankton concentrations and population growth rates. Every parameter in our NPZ model was inferred as a probability distribution given empirical constraints, which provides a more objective method to identify a model parameterization given available empirical evidence, rather than fixing or tuning individual parameter values. Our model explains around 75% of variation in the seasonal dynamics of phytoplankton concentrations, 30% of variation in their population rates of change, and correctly predicts the phases of population growth and decline. Our parameter-inferred model supports the DRH, revealing the sustained reduction of grazing due to mixed-layer deepening as the driving mechanism behind bloom initiation, with the relaxation of nutrient limitation being another contributory mechanism. Our results also show that the continuation of the bloom is caused in part by the maintenance of phytoplankton concentrations below a level that can support positive zooplankton population growth. Our approach could be employed to formally assess alternative hypotheses for bloom formation.

Key Points:
- Inferred parameter model supports disturbance-recovery hypothesis for plankton blooms
- Bloom initiation occurs in winter when mixed layer deepens below euphotic zone depth
- Bloom prolonged by grazer population decline due to diluted phytoplankton concentrations

Supporting Information:
- Supporting Information S1

Correspondence to:
M. J. Smith,
Matthew.Smith@Microsoft.com

Citation:
Smith, M. J., D. P. Tittensor, V. Lyutsarev, and E. Murphy (2015), Inferred support for disturbance-recovery hypothesis of North Atlantic phytoplankton blooms, *J. Geophys. Res. Oceans*, *120*, 7067–7090, doi:10.1002/2015JC011080.

Received 25 JUN 2015
Accepted 6 OCT 2015
Accepted article online 8 OCT 2015
Published online 31 OCT 2015

1. Introduction

Phytoplankton blooms are commonly characterized as periods of anomalously high concentrations of phytoplankton over a period of time (usually days to weeks) [*Smayda*, 1997]. While there are different precise definitions for phytoplankton blooms [*Smayda*, 1997], our study here focuses on the sustained period (sometimes up to months) of positive population growth rates that leads to high phytoplankton concentrations. Phytoplankton blooms have been the subject of scientific research for decades, following the study of *Sverdrup* [1953]. The annual spring bloom event in the North Atlantic represents one of the most important examples, occurring over a large geographical area and with significant consequences for the dynamics of higher tropic levels, commercial fisheries, and the marine carbon cycle [*Beaugrand et al*., 2003; *Koeller et al*., 2009; *Behrenfeld and Boss*, 2014; *Siegel et al*., 2014]. This bloom is most noticeable in the mid to late spring when phytoplankton concentrations reach their peak. Attempting to dissect the ecological and climatic factors responsible has been a matter of significant scientific interest [*Lindemann and St. John*, 2014; *Behrenfeld and Boss*, 2014].

For decades, the Critical Depth Hypothesis (CDH), as published by *Sverdrup* [1953], has been the leading hypothesis for the spring bloom. The principal mechanism behind the CDH is the change in the depth of the mixed layer. In the winter, the mixed layer is relatively deep and phytoplankton spend a large fraction of their time in waters that have insufficient light to support photosynthesis. The CDH predicts that during the winter the losses due to mortality exceed the gains due to cell division (phytoplankton growth) and the population growth rate is negative. However, at the beginning of spring, the mixed layer progressively shallows such that phytoplankton spend an increasing fraction of their time in water suitable for photosynthesis. The "critical depth" is the depth of the mixed layer above which gains from photosynthesis exceed

losses, allowing for positive population growth rates. The CDH postulates that phytoplankton blooms are triggered when the mixed layer becomes shallower than the critical depth.

A large body of research has aimed to test the CDH and alternative hypotheses [*Behrenfeld and Boss*, 2014; *Franks*, 2014; *Taylor and Ferrari*, 2011]. One major challenge has been a lack of empirical data; however, advances in monitoring the oceans, such as through Lagrangian floats, satellites, and ship-towed plankton recorders, have generated a wealth of new data, leading to new insights [*Mahadevan et al*., 2012; *Barton et al*., 2015; *Brody and Lozier*, 2014; *Chiswell*, 2011; *Taylor and Ferrari*, 2011; *Platt et al*., 2009; *Behrenfeld*, 2010; *Boss and Behrenfeld*, 2010; *Brody et al*., 2013]. Some of these studies have suggested that phytoplankton population growth rates become positive prior to midwinter, well before the mixed layer has shallowed to a hypothetical critical depth and indeed even while the mixed layer is deepening [*Behrenfeld*, 2010; *Boss and Behrenfeld*, 2010]. This is not a straightforward phenomenon to detect: phytoplankton concentrations are decreasing at this time because of dilution through addition of phytoplankton-free waters from below as the mixed layer deepens.

The most prominent alternative hypotheses to the CDH are the critical turbulence hypothesis (CTH) [*Huisman et al*., 1999; *Taylor and Ferrari*, 2011], the eddy stratification hypothesis (ESH) [*Mahadevan et al*., 2012], and the disturbance-recovery hypothesis (DRH) [*Behrenfeld*, 2010]. Under the CTH, a critical reduction in turbulence causes phytoplankton in the upper fraction of the mixed layer to spend more time in environmental conditions favoring photosynthesis, leading to positive population growth rates. A similar mechanism is proposed under the ESH except it is the development of eddies in the mixed layer that stratify the upper ocean that causes the phytoplankton to experience more favorable conditions. However, both the ESH and the CTH have been criticized because the conditions that lead to positive phytoplankton population growth rates occur during or after the time of the maximum depth of the mixed layer, and analyses of satellite chlorophyll data suggest that plankton blooms can start prior to this event [*Behrenfeld and Boss*, 2014]. Neither the ESH nor the CTH are the focus of our study and so we do not describe them further here, but do return to them in section 4. The DRH postulates that bloom initiation is caused by a physical or ecological disturbance to factors acting to reduce phytoplankton population growth rates (which may be different in different circumstances) that causes the phytoplankton population growth rates to become positive for a sustained period of time, allowing for the accumulation of mass and the formation of a bloom [*Behrenfeld et al*., 2013]. For the North Atlantic, this disturbance is hypothesized to be the deepening of the mixed layer in the autumn and winter which causes entrainment of water with low phytoplankton concentrations and higher nutrient concentrations from below. This dilution causes the phytoplankton concentrations to decrease, which is hypothesized to reduce the grazing rate by zooplankton, with the consequence of positive phytoplankton population growth rates, even when their concentration is declining. The hypothesis therefore predicts that bloom initiation begins in the late autumn/early winter, while the mixed layer is deepening and prior to the winter solstice, consistent with empirical data [*Behrenfeld*, 2010; *Boss and Behrenfeld*, 2010].

A few studies have sought to investigate whether the mechanisms responsible for bloom formation proposed by the DRH can be reproduced by mathematical models [*Behrenfeld et al*., 2013; *Kuhn et al*., 2015; *Lévy*, 2015]. Several of these have shown that bloom initiation in the North Atlantic can occur in the late autumn/early winter, associated with the reduction of grazing pressure via mixed-layer deepening. *Behrenfeld et al*. [2013] found support for the DRH by analyzing predictions of the relatively complex Biogeochemical Element Cycling-Community Climate System Model (BEC-CCSM) [*Behrenfeld et al*., 2013]; however, a complication is that the BEC-CCSM has so many parameters and degrees of freedom that uncertainty remains about the extent to which the results are a function of the specific model configuration (parameters values and functional forms) [*Franks*, 2002, 2009; *Anderson*, 2005]. *Lévy* [2015] found support for the DRH using a set of much simpler models but also found that mechanisms postulated under the critical turbulence hypothesis could play an important role in bloom formation. *Kuhn et al*. [2015] used optimization techniques to parameterize a simple a nutrient-phytoplankton-zooplankton-detritus (NPZD) model, which they then used to investigate North Atlantic bloom formation. Their model predicted that positive population growth rates in early winter were associated with mixed-layer deepening, consistent with the DRH, but also found that physical disruption of grazing by mixed-layer deepening was not necessary for bloom formation and that changes in zooplankton biomass caused by mixed-layer deepening may not in fact have a major effect on the timing of onset of the blooms.

The study of *Kuhn et al*. [2015] showed that a simple model with inferred parameters could both accurately reproduce average seasonal phytoplankton dynamics and be used to investigate the mechanisms of bloom formation. Inferring parameters and state variables based on the match between model predictions and empirical data is an increasingly used technique in ecological applications [*Matear*, 1995; *Friedrichs et al*., 2007; *Jones et al*., 2010; *Dowd*, 2011; *Liu and Scavia*, 2010; *Hahn-Woernle et al*., 2014; *Kuhn et al*., 2015]. A common term for this approach is inverse modeling, in which rather than aiming to predict the system dynamics from a model, the aim is to predict the model parameters based on the match between the model predictions and the data. In an early example, *Matear* [1995] inferred parameters for three models to investigate the extent to which they could capture the plankton dynamics observed at a North Atlantic site. In that study, all models could capture the quantitative dynamics well but the number of poorly constrained parameters increased with model complexity. Such poorly constrained parameters imply that multiple alternative parameter values are equally likely. Poorly constrained parameters can also generate highly uncertain predictions, if their full possible range is translated into a range of predictions [*Denman*, 2003; *Anderson*, 2005; *Franks*, 2009]. However, recent increases in data availability enable most of the parameters of simple models to be well constrained, as illustrated by *Kuhn et al*. [2015]. Even with well-constrained parameters and good predictions, incorrect conclusions about the underpinning mechanisms can still occur because the assumed model formulation (the equations and functional forms in the model) is inapplicable. This could especially be a problem with simple models, where often grossly simplifying assumptions are made. This problem will be alleviated over time as more comparisons are done with models with differing complexities, alternative assumptions, and improvements in their parameterizations.

Our study here contributes to this ongoing process of increasing our understanding of phytoplankton bloom formation through the use of data-constrained models. In a previous study, *Behrenfeld and Boss* [2014] showed that the simple nutrient-phytoplankton-zooplankton (NPZ) model of *Evans and Parslow* [1985] can generate plankton blooms in a way that is consistent with the DRH. However, they did not show that the model predictions could actually match observational data. Given that dynamics consistent with multiple alternative hypotheses could be obtained with a model depending on its parameterization, it is appropriate to evaluate whether there exists a parameterization of the *Evans and Parslow* [1985] model that matches the observations, in terms of the spatial and temporal dynamics of phytoplankton and their population growth rates, and whether it is consistent with the DRH. In this study, we had two aims: (i) to assess how well the NPZ model of *Evans and Parslow* [1985], fitted to the data using Bayesian parameter estimation, could quantitatively capture the dynamics of phytoplankton in the North Atlantic, and (ii) to use such a model to assess whether it provides support for the DRH for the North Atlantic phytoplankton blooms.

Our approach resembles that of *Kuhn et al*. [2015] in terms of using a simple data-constrained model to investigate mechanisms underpinning North Atlantic phytoplankton blooms, but it differs in a number of important details. First, we use Bayesian parameter inference rather than parameter optimization. This enables us to infer the model parameters as a joint probability distribution given uncertainty in the observational data sets, rather than individual best estimates. One advantage of this is that we can directly assess how well the inferred parameter values are constrained and whether they are correlated in their probability distributions (this latter analysis does not turn out to be insightful in our study). However, the key advantage Bayesian parameter inference brings is that it naturally enables us to evaluate of how the resulting parameter uncertainty propagates through predictions of the final models and affects conclusions.

Another departure from *Kuhn et al*. [2015] is that we do not average our time series data on phytoplankton concentrations into seasonal averages because we want to assess the predictive ability of our model even when the data includes year to year variability. A resulting benefit is that we can assess the model predictive ability against data using years omitted from the parameter inference process, which is good practice to help avoid including mechanisms that only help the model explain the data it is being trained to (known as overfitting). Using multiyear data also enables us to use time series data with gaps: as is frequent, the satellite data we use contain gaps for some sites due to a lack of sunlight in the winter. *Kuhn et al*. [2015] had to omit the more northerly sites from their analysis because of this difficulty. Another difference to *Kuhn et al*. [2015] is that we data constrain a different model formulation, in particular ours is not vertically resolved (it also does not contain a detritus component). While it was tempting to include this detail because it could potentially have made simulations of vertical dynamics more realistic, we chose to keep our model matching the *Evans and Parslow* [1985] formulation to address our aims.

Table 1. Parameter Definitions and Values for the Models Used in This Study

Textual Definition	Symbol	Units	Prior Range	Justification for Value
Nitrate uptake half saturation	j	mM m^{-3}	0.25–2	0.25–400% value in EP85
Phytoplankton background mortality rate	ω	d^{-1}	0.0175–0.28	0.25–400% value in EP85
Grazing half saturation	K	mM m^{-3}	0.25–4	0.25–400% value in EP85
Maximum grazing rate	c	day^{-1}	0.25–4	0.25–400% value in EP85
Grazing efficiency	f		0.125–1	0.25–200% value in EP85; $f > 1$ does not make biological sense
Zooplankton background mortality rate	g	day^{-1}	0.0175–0.28	0.25–400% value in EP85
Maximum photosynthetic rate (supporting information Text S1)	Q	day^{-1}	0.5–8	0.25–400% value in EP85
Lowlight photosynthetic slope (supporting information Text S1)	s	(W m^{-2})$^{-1}$	420–6700	0.25–400% value in EP85
Light attenuation by water (supporting information Text S1)	k	m^{-1}	0.025–0.4	0.25–400% value in EP85
Light attenuation by phytoplankton (supporting information Text S1)	l	mM m^{-1}	0.03–0.48	0.25–400% value in EP85
Nitrate diffusion rate	m_N	m d^{-1}	0.75–12	0.25–400% value in EP85
Phytoplankton diffusion rate out of mixed layer	m_P	m d^{-1}	0.75–12	0.25–400% value in EP85
Nitrate conc. below the mixed layer	N_d	mM m^{-3}	2.5–40	0.25–400% value in EP85

2. Methods

We will first describe the model and data sets we used in our study, and the numerical technique we used to drive the model with time series data on the physical environment (such as mixed-layer depth). We will then describe our Bayesian parameter inference technique in some detail because while it is increasingly commonly used [*Jones et al*., 2010; *Dowd*, 2011; *Liu and Scavia*, 2010], it will still be unfamiliar to many.

2.1. Models

We used two models both based on *Evans and Parslow* [1985, hereon denoted EP85]. The first model (Model 1) is almost identical to EP85, while the second (Model 2) has a subtle but important modification to the effects mixed layer depth changes on phytoplankton and zooplankton concentrations.

Both models have the classic three pool nutrient-phytoplankton-zooplankton structure (nitrate being the only nutrient we consider here), with changes in their concentrations given as:

$$\frac{dN}{dt} = \omega P + gZ + \frac{(1-f)cP^2Z}{K+P^2} - \frac{\alpha(L_d, M_d, P)PN}{j+N} - (N-N_d)\zeta_N(M_d)/M_d, \tag{1a}$$

$$\frac{dP}{dt} = \frac{\alpha(L_d, M_d, P)PN}{j+N} - \omega P - \frac{cP^2Z}{K+P^2} - P\zeta_P(M_d)/M_d, \tag{1b}$$

and

$$\frac{dZ}{dt} = \frac{fcP^2Z}{K+P^2} - gZ - Z\zeta_Z(M_d)/M_d, \tag{1c}$$

where N, P, and Z are the concentrations of nitrate, phytoplankton, and zooplankton respectively (all with units mM N m^{-3}); α a function that predicts the phytoplankton growth rate (day^{-1}) in nitrate unlimited conditions as a function of light throughout the mixed layer (identical to EP85; see supporting information Text S1 of this paper for the equations); L_d and M_d are, respectively, the light intensity reaching the water surface (W m^{-2}) and the mixed-layer depth (m); both of these varied over time according to driver data (see driver data below); ζ_N, ζ_P, and ζ_Z determine the effects of mixed-layer depth and its rate of change on nitrate, phytoplankton, and zooplankton concentrations within the mixed layer; and the other parameter definitions are given in Table 1.

Models 1 and 2 differ only in how the properties of the mixed layer affect nitrate, phytoplankton, and zooplankton concentrations. For Model 1 we make the same assumptions as EP85, with

$$\zeta_N(M_d) = m_N + \max(0.0,\, dM_d/dt), \tag{2a}$$

$$\zeta_P(M_d) = m_P + \max(0.0,\ dM_d/dt), \tag{2b}$$

and

$$\zeta_z(M_d) = dM_d/dt. \tag{2c}$$

According to this model, nitrate in the mixed layer changes as a result of diffusion across the mixed layer (with rate proportional to m_N), the difference in nitrate concentrations between the mixed layer and the rest of the water column, and due to the entrainment of water below the mixed layer as it deepens. Phytoplankton are also assumed to diffuse out of the system and are also diluted when phytoplankton-free water is entrained as the mixed layer deepens. Importantly, it is assumed that the phytoplankton and zooplankton concentrations below the mixed layer are zero. Zooplankton are assumed to actively avoid leaving the mixed layer and so their concentrations increase in response to mixed-layer shallowing and decrease in response to mixed-layer deepening (all else being equal). Phytoplankton are assumed to be entirely passive and are thus lost from the mixed layer as it shallows (their concentrations in the mixed layer are unaffected) and are diluted as the mixed layer deepens.

Since EP85, studies have shown that some autotrophic organisms can employ active mechanisms to maintain their presence within the mixed layer [*Cullen*, 2015]. In addition, studies have shown that phytoplankton concentrations can remain relatively high down to the depth at which light has been reduced to 1% of surface light, the euphotic zone depth, M_{eu}, beyond which their concentrations are much reduced [*Cullen*, 2015; *Westberry et al*., 2008]; this depth can sometimes be greater than the mixed-layer depth. We therefore created Model 2 to reflect these empirical observations. For nitrate and phytoplankton, we modeled the effects of mixed-layer deepening as

$$\zeta_{NP}(M_d) = 0,\ \text{ when }\ 0 < M_d < M_{eu}, \tag{3a}$$

but

$$\zeta_{NP}(M_d) = (F(dM_d/dt) + \max(0.0,\ dM_d/dt)),\ \text{ when }\ M_d > M_{eu}, \tag{3b}$$

where $NP{=}N$ for nitrate and $NP = P$ for phytoplankton, $F = m_{NP}$ when $dM_d/dt < 0$ and $F{=}0$ when $dM_d/dt > 0$. Thus, we assume that nitrate and phytoplankton concentrations are not affected by mixed-layer deepening down to M_{eu} (equation (3a)). However, beyond that they are diluted when the mixed layer is deepening ($dM_d/dt > 0$ in equation (3b)) and can diffuse out of the mixed layer when it is shallowing ($F{=}m_P$, in equation (3b)). Our formulation of $F(dM_d/dt)$ reflects a simplifying assumption that diffusion out of the mixed layer is insignificant as it is deepening because anything that sinks out of the mixed layer is soon incorporated back into the mixed layer. Experiments with the alternative assumption that $F{=}m_{NP}$ when $dM_d/dt > 0$ led to trivial quantitative effects on our results.

We model the effects of mixed-layer deepening on zooplankton as

$$\zeta_Z(M_d) = 0,\ \text{ when }\ 0 < M_d < M_{eu}, \tag{4a}$$

but

$$\zeta_Z(M_d) = dM_d/dt,\ \text{ when }\ M_d > M_{eu}. \tag{4b}$$

Thus, for the second model, we assume that their concentrations are not affected by mixed-layer deepening down to M_{eu} but beyond that the zooplankton swim to maintain their presence within the mixed layer.

2.2. Study Area, Driver, and Comparison Data

We chose a slightly smaller but similar region of the North Atlantic to that of *Behrenfeld* [2010], focusing on the rectangular region between 45°W and 15°W and between 40°N and 55°N (Figure 1). This region includes the three northern bins NA-7, NA-8, and NA-9 that were excluded by *Kuhn et al*. [2015] because they contained gaps in their time series in the midwinter. Our time period of interest was the 10 years from the beginning of 1998 to the end of 2007. For this period, it is possible to obtain remotely sensed time series data representative of phytoplankton concentrations in the mixed layer as well as the environmental variables necessary to drive our model (such as mixed layer depth). Data were aggregated to 12.5° latitude × 10° longitude "bins," resulting in the 3 × 3 bin lattice labeled regions NA-1 to NA-9 by *Behrenfeld* [2010], a naming we adopt here (Figure 1). This spatial resolution is coarser than the original resolution of the

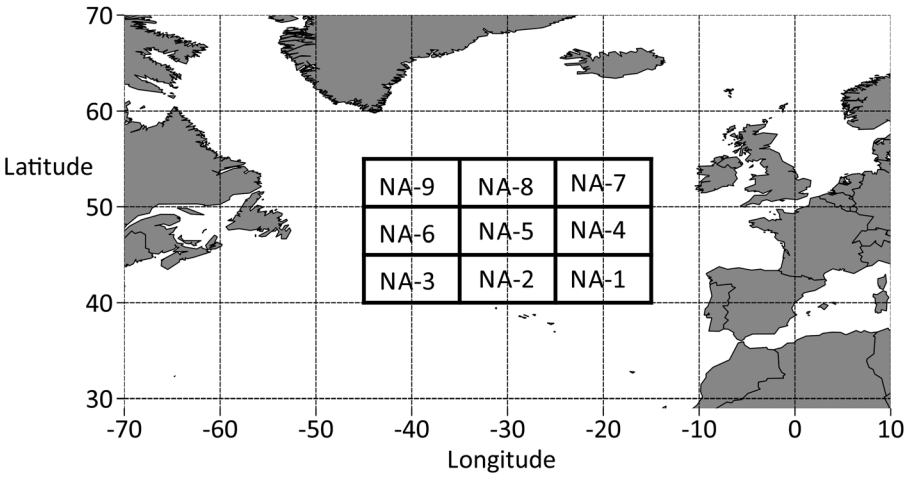

Figure 1. The nine North Atlantic bins (rectangular geographical regions) that are the focus of our study here are cells of a 3 × 3 rectangular grid from −45 to −15 longitude in steps of 10° and from 40 to 55 latitude in steps of 5°. The identities of these grid cells (NA-1 to NA-9) are identical to those used by *Behrenfeld* [2010].

phytoplankton and environmental data sets to reduce the effects of lateral advection on the measured temporal dynamics of phytoplankton; however, the benefit of studying multiple regions is that we can observe how our results vary spatially.

We used three different but nonindependent time series data sets as constraints in our parameter estimation routine: phytoplankton carbon concentrations (in mg m^{-3}; the only type of time series data used as a parameter constraint by *Kuhn et al.* [2015]), the rates of change of phytoplankton carbon concentrations, and euphotic zone depths (in m). The original source data are satellite products derived from applying algorithms to measurements of water leaving radiances by the Sea-viewing Wide Field-of-view Sensor (SeaWiFS) remote sensing instrument, as described in more detail by *Behrenfeld* [2010]. The data were downloaded from the Ocean productivity web portal (http://www.science.oregonstate.edu/ocean.productivity/) with a temporal resolution of 8 days and a spatial resolution of one sixth of a decimal degree. We coarsened these data sets to the 12.5° latitude × 10° longitude bins by taking the mean and standard deviation across the log transformed time series for each grid cell.

The satellite product on phytoplankton concentrations reports them in mg carbon m^{-3} [*Behrenfeld*, 2010] but our model predicts concentrations in m*M* nitrogen m^{-3}. Therefore, to convert the model predictions to compare with data, we assumed the Redfield ratio of 106 C atoms to 16 N atoms [*Redfield*, 1934].

The data on the rates of change of phytoplankton carbon concentrations were computed as $gr = \ln(P(t)/P(t-\Delta t))/\Delta t$, where gr is the rate of change of phytoplankton concentration, $P(t)$, at time t, and Δt is the time window. Note that the rate of change of phytoplankton concentrations includes dilution and, thus, is different from the rate of change of phytoplankton populations which discounts the effects of dilution (detailed below).

The data on euphotic zone depths were computed by applying the formulae of *Morel and Berthon* [1989]

$$M_{eu} = 568.2[chl]^{-0.746}, \text{ when } M_{eu} < 102 \text{ m}, \tag{5a}$$

and

$$M_{eu} = 200[chl]^{-0.293}, \text{ when } M_{eu} > 102 \text{ m}, \tag{5b}$$

where $[chl]$ is the total pigment content in the euphotic layer (mg m^{-2}). This circular relationship occurs because the formulae were derived from a nonlinear empirical relationship between chlorophyll concentrations and euphotic zone depths. The data on $[chl]$ used to compute M_{eu} via. equation (5) were downloaded from http://www.science.oregonstate.edu/ocean.productivity/).

Computing model predicted euphotic zone depth to compare with that derived from empirical data was straightforward: our photosynthesis model assumes light intensity decays through the water column

according to $\exp -(k+lP)h$, where k and l are inferred light decay coefficients, P is the phytoplankton concentration, and h is the distance from the water surface (supporting information Text 1) [*Evans and Parslow*, 1985]. This can be used to calculate the predicted M_{eu}, the distance from the water surface at which light intensity is 1% of that at the surface.

We used two data sets as drivers for our model simulations: mixed-layer depths and sea surface light intensities (*Kuhn et al*. [2015] used an astronomical formula rather than estimated actual light intensities). As in *Behrenfeld* [2010], we used mixed-layer depth (MLD) data derived from data-constrained ocean circulation models: the Fleet Numerical Meteorology and Oceanography Centre (FNMOC) model [*Clancy and Sadler*, 1992] for the years 1998 to 2004, the Simple Ocean Data Assimilation (SODA) model from day 169 in year 2005 until 2008, and the Thermal Ocean Prediction Model (TOPS) for the first 168 days in year 2005. Outputs from these three different models were used because they cover different time periods. These were all downloaded from the Ocean productivity web portal (http://www.science.oregonstate.edu/ocean.productivity/) with a temporal resolution of 8 days and a spatial resolution of one sixth of a decimal degree.

Data on light reaching the sea surface were obtained from NCEP Reanalysis Daily Averages Surface Flux data, downloaded from http://www.esrl.noaa.gov/psd/data/gridded/data.ncep.reanalysis.surfaceflux.html. These had a temporal resolution of 1 day and a spatial resolution of 1.92 decimal degrees latitude $\times$ 1.875 decimal degrees longitude.

2.3. Calculation of the Rate of Change of Phytoplankton Populations and Other Key Metrics

We were particularly interested in the rates of change of phytoplankton populations as a metric; that is, the component of the rate of change of phytoplankton concentrations that occurs separately from dilution effects. For the model simulations, these were calculated using equation (1b) with the effects of mixed-layer deepening removed (i.e., setting $dM_d/dt=0$). For the satellite data, we calculated the rate of change of the phytoplankton population using the method defined by *Behrenfeld* [2010]. That is,

$$r = \frac{\ln\left(\frac{\sum P(t)}{\sum P(t-\Delta t)}\right)}{\Delta t}, \text{ when } M_d(t) > M_d(t-\Delta t) \text{ and } M_d(t) > M_{eu}, \tag{6a}$$

and

$$r = \frac{\ln\left(\frac{P(t)}{P(t-\Delta t)}\right)}{\Delta t}, \text{ when } M_d(t) < M_d(t-\Delta t) \text{ or } M_d(t) < M_{eu}, \tag{6b}$$

where r is the rate of change of the phytoplankton population at time t, Δt is the length of time between consecutive datapoints (mostly 8 days and we exclude larger gaps in the calculation of r), P is the satellite-derived concentration of phytoplankton carbon, $\sum P = PM_d$ is the depth-integrated estimate of phytoplankton carbon content, and M_{eu} is the euphotic depth, which was calculated according to equation (5) above.

Behrenfeld [2014] hypothesized that the sustained acceleration in phytoplankton growth (cell division) rates after bloom initiation could be the mechanism driving the prolonged period of growth in phytoplankton populations throughout the late winter and early spring. To assess this hypothesis, we computed the acceleration in growth rates using the same formulation as *Behrenfeld* [2014]

$$\Delta\mu_{rel} = \frac{\mu_1 - \mu_0}{(\mu_1 + \mu_0)/2}, \tag{7}$$

where μ_1 is the phytoplankton growth rate at a time point and μ_0 is the growth rate 8 days earlier.

In our analyses, we also calculated a threshold concentration of phytoplankton above which the zooplankton feeding rate is sufficient to offset losses and the zooplankton concentration increases, and below which it decreases. This is derived by assuming the rate of change of zooplankton in equation (1c) is zero and rearranging to give

$$\tilde{P} = \sqrt{\frac{(g + \zeta_Z(M_d))K}{fc - (g + \zeta_Z(M_d))}}. \tag{8}$$

This expression has no solutions under two circumstances for realistic parameter values (see Table 1 for parameter definitions): the first is when the maximum zooplankton growth rate is less than the loss rate

due to dilution and mortality (when $fc < (g+\zeta_Z(M_d))$ and the other is when mixed-layer shallowing is so rapid (when $\zeta_Z(M_d) < 0$ and $abs(\zeta_Z(M_d)) > g$) that the zooplankton concentration inevitably increases, irrespective of the phytoplankton concentration.

In our analyses below, we also calculated the strength of nitrate limitation on growth rates as

$$N_{lim} = 1 - \frac{\mathrm{N}}{j+N},$$

which is the quantity that maximum photosynthesis rates are reduced by nitrate limitation, where j is the inferred nitrate uptake half saturation coefficient.

2.4. Numerical Solution and Knock-Out Experiments

A basic forward Euler method was used to simulate our models, adopting a daily time step; early experimentation with adaptive time step solvers revealed that the time step chosen was always approximately daily and we obtained trivial numerical differences between the two schemes. Environmental driver data were used each time step, adopting linear interpolation to obtain values for time steps without data (our coarsest temporal resolution for driver data was 8 days). Simulations also required specification of initial conditions for nitrate, phytoplankton, and zooplankton concentrations. Early experimentation with inferring these as model parameters indicated that values were being selected that were biased toward obtaining a good fit in the first year because the effects of variations in the initial conditions were largely lost after the first year of simulation. To prevent this bias, we assumed that the multiyear dynamics followed a constant average trajectory (did not increase or decrease) and simulated the model twice, the first time with arbitrary initial conditions and the second time with the initial conditions determined by the average concentrations over the end of the final 4 years of the first simulation. This procedure was used to obtain initial conditions for both the training and evaluation simulations (each simulating 5 years) detailed below.

We performed a series of simple knock-out experiments to explore the relative importance of different mechanisms, by removing variation in particular mechanisms on the seasonal dynamics of the rate of change of phytoplankton populations. These were performed using the mean parameter values derived from the joint posterior distribution of the inferred parameters. All of the knock-out experiments were first spun-up under identical conditions to the standard simulation for each bin, simulating from 1998 to 2002. The final values of nutrients, phytoplankton, and zooplankton from that simulation were then used as initial conditions in the knock-out experiments. Then the knock-out was applied by altering one or more of the mechanisms and repeating the simulation from 1998 to 2002 ten times, each time feeding the final states into the initial states of the next simulation. This was sufficient to remove transient dynamics. The final 5 years of simulation were used in the analysis. The knock-outs we explored were LightMean in which light intensity was set to the mean light level from the data for that bin, NoZ—zooplankton was fixed to zero, MLDMean—mixed-layer depth (MLD) was set to the mean MLD from the data for that bin, NUnlim—in which nitrate levels were set to be a constant value, sufficiently high to be unlimiting to phytoplankton growth rates, LightOnly—zooplankton was set to zero and MLD was set to the mean MLD, MLDOnly—zooplankton was set to zero and light was set to the mean light level, MLDOnlyNUnlim—as in MLDOnly but nitrate was also set to be unlimiting to phytoplankton growth rates, and ZOnly—MLD and Light were set to their mean levels.

2.5. Bayesian Parameter Estimation Method and Comparing Model Predictions to Data

Our goal with parameter estimation was to estimate the probability distributions of the model's parameter values given the empirical data and accounting for uncertainty in the data. Bayes's theorem [*Gilks et al.*, 1996] enables calculation of the probability of a model's parameters given empirical data. Translated into our parameter estimation problem it is

$$\mathcal{P}(\theta|\mathcal{X}) = \frac{\mathcal{P}(\mathcal{X}|\theta)\mathcal{P}(\theta)}{\int \mathcal{P}(\mathcal{X}|\theta)\mathcal{P}(\theta)d\theta}, \tag{9}$$

where we use $\mathcal{P}$ to denote a probability, $\mathcal{X}$ to denote our data (phytoplankton carbon concentrations, their rates of change and euphotic zone depths), and θ to denote our vector of model parameters (those in Table 1) [*Gilks et al.*, 1996]. Equation (9) states that we can calculate the probability distributions for our model parameters (the posterior distribution) by calculating the three terms on the right-hand side,

although for the purposes of this study we only need to compute two terms in the numerator, with the denominator acting as a normalizing constant (which we assume equal to 1) [*Gilks et al*., 1996]. The term for the probability of the parameters $\mathcal{P}(\theta)$ represents our prior belief about the true distribution of the model parameters. We chose to define broad yet biologically or physically reasonable prior ranges, with their probability uniformly distributed between maxima and minima, because of a general lack of adequate empirical data. As a consequence, our problem reduced to calculating

$$\mathcal{P}(\theta|\mathcal{X}) \propto \mathcal{L}(\mathcal{X}|\theta), \tag{10}$$

where $\mathcal{L}$ is typically referred to as the likelihood of the data given the model parameters.

To calculate $\mathcal{L}(\mathcal{X}|\theta)$, we need to formulate an expression for the probability of the data given the predictions of our model. We take advantage of the fact that the data are itself uncertain to express this as

$$\log\left(\mathcal{L}(\mathcal{X}|\theta)\right) = \sum_{b=0}^{B} \log\left(\mathcal{N}(\mathcal{X}_b, \mathcal{U} = \mathcal{Y}_b(\theta), \mathcal{S} = \mathcal{V}_b)\right), \tag{11}$$

where $\mathcal{N}$ is a normal distribution with mean $\mathcal{U}$ and variance $\mathcal{S}$, $\mathcal{Y}_b(\theta)$ is the deterministic model prediction associated with data item *b*, B is the total number of data items, and $\mathcal{V}_b$ is the variance associated with data item *b*. In words, this states that the log likelihood of the observations given the model parameters is given by the probability that each observation was drawn from a normal distribution, centered on the model predictions for that data item, with standard deviation given by the standard deviation across observations within each bin. Thus, we were assessing the probability that the model produced the data given that the only source of error was spatial heterogeneity within each bin (this issue is addressed in section 4). We used three different observational data sets (detailed above) and thus roughly every 8 days of simulation the probability of those observations was calculated and their logarithms summed. Spatial variation about mean phytoplankton concentrations within each bin was lognormally distributed and so these data were logged before computing the variance $\mathcal{V}_b$ and these log-transformed data were compared with the log of phytoplankton concentrations in the likelihood function.

Calculating the likelihood exactly is practically impossible for complex nonlinear models with more than just a few parameters and so we used the Markov Chain Monte Carlo (MCMC) algorithm combined with the Metropolis-Hastings algorithm [*Gilks et al*., 1996] to solve equation (11). We omit full details of how this algorithm works for brevity (detailed in *Gilks et al*. [1996] and *Jones et al*. [2010]) but ultimately the successful application of this technique leads to a list of log likelihoods that has a frequency distribution that approximates $\log\left(\mathcal{L}(\mathcal{X}|\theta)\right)$. This list of log likelihoods also has a list of associated parameters that have a frequency distribution that approximates the posterior joint probability distribution of the parameters $\mathcal{P}ost(\theta)$.

We used the Filzbach software library (http://research.microsoft.com/en-us/um/cambridge/groups/science/tools/filzbach/filzbach.htm) to carry out Bayesian parameter inference using the MCMC technique described above. A potential problem with MCMC parameter estimation is that it can become stuck in local maxima of $\mathcal{L}(\mathcal{X}|\theta)$. Multiple chains were therefore run to check (and confirm in our case) that they always converged to the same $\mathcal{L}(\mathcal{X}|\theta)$. We used a burn in-phase length of 50,000 iterations (used to configure internal parameters to the algorithm) and a sampling phase of 100,000 iterations (used to estimate $\mathcal{P}ost(\theta)$). A complication in interpreting the individual parameter values within $\mathcal{P}ost(\theta)$ is that they might be strongly correlated with each other. However, the majority of our parameters had very low degrees of correlation with each other (correlation coefficients between -0.4 and $+0.4$). Those few pairs that were exceptions were more tightly correlated for understandable reasons from inspecting our model equations, e.g., the maximum photosynthetic rate and the phytoplankton mortality rate tended to be positively correlated.

As part of our analysis, we wished to assess evidence for spatial variation in inferred parameter values across the bins in our model (Figure 1), to understand the extent to which bloom formation mechanisms might vary spatially, as well as the extent to which spatial variation in inferred parameter values improved the model predictions. We therefore performed Bayesian parameter estimation with our two models assuming either that the values of the estimated parameters could vary between North Atlantic bins or were common to all bins. In the first case, each bin had an independent set of parameters and the model was constrained against the data for each bin independently, we refer to these models as having bin-specific parameters. In the second case, the bins shared the same parameters and so the model was constrained against all of the

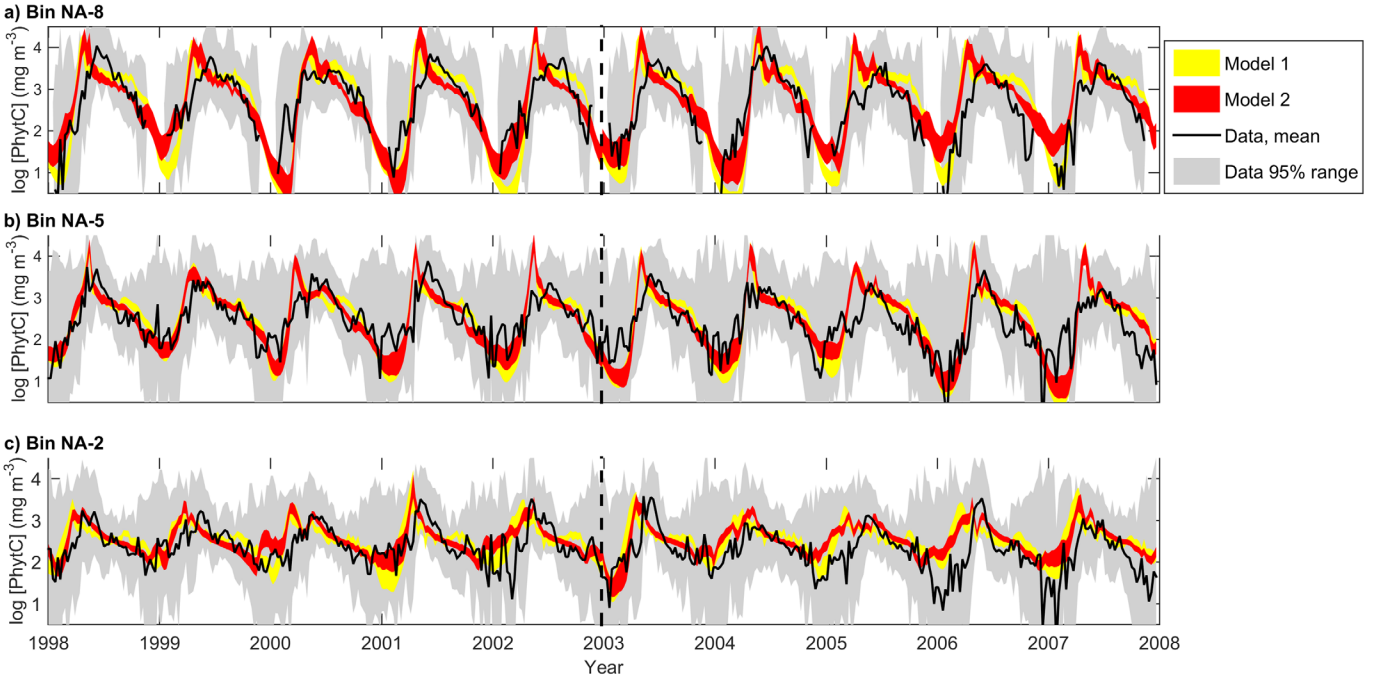

Figure 2. Simulated and observed phytoplankton concentrations for three North Atlantic bins and for both model training (1998-end of 2002) and test (2003-end of 2007) data. Simulated trajectories are distributions resulting from 1000 simulations, each using draws from the joint posterior parameter distribution from the data-constrained models. Training and test data simulations were run independently. The data 95% ranges are the 95% confidence intervals computed from spatial variation in the data within each of the nine North Atlantic bins identified in Figure 1.

data for all of the bins together (the log likelihoods for each bin were summed at each MCMC step to calculate the overall log likelihood), we refer to these models as having bin-shared parameters. We also partitioned all of our time series data into training and evaluation sets, selecting the years 1998–2002 as the training years and 2003–2008 as the testing years. This was important to assess how well the results generalized to data that the model had not been trained to and to guard against overfitting our models.

2.6. Definition of Parameter Ranges

Our parameter definitions (to equations (1)–(5) and defined in Table 1) were, where possible, the same as in the original EP85 model formulation, though with some being made redundant by having environmental data (e.g., sea surface irradiance), and some new parameters being added (such as separate nitrate and phytoplankton diffusion rates). EP85 gave a list of parameter values and their justifications for deciding those values. We inferred the most likely parameter values given the observed data, with the limitation that they must fall within biologically realistic ranges. A literature survey indicated that it would be difficult to specify meaningful prior parameter ranges for most parameters and so we chose to base our prior ranges on the values adopted by EP85, setting the ranges between 0.25% and 400% of those used in EP85 (this choice is addressed in section 4).

3. Results

3.1. How Well Do the Models Predict the Data?

After parameter estimation, the predictions of both models closely followed variation in the mean of the time series data on phytoplankton concentrations and euphotic zone depths, when assessed against both the training data (1998–2002) and the evaluation data (2003–2007; see Figures 2 and 3 for the simulated trajectories for three bins and supporting information Figures S1 and S2 for trajectories for all bins). Around 40% of the variation in the data on phytoplankton concentrations was explained by Model 1 and around 50% by Model 2 although these performance metrics overlap in their standard deviations; these numbers calculated by averaging across the training and test data performance metrics, (based on the correlation coefficients between predictions and observations for [P] in Table 2 and supporting information Table S1;

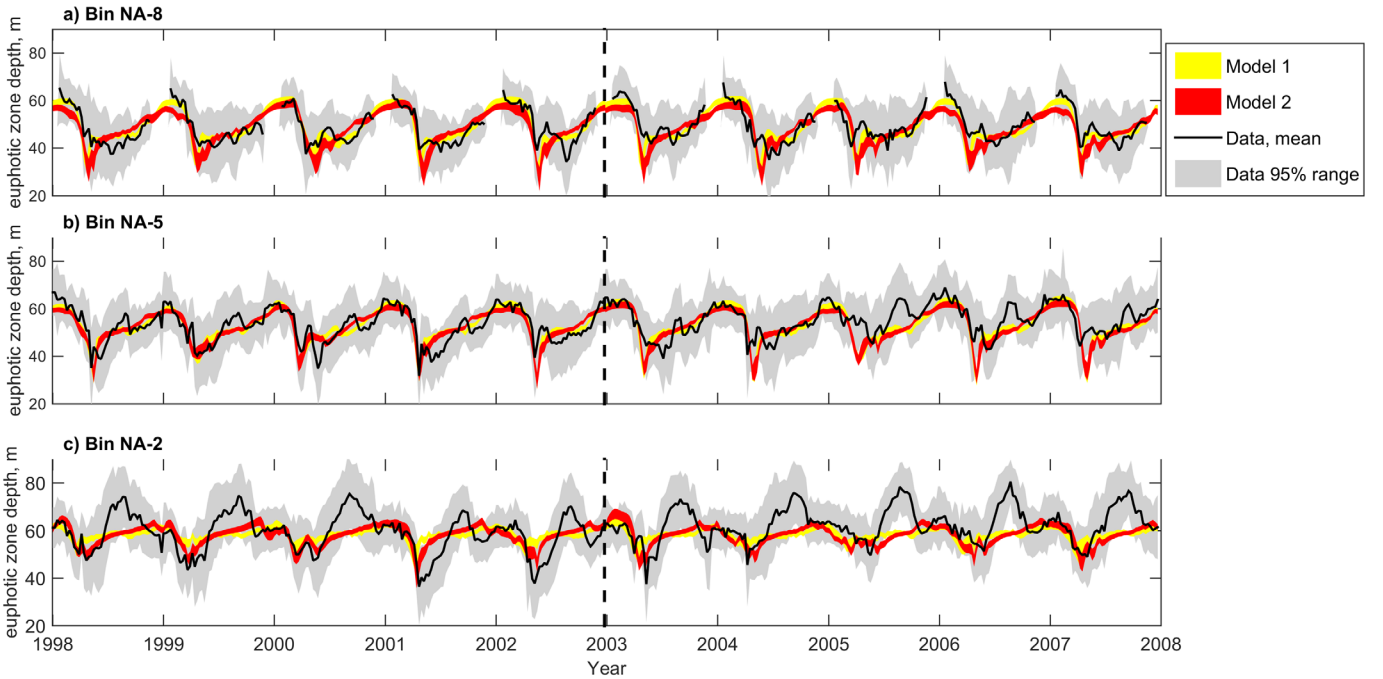

Figure 3. Simulated and observed euphotic zone depths for three North Atlantic bins and for both model training (1998-end of 2002) and test (2003-end of 2007) data. Simulated trajectories are distributions resulting from 1000 simulations, each using draws from the joint posterior parameter distribution from the data-constrained models. Training and test data simulations were run independently. The data 95% ranges are the 95% confidence intervals computed from spatial variation in the data within each of the nine North Atlantic bins identified in Figure 1.

Table 2 shows results for the test data, training data results are given in supporting information Table S1). Predictions of euphotic zone depths explained around 37% of the variation in the empirical mean data for Model 1 and around 41% for Model 2 (based on correlation coefficients for M_d in Table 2 and supporting information Table S1; again, overlapping standard deviations). In contrast, neither model predictions correlated with the mean time series of the rate of change of phytoplankton concentrations (d[P]/dt; Table 2 and supporting information Table S1 and Figure S3). The large uncertainty relative to the mean empirical data trajectory indicated that there was particularly low confidence in the high-frequency fluctuations in mean population growth rate. However, Model 2 tended to explain over twice as much variation in the rate of

Table 2. Model Predictive Performance for Test Data (Years 2003–2008) After Parameter Estimation Had Been Performed Using the Training Data (Years 1998–2002)[a]

		Model 1		Model 2	
	Model Parameter Setup	Bin-Specific Parameters	Bin-Shared Parameters	Bin-Specific Parameters	Bin-Shared Parameters
Correlation coefficient (R^2) between predictions and observations for. . .	[P]	0.40 (0.18)	0.42 (0.09)	0.45 (0.16)	0.54 (0.04)
	d[P]/dt	0.01 (0.01)	0.01 (0.1)	0.01 (0.01)	0.01 (0.01)
	M_d	0.39 (0.21)	0.33 (0.17)	0.43 (0.17)	0.37 (0.16)
	Phyt. pop. growth rate	0.03 (0.03)	0.03 (0.02)	0.09 (0.05)	0.10 (0.04)
	Avg. annual [P]	0.58 (0.20)	0.64 (0.08)	0.66 (0.19)	0.77 (0.03)
	Avg. annual phyt. pop. growth rate	0.08 (0.06)	0.07 (0.04)	0.29 (0.13)	0.35 (0.14)
Mean relative likelihood for. . .(larger = better)	All data	−145	−219	−135	−195
	[P]	−54	−65	−40	−44
	d[P]/dt	−31	−31	−31	−31
	M_d	−63	−126	−67	−124

[a]Parameter estimation was either performed fitting the models to data from individual bins or all bins together. Mean relative likelihoods are the average differences in log likelihoods across individual bins relative to the model that had the highest likelihood for each bin (Model 2 with bin-specific parameters for the training data). Correlation coefficients were either calculated by comparing predicted versus observed time series (n = 230 for each bin) or predicted versus observed for the mean annual time series for each bin (n = 48 for each bin). See supporting information Table S1 for model performance assessed against the training data, including information criteria scores.

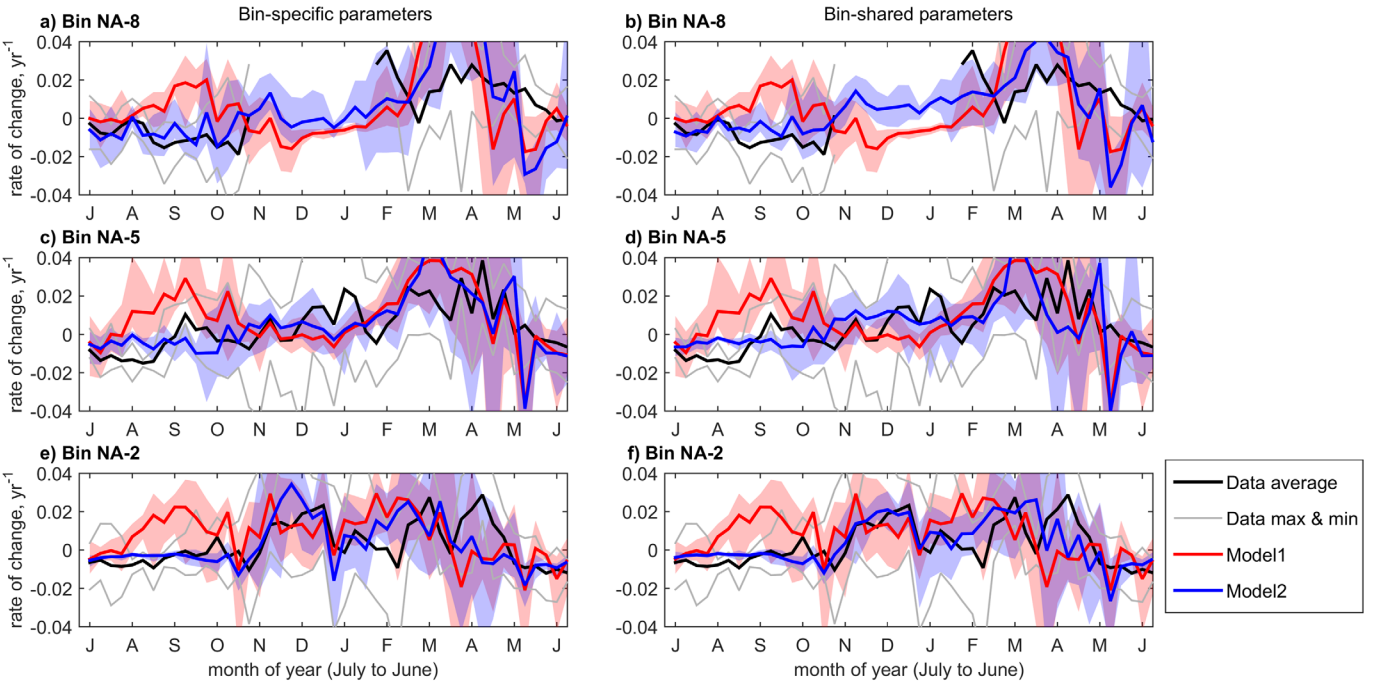

Figure 4. Predicted and observed mean annual dynamics for the rates of change of phytoplankton populations for North Atlantic bins 2, 5, and 8 for the years 1998–2002, (training data, although the same results were observed for the testing data). Solid lines for Models 1 and 2 indicate mean values while the shading indicates the 95% confidence range arising from 1000 simulations, each using draws from the joint posterior parameter distribution from the data-constrained models either with bin-specific or bin-shared parameters. The main differences between Models 1 and 2 are the rates of change predicted in the autumn (August–October).

change of phytoplankton populations (around 10% for phytoplankton population growth rate in Table 2 and supporting information Table S1) compared to Model 1 (around 4%), with little overlap in their standard deviations.

In terms of average annual phytoplankton dynamics, Model 2 explains around 70% of the variation in phytoplankton concentration dynamics compared to around 60% for Model 1, with overlapping standard deviations (average annual [P] in Table 2 and supporting information Table S1). However, Model 2 explained around 30% of the variation in the rate of change of phytoplankton populations compared to 8% for Model 1 (average annual phytoplankton population growth rate in Table 2 and supporting information Table S1), with no overlapping standard deviations. Therefore, Model 2 tended to perform better than Model 1 in terms of capturing the dynamics of phytoplankton concentrations and euphotic zone depths and also in terms of predicting the rates of change of phytoplankton populations. We conclude that from the perspective of making predictions that best fit the data, Model 2 is better than Model 1. Analysis of the model likelihoods and information criteria supports this conclusion (Table 2 and supporting information Table S1).

Comparing the predictions of the model with bin-specific and bin-shared parameters shows that the models with bin-shared parameters predict the data with much lower uncertainty (evident from comparing the standard deviations in Table 2). This is an expected consequence of using more data; the data from all bins rather than from just one bin, to infer the parameters of the bin-shared model. In terms of predictive performance, the models with bin-shared parameters and bin-specific parameters predict a similar amount of variation in all of the data sets (Table 2), with overlapping standard deviations. However, comparison of the model likelihoods shows that the fit of the data is better for the models with bin-specific parameters (Table 2).

3.2. What Do the Models Predict for the Seasonal Dynamics of Phytoplankton Population Growth Rates?

The model performance analysis above highlights that Model 2 particularly outperforms Model 1 in predicting the seasonal dynamics of the phytoplankton population growth rates. Graphical comparison of the average seasonal dynamics of these highlights how Model 2 captures the data better than Model 1 (Figure 4;

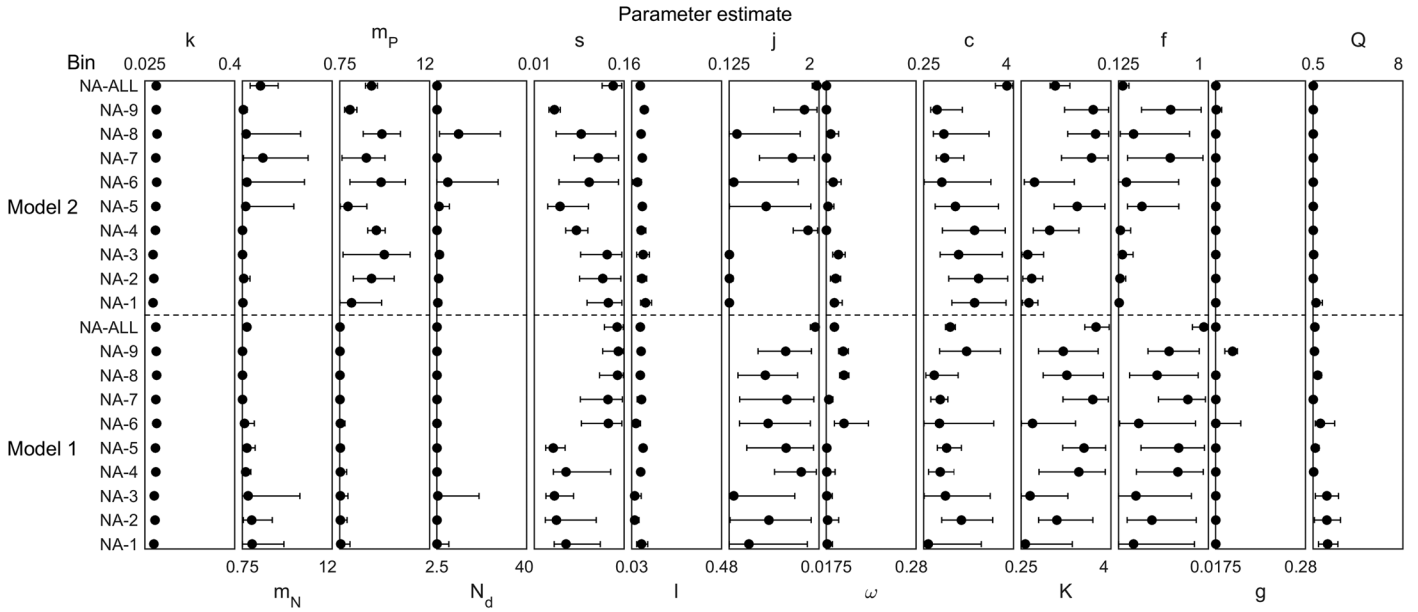

Figure 5. Posterior parameter distributions, showing means (dots), 2.5th and 97.5th percentile (whiskers). Bins labeled "NA-ALL" show the posterior distributions when assuming all bins shared the same parameters, the others being the distributions assuming bin-specific parameters. See Table 1 for parameter definitions and Figure 1 for the bin locations. The horizontal axis extents define the allowable range for parameter variation during parameter inference (the prior range).

results for all bins with bin-specific parameters are shown in supporting information Figure S4 and with bin-shared parameters in supporting information Figure S5). As shown by *Behrenfeld* [2010], empirical data (black lines in Figure 4) indicate that the switch from negative to positive population growth rates associated with the blooms begins in the late autumn (approx. November) and tends to stay positive until the end of the bloom (approx. May). However, Model 1 predicted a positive but varying population growth rate throughout the year, while Model 2 tended to capture the observed seasonal timing correctly. This is also true when bin-shared parameters are used for both models (Figure 4).

Our model predictions also imply latitudinal variations in the patterns of phytoplankton population growth rates (Figure 4). In the most northern bins, the models indicate that the rates of change of phytoplankton populations tended to decline toward zero after becoming positive in the autumn, approaching zero in the winter. Unfortunately, satellite data for the northern bins with which to test this prediction do not exist, although it makes logical sense given that there is almost no light to support photosynthesis in the midwinter in the most northern bins. Further south, the midwinter decline in population growth rates is less prominent for both the empirical data and model predictions.

3.3. How Do the Inferred Parameters Differ Between Models and Between Bins?

Analysis of the inferred parameters for different bins and for different models aids in understanding why the different models make different predictions, and why their predictions vary spatially. However, the only inferred parameter that clearly differed between Models 1 and 2 was the rate of phytoplankton diffusion out of the mixed layer (m_P), which tended to be higher for Model 2 than for Model 1 (Figure 5). This is consistent with the fact that the major structural difference between the two models was in how mixed-layer deepening influenced dilution and diffusion rates, where diffusion occurred all the time in Model 1 but only applied in Model 2 when the mixed layer was shallowing and below the euphotic zone depth. This implies that the prevention of diffusion and dilution while the mixed layer was shallower than the euphotic zone depth enabled Model 2 to capture the increase in phytoplankton population growth rates in the late autumn.

The parameters controlling grazing rates; grazing efficiency (f), half saturation of grazing rate (K), and the maximum grazing rate (c), all had quite large posterior ranges for both models, but also showed evidence of varying with latitude. It is difficult to interpret the consequences of that variation for grazing rates from

Figure 5 alone and so we used the joint posterior of the parameters to reconstruct the resulting functional relationships between phytoplankton concentrations and grazing rate, $fP^2/(K+P^2)$. These showed that grazing rates tended to be slightly higher for Model 1 than for Model 2, with overlapping uncertainty distributions, but did not show clear variation with latitude for the estimates with bin-specific parameters (supporting information Figure S6).

The half saturation constant controlling nitrogen limitation on phytoplankton division rates j was lower for the most southerly bins than the northerly bins in both models. This is connected to our finding reported below that nitrate limitation is particularly strong in the most southerly bins (at least for Model 2) and so we defer further interpretation of this pattern until then.

The parameters controlling light decay through the water column (parameters k and l in Figure 5), the maximum photosynthetic rate (Q), and the background mortality rates of phytoplankton and zooplankton (ω and g) were all constrained to similar values for both models. This is also true for the inferred nitrate concentration below the mixed layer (N_d) and the nitrate diffusion rate out of the mixed layer (m_N).

3.4. How Do the Mechanisms Controlling Phytoplankton Population Growth Rates Vary Seasonally?

We now analyze the relative importance of changes in the rates of phytoplankton growth, grazing, dilution, and diffusion out of the mixed layer, the only components contributing to changes in phytoplankton population growth rates in our model, to investigate what causes the dynamics of the rate of change of phytoplankton populations. In light of the fact that our structural modification clearly improves the model performance (Table 2), we omit further analysis of Model 1 and report only on what Model 2 implies about the mechanisms underpinning the phytoplankton blooms.

Phytoplankton concentrations, in both predictions and empirical data, tend to decline throughout the latter half of the year and begin to increase again between January and February (Figures 6a–6c). The components contributing to changes in the population growth rates of phytoplankton also vary seasonally (Figures 6d–6f), driven by the seasonal variation in the mixed-layer depth and light intensity (Figures 6g–6i). Phytoplankton growth and grazing rates track each other closely for most of the year but diverge toward late winter and spring (Figures 6d–6f), associated with the period in which the mixed layer deepens below the euphotic zone depth (Figures 6g–6i) and positive dilution rates (Figures 6d–6f). Losses due to diffusion out of the mixed layer only occur from approximately January through May (Figures 6d–6f) as the mixed layer shallows while it is deeper than the euphotic zone depth (Figures 6g–6i).

To analyze the relative magnitudes of the factors controlling the phytoplankton population growth rates, we normalize them relative to the maximum possible phytoplankton growth rate given the light intensity α (Figures 6j–6l; i.e., if nitrate is not limiting). These show for all bins that grazing rates closely track phytoplankton growth rates from May to around October but then diverge, initially declining then increasing between October and May. These periods of positive difference between phytoplankton growth and grazing rates are associated with the period in which the rate of change of the phytoplankton population is positive (Figures 6m–6o). Therefore, it appears as though the initiation of the bloom is attributable to the decoupling of the dynamics of phytoplankton growth rates and grazing rates. These findings are also true if we analyze the predictions of the model with bin-shared parameters (supporting information Figure S7). However, nitrate limitation on phytoplankton growth rates also decreases during the early spring as more nitrate is entrained from deep water (Figures 6j–6l), which will also contribute to the increase in phytoplankton population growth rates.

Our model predicts differing strengths of nitrate limitation in different bins when the parameters have been inferred specifically for those bins, increasing from north to south (Figures 6j–6l). In the northernmost bins, the strength of nitrate limitation is relatively low throughout the year and does not decline substantially when the mixed layer deepens (Figure 6j). However, nitrate limitation does decline substantially in the more southerly bins (Figures 6k and 6l), which will inevitably contribute to the increase in the phytoplankton rate of change. It therefore appears as though the initiation of the plankton blooms in our model is associated with both a decrease in grazing rates and a relaxation of nitrate limitation, the latter being applicable in the more southerly bins. Diffusion through the mixed layer is predicted to decrease spring phytoplankton population growth rates similarly for all bins, but not to an extent that causes the phytoplankton population growth rates to go negative (Figures 6j–6l).

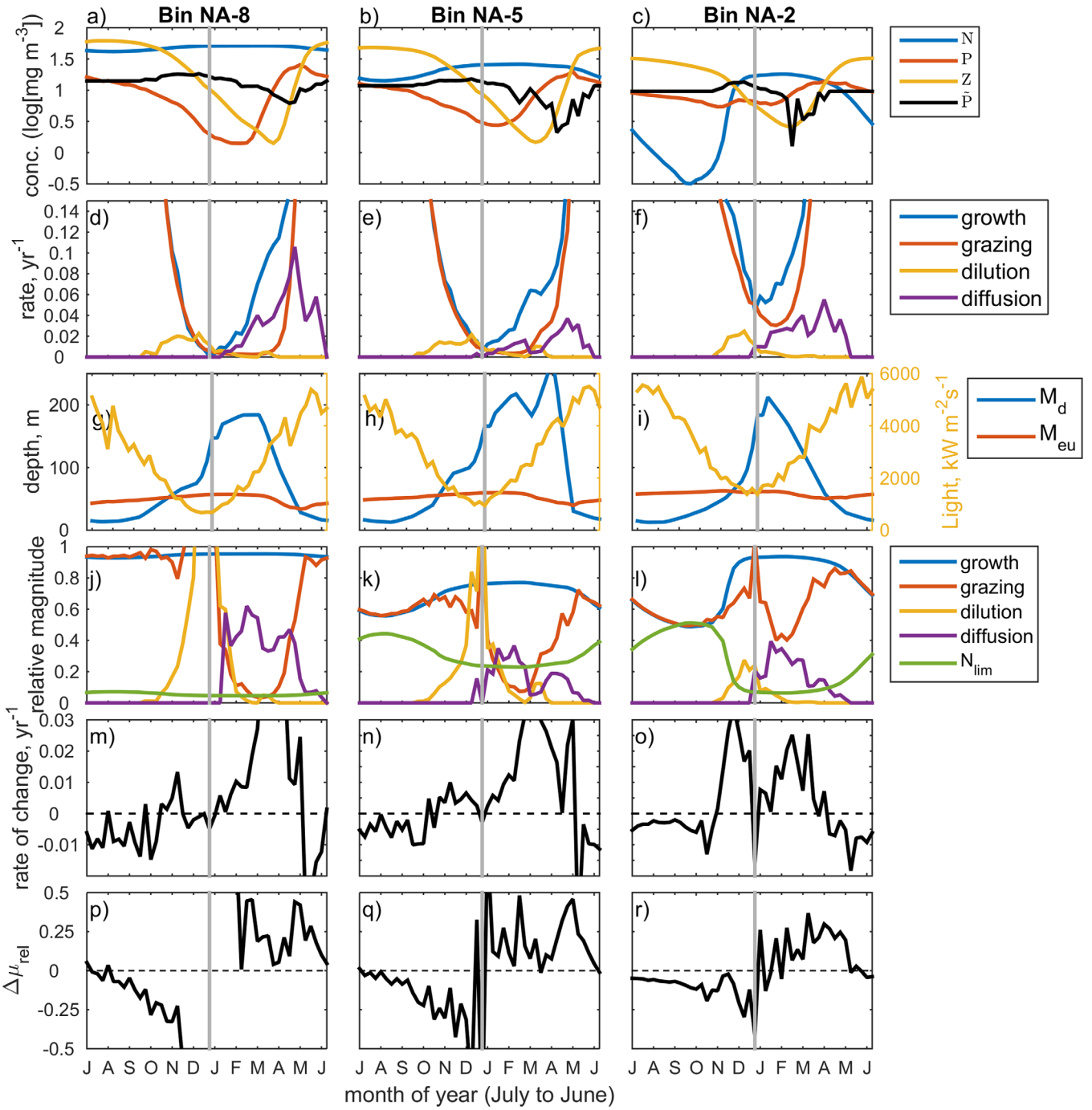

Figure 6. Mean annual variation in properties of the NPZ dynamics for three different bins. Dynamics shown are the median prediction from Model 2 with bin-specific parameters. (a–c) Dynamics of nitrate (N), phytoplankton (P), and zooplankton (Z), as well as the nullcline ($\tilde{P}$): the phytoplankton concentration at which the rate of change of zooplankton is zero (defined in equation (8)). (d–f) Components contributing to the rate of change of phytoplankton. (g–i) Mean seasonal dynamics of mixed-layer depth (M_d), euphotic zone depth M_{eu}, and light intensity. (j–l) Rate components normalized by nutrient unlimited phytoplankton growth rate, where N_{lim} represents the reduction of maximum growth rate by nitrate limitation. (m–o) Rates of change of phytoplankton populations. (p–r) Relative change in the phytoplankton growth rates over each 8 day time window as defined in equation (7). Gray vertical line in all plots indicates the winter solstice.

The blooms appear to be initiated by the same mechanisms when the model parameters have been inferred for all bins, with the exception that the dynamics of the different factors contributing to the rate of change of phytoplankton populations do not differ substantially with latitude (supporting information Figure S7). This latter finding implies that the differences in relative strengths of the rate controlling factors observed with our model with bin-specific parameters are due to inferred parameter differences. The lower

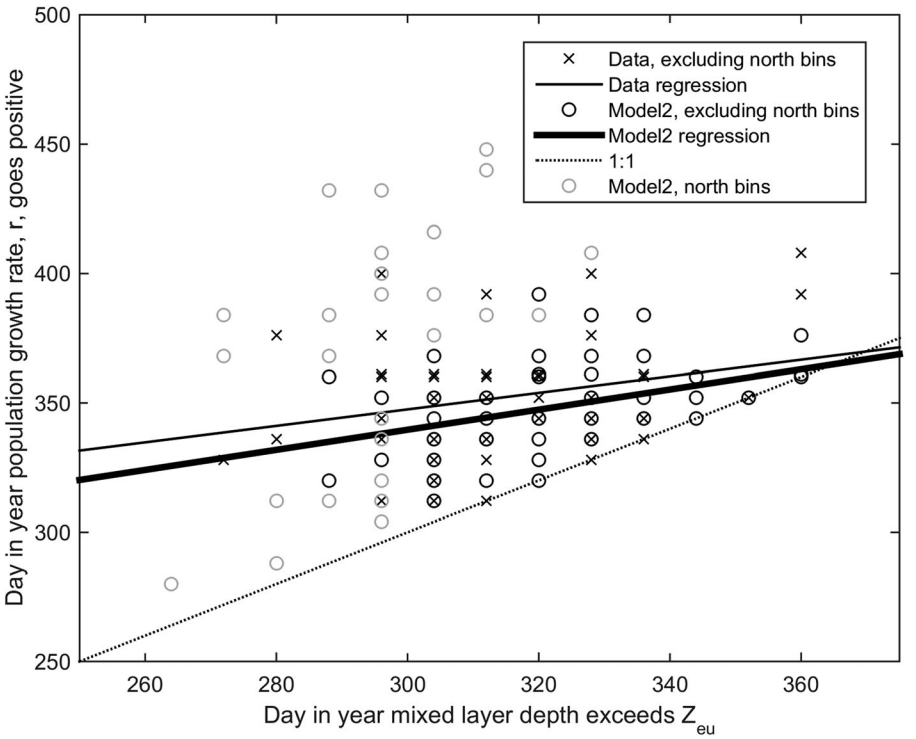

Figure 7. Relationship between days in the year that mixed-layer depth exceeds euphotic zone depth and days in the year that phytoplankton population growth rates go positive, marking the initiation of the phytoplankton bloom, for both empirical data and model predictions. We applied a 120 day moving average to smooth the time series of phytoplankton population growth rate to enable detection of when it switched from negative to positive. The linear regressions are significant with $p < 0.05$ and both exclude the data from the northernmost bins.

saturation constants *j* inferred for the most southerly bins (Figure 5), as reported above, are probably related to the fact that nitrate limitation is predicted to be relatively high in those bins. Lower values of *j* would allow higher growth rates at low nitrate concentrations and the nitrate concentrations in the most southerly bins are predicted to get lower than for the more northerly bins (Figures 6j–6l). This is likely because of the relatively low replenishment from mixed-layer deepening; dilution rates are lower in the most southerly bins (Figure 6l) and nitrate concentrations are not replenished as much as for the more northerly bins (Figures 6a–6c). Thus, it appears as though lower values of *j* are inferred to enable predicted phytoplankton concentrations to match the data when nitrate concentrations are predicted to be limiting in the most southerly bins.

3.5. Does Temporal Variation in Mixed-Layer Deepening Cause Temporal Variation in Bloom Initiation?

The analysis above implies that dilution-induced reductions of grazing pressure and nutrient limitation on phytoplankton growth rates are associated with the initiation of the bloom in the late autumn/winter. This leads to the prediction that the timing of bloom initiation should be related to the timing of mixed-layer deepening below the euphotic zone depth. We found a significant positive relationship between the timing of bloom initiation (defined as the time at which the growth rate of the phytoplankton population goes positive) and the time at which the mixed-layer depth exceeds the euphotic zone depth in both the data and in our model predictions (Figure 7; significant at $p < 0.05$), although these relationships have low predictive power ($r^2 = 0.07$ for the data, $r^2 = 0.13$ for the model). In addition, the deepening of the mixed layer below the euphotic zone depth always precedes bloom initiation (Figure 7). Difficulties in obtaining a smooth time series of population growth rates for the northernmost bins owing to the gaps in the data meant that we had to omit that data from this analysis. In addition, predicted population growth rates do not consistently change from negative to positive at the onset of dilution in northernmost bins implying that a more sensitive method is required to detect the bloom onset.

3.6. What Sustains the Blooms?

Analysis of the rates contributing to the population growth rate indicates that while phytoplankton growth rates increase after midwinter due to increasing light intensity, grazing rates continue to decline before increasing again in the early spring (Figures 6d–6l). The increases in phytoplankton growth rates are shortly followed by increases in phytoplankton concentrations in the mixed layer, corresponding to when the mixed-layer depth stops deepening and begins to shallow. However, the zooplankton concentrations do not begin to recover until 1–2 months after this time. This implies that there is a period of time in which the phytoplankton concentrations are too low to sustain positive growth in zooplankton populations, confirmed by analysis of the dynamics of the nullcline, as defined by equation (8) (Figures 6a–6c). Analysis of the nullcline changes over time in different bins indicates that zooplankton are actually predicted to experience positive population growth rates for a short time period in the year; typically around 4 months (Figures 6a–6c).

Behrenfeld [2014] hypothesized that a sustained acceleration in phytoplankton growth rates after the winter solstice could be the mechanism that could underpin a prolonged period of growth for the phytoplankton population throughout the spring. Computing this acceleration according to equation (7) highlights that our model indeed predicts that growth rates accelerate during this period (Figures 6p–6r). However, the fact that that the phytoplankton concentrations are insufficient to support the population growth of zooplankton in late winter and early spring means that zooplankton concentrations could not "catch up" with phytoplankton in the early spring even if the acceleration of phytoplankton growth rates was zero. Therefore, a period of acceleration of phytoplankton growth rates does not appear necessary to explain the period of sustained phytoplankton population growth. Moreover, the period of rapid decrease in the phytoplankton population growth rate at the end of the bloom in May to July is associated with a rapid increase in the grazing rate while the phytoplankton growth rate is still accelerating, highlighting that an accelerating growth rate alone is insufficient to prevent zooplankton from ending the bloom (Figure 6).

3.7. Are Zooplankton, Nitrate Limitation, and Seasonal Mixed Layers Necessary for the Timing of the Blooms?

Our analysis above has revealed how time-varying processes contribute to bloom formation in our model. It is natural to then ask about which of these processes might be necessary or sufficient to generate phytoplankton blooms more generally. We address that question here by sequentially removing rate controlling processes and observing the impact on predicted bloom dynamics. We performed the knock-out experiments for Model 2 for various bins, and with model parameters inferred specifically for each bin or shared between bins, and our qualitative conclusions were the same. However, in Figure 8 we only show the results for Bin NA-5 with bin-specific parameters for brevity.

Only two of the model manipulations in which we changed single factors had a similar seasonal trend in the rate of change of phytoplankton populations to the original model. These were setting light to be constant and setting nitrate concentration such that it did not limit phytoplankton growth (Figures 8c and 8f compared to Figure 8b). These imply that seasonal light and nutrient limitation are not necessary conditions for the observed bloom dynamics (in terms of the rates of population change). Setting light to be constant had a more noticeable effect on the seasonal dynamics of phytoplankton concentrations, compared to the original model (Figure 8a, green versus black), in particular allowing increases in the average phytoplankton concentration even while the mixed layer is deepening (Figures 8b and 8c). When we knocked out seasonal MLD only, the onset of the plankton bloom coincided with the end of the winter solstice (Figure 8e), indicating that knocking out the effects of seasonal MLD removes the phase of positive population growth rates at the end of the year observed in the data.

Knocking out zooplankton entirely caused the phytoplankton concentrations to increase overall, both when it was the only manipulation and when it was combined with manipulating other factors (Figures 8a, 8d, 8g, and 8h). However, bloom dynamics were still observed, implying that phytoplankton blooms can occur through the interaction with physical forcing variables only, although the specific timings and trajectories are different. The simulation with no zooplankton and no seasonal MLD shows phytoplankton population growth rates roughly following the seasonal dynamics of light, being positive when light is increasing and negative when light is decreasing (Figure 8h). The simulation with no zooplankton and no seasonal light has positive phytoplankton population growth rates throughout the year (Figure 8g) except in January and

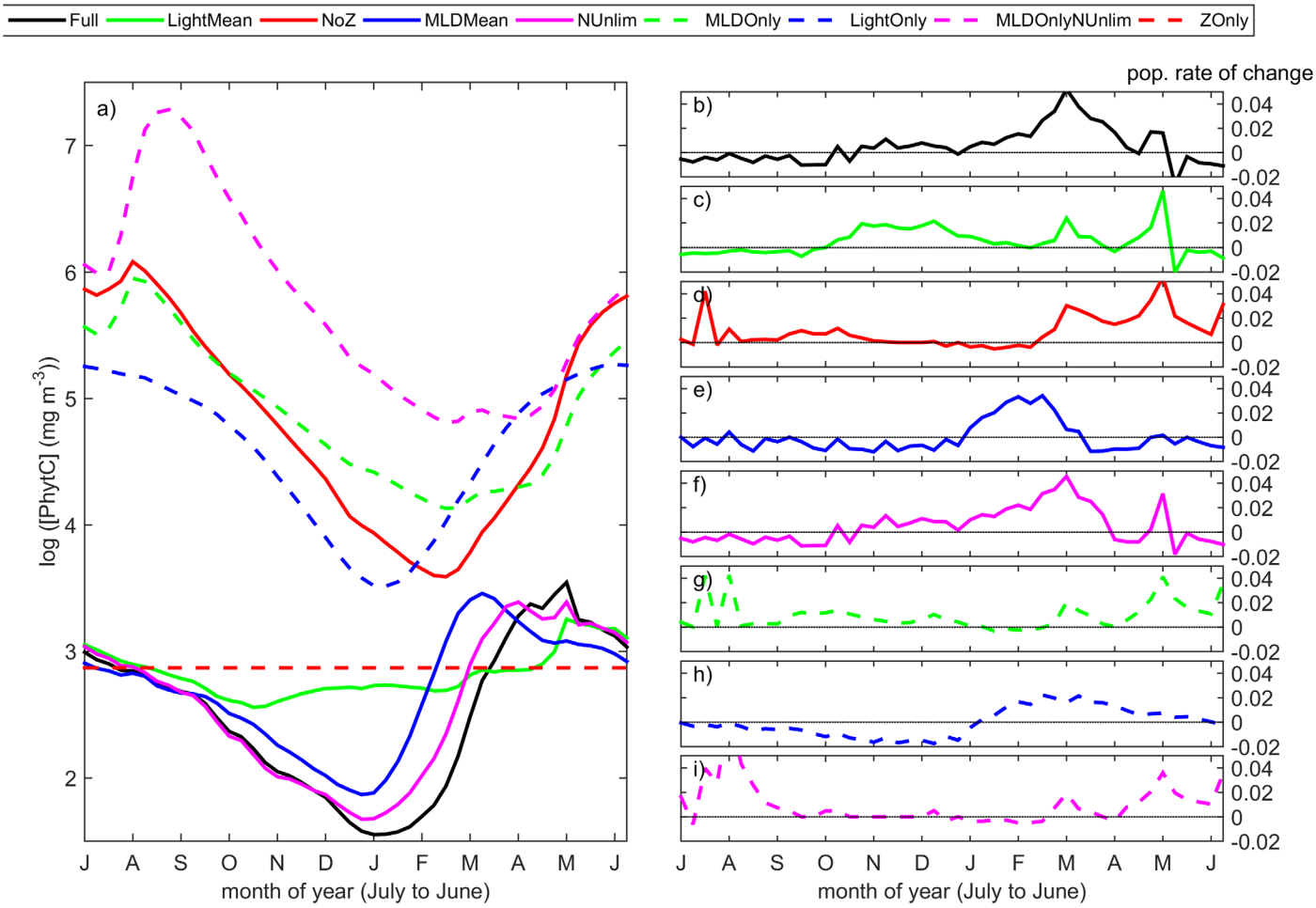

Figure 8. Knock-out experiments to assess the relative importance of different factors in controlling the trajectory and timing of phytoplankton concentrations and the rates of change of phytoplankton populations. Average-predicted seasonal dynamics of (left) phytoplankton concentrations and (right) rates of change of phytoplankton populations for bin NA-5 using Model 2 with the mean posterior parameter values and different manipulations of the driver data or initial conditions. Legend definitions are: LightMean—light intensity was set to the mean light level from the data for that bin, NoZ—zooplankton was fixed to zero, MLDMean—mixed-layer depth (MLD) was set to the mean MLD from the data for that bin, NUnlim—nitrate levels were set to be a constant value, sufficiently high to be unlimiting to phytoplankton growth rates, LightOnly—zooplankton was set to zero and MLD was set to the mean MLD, MLDOnly—zooplankton was set to zero and light was set to the mean light level, MLDOnlyNUnlim—as in MLDOnly but nitrate was also set to be unlimiting to phytoplankton growth rates, and ZOnly—MLD and Light were set to their mean levels.

February. The changes in phytoplankton population growth rates in this experiment could be controlled by changes in the integrated quantity of light in the mixed layer and in nutrient concentrations within the mixed layer. The high phytoplankton population growth rates in March to August coincide with a shallowing mixed layer, increasing the integrated quantity of light. The slightly positive population growth rates in the autumn and early winter may be due to either increasing light in the water column or reduced nutrient limitation but the former seems unlikely given that the mixed layer is deepening (increasing the time the phytoplankton spend at greater depths) and the phytoplankton population growth rates are positive. Thus, it appears as though the positive phytoplankton population growth rates are associated with a relaxation of nutrient limitation. To test this, we performed the same knock-out but without nitrate limitation (Figure 8i). This simulation had a zero phytoplankton population growth rate throughout the autumn and early winter, implying that the relaxation of nutrient limitation in the autumn and early winter due to mixed-layer deepening is a sufficient condition for positive phytoplankton population growth rates.

When physical forcing is knocked out entirely, the simulations do not show any seasonal dynamics: when zooplankton and phytoplankton are present together then they approach their stable equilibria (Figure 8a, ZOnly) and when phytoplankton is present on its own (not shown) it approaches an alternative stable equilibrium.

4. Discussion

4.1. Model Predictive Accuracy

Our results show that a simple data constrained model, slightly modified from that of *Evans and Parslow* [1985], can capture the qualitative and quantitative temporal and spatial dynamics of phytoplankton blooms in the North Atlantic (Figures 2 and 4, and 6 and Table 2). This extends the finding of *Kuhn et al*. [2015] that a simple data-constrained model could capture the seasonal average dynamics of phytoplankton, by showing that a simple model can also predict multiannual time series data.

The accuracy of our model at predicting the seasonal dynamics of phytoplankton, explaining around 66 and 77% of the variation of the training and evaluation data respectively, is slightly lower than that achieved by *Kuhn et al*. [2015] who explained around 85% for the more southerly bins. This difference is mostly likely to be due to a difference in our parameter inference techniques. The technique used by *Kuhn et al*. [2015] involved finding the single set of parameters that gave the best fit to the annually average seasonal dynamics of phytoplankton, weighting toward capturing the dynamics of the bloom by using a variable weighting scheme throughout the simulated year. We expect this to give a better fit to the annual average time series for several reasons; their model was trained to capture that specific time series whereas ours was trained to capture multiyear dynamics and for multiple data sets, their method did not take into account uncertainty in the data whereas ours did, and the ratio of the number of model parameters to be estimated to the number of data points is much higher in their study than ours (we use 5 years of data for three different data types to infer parameters).

Another reason for differences in predictive accuracy between our model and that of *Kuhn et al*. [2015] could be that their model was vertically resolved, which enables the more realistic simulation of vertical dynamics. More realistically representing the effects of mixed layer deepening on phytoplankton, zooplankton, and nutrients led to improved accuracy in our model. A clear extension of our study is therefore to investigate how much further improvement is made by explicitly representing the vertical dimension, a matter which we return to below. However, the fact that our simpler model does relatively well indicates that extending to the additional complexity of a vertically resolved model may not always be necessary in applications that still require a certain degree of qualitative and quantitative accuracy.

One reason for why our model might perform relatively well at capturing the time series of phytoplankton dynamics is because we could obtain time series data on mixed layer depths and light. The use of time series data on light contrasts with the study of *Kuhn et al*. [2015] who used a simple formula to estimate incident light. We did not, however, account for other important changes in environmental variables such as temperature (as was done by *Kuhn et al*. [2015]) but it would be straightforward to investigate the benefits of including such details in future.

Our study potentially increases the utility of the simple model of *Evans and Parslow* [1985] (albeit, with modifications), and other similarly simple models, from simply investigating and illustrating the hypothesized mechanisms driving phytoplankton dynamics to being useful for capturing quantitative dynamics when driven by physical forcing variables (as also implied by the study of *Kuhn et al*. [2015]). This supports the supposition by *Franks* [2002] that NPZ models should not necessarily be rejected outright for forward prediction (forecasting) simply on the basis that they appear too simple. These findings imply that simple data constrained models could be used to quantitatively forecast changes in phytoplankton dynamics and be useful for identifying possible reasons for anomalous changes in phytoplankton dynamics, such as under climate change [*Behrenfeld*, 2014]. While this would provide a model that can be analyzed more easily than more complex models, a disadvantage is that the model may not contain the richness of incorporated mechanisms necessary to explain the observations. However, within our Bayesian framework, models can be directly compared and assessed for their ability to predict data, and thus the relative importance of additional mechanisms quantified.

As it stands, our model can be used to investigate new relationships between phenomena it was designed to capture; the relationship between the timing of MLD deepening and the initiation of the bloom is a novel prediction that appears to be consistent with data (Figure 7), although has low predictive power. Our model can also be used to explore emergent relationships such as the inferred latitudinal variation in parameter values or dynamical properties. Like *Behrenfeld et al*. [2013], our model also predicts that the deeper the

maximum mixed-layer depth the higher the maximum phytoplankton population (concentrations integrated over depth), which is also consistent with the data (details omitted for brevity).

A partial reason why phytoplankton dynamics are predicted so well in our analysis is because the physical driving variables are assumed known. In genuine forecasting, these driving variables are also unknown, although can be predicted from physical models and climatological averages. Thus, the assessment of the forecast ability of a model should be made by also taking into account forecast uncertainty in the physical forcing variables. This does not mean that the predictability of phytoplankton dynamics will match that of the physical driving variables because it partly depends on how long the system retains some memory what happened in the past. For example, *Séférian et al*. [2015] illustrated that marine productivity can actually be forecasted further into the future than the physical variables such as sea surface temperature because their changes over time and space persist for longer.

4.2. Support for Mechanisms Underpinning Bloom Initiation

Our results add to the support of other recent studies that disturbance to ecosystem dynamics in the autumn caused by the deepening of the mixed layer can initiate the bloom [*Behrenfeld et al*., 2013; *Kuhn et al*., 2015; *Lévy*, 2015]. We selected the *Evans and Parslow* [1985] model because it had been proposed to exhibit phytoplankton bloom dynamics that were consistent with the disturbance-recovery hypothesis but had not yet been shown to capture the quantitative dynamics of phytoplankton in the North Atlantic; however, our results indicate that the original model is probably too simple to represent the system effectively enough to accurately predict the timing of the blooms. This appears to be because population growth rates in the original model needed to be higher to compensate for high diffusion rates out of the mixed layer during the autumn when the mixed layer is shallowest. This caused phytoplankton population growth rates to be positive throughout the late summer and autumn for that model (Figure 4). This was prevented in Model 2 by assuming that phytoplankton concentrations beneath the mixed layer were the same as above it when the mixed layer was shallower than the euphotic zone depth.

Bloom initiation in our model occurs due to the reduction in grazing rates and nutrient limitation on phytoplankton growth rates after the mixed layer increased above the euphotic zone depth, consistent with *Kuhn et al*. [2015]. However, unlike *Kuhn et al*. [2015] our knock-out experiments indicate that nutrient limitation is not a necessary condition for the observed timing of the onset of the phytoplankton blooms (Figure 8f), even in the southerly bins where nitrate limitation appears greatest. Instead, they imply that the presence of zooplankton and of seasonality in mixed-layer depths is necessary conditions in order to match the seasonal dynamics of phytoplankton population growth rates. It appears as though the effects of nutrient limitation in the study of *Kuhn et al*. [2015] had greater bottom-up control on phytoplankton population growth rates relative to grazing rates than we found. This could be a consequence of their use of a vertically resolved model, which enabled the simulation of the spatiotemporal dynamics of the different dynamical components, including nitrate, whereas we assumed a uniform nitrate concentration within the mixed layer and a constant nitrate concentration in the deep water below the mixed layer.

Taken together, the recent modeling investigations into phytoplankton bloom initiation are unanimous in recognizing the effects of multiple controls [*Behrenfeld et al*., 2013; *Kuhn et al*., 2015; *Lévy*, 2015, this study]. The analysis of *Lévy* [2015] illustrated how bloom formation in seasonally forced NPZ models could conform to both the critical turbulence hypothesis and the disturbance-recovery hypothesis. Key to that illustration was that the model was also vertically resolved, enabling seasonal changes in turbulence rates to be captured within the model. This argues for the importance of modeling the vertical mixing rates separately to enable the mixing layer (determined by turbulence) to be distinguished from the mixed layer (determined by temperature), as also argued by *Franks* [2014], which we did not do in our study. The model structure used in our study is not realistic enough to account for these effects of turbulence, which would require at least the modeling of the vertical dimension; nonetheless, we captured a significant proportion of variation in phytoplankton populations. Thus, while we have been able to show that our model, despite simple, is complex enough to accurately capture the data and provide support for the DRH, it probably is not complex enough to use to assess evidence for alternative competing hypotheses for the phytoplankton blooms using the same model structure. However, together, our study and those of *Kuhn et al*. [2015] and *Lévy* [2015] imply that it should now be possible to identify a model or set of models that can be constrained

using empirical data and rigorously assessed in terms of the relative support for different mechanisms underpinning phytoplankton blooms.

4.3. Insights Into Bloom Continuation

A key point made by *Behrenfeld and Boss* [2014] was that after bloom initiation, some mechanism is needed to prevent the zooplankton population from rapidly catching up with any increase in phytoplankton abundance to allow the bloom to persist over several months. *Behrenfeld* [2014] provides empirical evidence that this is due to the acceleration in the phytoplankton growth rate, which is supported by the theoretical studies of *Evans and Parslow* [1985] and *Truscott* [1995]. Our study highlights another potential component to that explanation that there can also exist a period in early spring in which the phytoplankton concentration is too diluted to sustain the zooplankton population, granting the phytoplankton population a period to grow while the zooplankton concentration is shrinking. Inspection of the study of *Evans and Parslow* [1985] highlights that this phenomenon is also present there and probably also in the simulations shown in *Behrenfeld and Boss* [2014]. *Behrenfeld et al*. [2013] do not show the predicted seasonal dynamics of zooplankton but do show predicted seasonal variation in grazing rates and these also match the predictions of our model (Figures 6d–6f, but especially the model with bin-shared parameters in supporting information Figures S7d–S7f) making us suspect that the same phenomenon exists in their model. Despite this, we are aware of no empirical evidence that supports our theoretical prediction although this is likely because no one has looked for it yet.

In our simulations, phytoplankton growth rates continue to accelerate even after the zooplankton population has begun to increase, which could act to prolong the blooms (Figures 6m–6o). However, the fact that the rate of change of the phytoplankton population decreases even while the phytoplankton growth rate is accelerating implies that the acceleration rate is in fact insufficient to prevent the zooplankton from catching up with the phytoplankton (Figures 6j–6l). Over this period diffusion rate out of the mixed layer is decreasing, which would help to maintain a positive population growth rate, and nitrate limitation is increasing, but not sufficiently to prevent the acceleration of the phytoplankton growth rate (compare Figures 6j–6l with Figures 6p–6r). Another influential mechanism could be the rapid shallowing of the mixed layer, which causes the zooplankton population to concentrate in a rapidly decreasing volume of water. At its most dramatic, this causes the zooplankton concentration to increase irrespective of the phytoplankton concentration, corresponding to when equation (8) has no positive solutions. This appears to be why zooplankton concentrations can recouple with phytoplankton concentrations even though phytoplankton growth rates are still accelerating.

4.4. Methodological Approach

A number of aspects of our methodology could be improved. Our method for quantifying uncertainty in the observational data only characterizes spatial heterogeneity observed across each environmental bin; however, uncertainty in the data arises from several other sources such as the accuracy with which the original algorithms used to compute the gridded data sets really capture the reported phenomena. Failure to account for such sources of uncertainty, a failure common to most other parameter inference studies in biology to our knowledge, could have biased our parameter estimates and so accounting for more uncertainty would increase the robustness of the findings. Our use of spatially averaged time series could also prevent important mechanisms underpinning the phytoplankton dynamics from becoming apparent, such as turbulence [*Franks*, 2014]. Thus, an important step in taking forward data-constrained modeling in this area is to account for these finer-scale horizontal spatial processes in the model. This would not necessarily make the model more complex in terms of numbers of parameters and modeled processes, but would make it more computationally demanding.

While our inference technique is conceptually similar to that employed by *Jones et al*. [2010] and *Dowd* [2011], both those studies take a different approach in how they account for uncertainty in the state variables (nutrients, phytoplankton, and zooplankton). They explicitly infer the state variables at each moment in time while accounting for observational uncertainty. An advantage of such an approach is that it reduces the effects that errors in the empirical data have on inferred model parameters, especially if such errors are biased, however a disadvantage is that such techniques can be computationally more demanding [*Palamara et al*., 2014]. Thus, a future improvement to our approach would be to improve how we take into account uncertainty in the system state.

Our study also highlights advantages of using Bayesian parameter inference over parameter optimization. The parameter optimization technique of *Kuhn et al*. [2015] resulted in a single set of parameters for each site. They then used using parameter perturbation techniques to assess how variation in those parameter values altered the consistency of the model predictions with data (as well as to obtain other insights). Using our technique, we can quantify directly how uncertainty in the data leads to uncertainty in our conclusions: about the model predictive accuracy and about the inferred parameter values. While our inference method could be improved (as discussed) the general technique of Bayesian parameter inference would provide a better way to compare multiple models while accounting for uncertainty than parameter optimization because it can explicitly take uncertainty into account.

A potential bias in our parameter estimates could have come from our specific selection of empirical data sets used to constrain the model parameters and assess model performance. The three data sets we used were ultimately all derived from the same remotely sensed data set. We were unsuccessful in obtaining more empirical constraints, both on other aspects of the system such as zooplankton concentrations, and those that could be used to define better priors for parameters. There is undoubtedly more empirical data that could be used in data-constrained modeling and we investigated the use of data from individual monitoring stations, databases and longitudinal surveys, in particular, to get more empirical constraints on zooplankton and nutrients. However, we found none that both clearly matched the spatiotemporal window of study and that could be easily used as data constraints.

Undoubtedly, the utilization of additional empirical constraints in future would improve confidence in the parameters and processes predicted using our method, both in a quantitative and qualitative sense: including more data will tend to lead to narrower probability distributions for the model parameters but more types of data would also improve confidence that the inferred parameter values and processes are indeed the best estimates given multiple sources of evidence. For example, we inferred several parameter values to be at one extreme of their prior range, as if their inferred values would occur beyond that range if we had let them (Figure 5). Placing stronger prior constraints would help to restrict more of these parameter values and generate greater confidence in the significance of the spatial variations remaining in the inferred parameters. One case in point is the parameter for the nitrate concentration below the mixed layer N_d, which we inferred to be at the minimum extent of its prior range (Figure 5). We hoped to prescribe this parameter from data on ocean nitrate concentrations but it was not clear which measurement to use, given that mixed layer depth varies seasonally, with latitude and with the vertical profile of nitrate varying through time throughout the water column. However, the inferred parameter of 2.5 m*M* nitrate m^{-3} appears too low given available data. *Kuhn et al*. [2015] adopted the approach of weakly nudging the dynamics of nitrate to approximately match the available climatological data set, which enabled the dynamics to be weakly constrained to match the observed climatology.

To enable the straightforward replication of our methodological approach, we conducted our study within a new framework for data-constrained modeling that allows models and empirical data analyses to be viewed and modified over a web-browser, as well as retaining a historical repository of changes (details given at http://research.microsoft.com/en-us/projects/msrceesdm/default.aspx). We propose that such modeling frameworks may facilitate and broaden the community of scientists doing similar research by enabling the simple comparison and sharing of models. This mirrors the intention of *Hemmings and Challenor* [2012] who developed a generic software tool to enable others to undertake data-constrained modeling using multiple empirical data sets. Ultimately, our intent is to facilitate multiple models to be constrained and assessed against empirical data sets within a common framework which allows for investigation both of hypotheses for ecological phenomena and the predictive performance of alternative models. Applying such an approach would constitute a strong assessment of present-day theoretical and empirical support for the DRH and alternative hypotheses.

5. Conclusions

Clearly, our simple NPZ model can capture both quantitative and qualitative dynamics of phytoplankton blooms in the North Atlantic. When our model does capture the dynamics of phytoplankton concentrations and population growth rates well, it does so using mechanisms that are consistent with the disturbance-recovery hypothesis of *Behrenfeld* [2010]. The key difference to previous work was our use of Bayesian

parameter inference techniques, which allowed us to more objectively identify a model parameterization given available empirical evidence. An important general next step will be to combine the methods illustrated in our study with insights from other recent modeling efforts to conduct a more comprehensive comparison of multiple models against empirical data to assess the relative abilities of alternative hypotheses to explain phytoplankton blooms in the North Atlantic and elsewhere. Thus, we hope that the key contribution of our study is laying the foundation for a comprehensive assessment of empirical and theoretical support for hypothesized mechanisms of phytoplankton blooms.

Acknowledgments

We thank Mike Behrenfeld for the extensive advice he gave us throughout this study, as well as providing the data that enabled us to begin this work. We thank Robert O'Malley for helpful advice in appropriate processing of the empirical data sets used in our study. We thank the editor Peter Brewer and two anonymous reviewers; their detailed and insightful comments improved our manuscript greatly. All of the data necessary to reproduce our findings are freely obtainable from the referenced sources detailed in section 2. The computational framework and complete computational experiment used to generate the results in this paper can be obtained from http://research.microsoft.com/en-us/projects/msrceesdm/default.aspx.

References

Anderson, T. R. (2005), Plankton functional type modelling: Running before we can walk?, *J. Plankton Res.*, *27*, 1073–1081.

Barton, A. D., M. Susan Lozier, and R. G. Williams (2015), Physical controls of variability in North Atlantic phytoplankton communities, *Limnol. Oceanogr. Methods*, *60*, 181–197.

Beaugrand, G., K. M. Brander, J. A. Lindley, S. Souissi, and P. C. Reid (2003), Plankton effect on cod recruitment in the North Sea, *Nature*, *426*, 661–664.

Behrenfeld, M. J. (2010), Abandoning Sverdrup's critical depth hypothesis on phytoplankton blooms, *Ecology*, *91*(4), 977–989.

Behrenfeld, M. J. (2014), Climate-mediated dance of the plankton, *Nat. Clim. Change*, *4*, 880–887.

Behrenfeld, M. J., and E. S. Boss (2014), Resurrecting the ecological underpinnings of ocean plankton blooms, *Annu. Rev. Mar. Sci.*, *6*, 167–194.

Behrenfeld, M. J., S. C. Doney, I. Lima, E. S. Boss, and D. A. Siegel (2013), Annual cycles of ecological disturbance and recovery underlying the subarctic Atlantic spring plankton bloom, *Global Biogeochem. Cycles*, *27*, 526–540, doi:10.1002/gbc.20050.

Berryman, A. A. (Ed.) (2002), *Population Cycles: The Case for Trophic Interactions*, 192 pp., Oxford Univ. Press, N. Y.

Boss, E., and M. Behrenfeld (2010), In situ evaluation of the initiation of the North Atlantic phytoplankton bloom, *Geophys. Res. Lett.*, *37*, L18603.

Brody, S. R., and M. S. Lozier (2014), Changes in dominant mixing length scales as a driver of subpolar phytoplankton bloom initiation in the North Atlantic, *Geophys. Res. Lett.*, *41*, 3197–3203, doi:10.1002/2014GL059707.

Brody, S. R., M. S. Lozier, and J. P. Dunne (2013), A comparison of methods to determine phytoplankton bloom initiation, *J. Geophys. Res. Oceans*, *118*, 2345–2357, doi:10.1002/jgrc.20167.

Chiswell, S. M. (2011), Annual cycles and spring blooms in phytoplankton: Don't abandon Sverdrup completely, *Mar. Ecol. Prog. Ser.*, *443*, 39–50.

Clancy, R. M., and W. D. Sadler (1992), The Fleet Numerical Oceanography Centre suite of oceanographic models and products, *Weather Forecasting*, *7*, 307–327.

Cullen, J. J. (2015), Subsurface chlorophyll maximum layers: Enduring enigma or mystery solved?, *Annu. Rev. Mar. Sci.*, *7*, 207–239.

Denman, K. L. (2003), Modelling planktonic ecosystems: Parameterizing complexity, *Prog. Oceanogr.*, *57*, 429–452.

Dowd, M. (2011), Estimating parameters for a stochastic dynamic marine ecological system, *Environmetrics*, *22*, 501–515.

Evans, G. T., and J. S. Parslow (1985), A model of annual plankton cycles, *Biol. Oceanogr.*, *3*, 327–347.

Franks, P. J. S. (2002), NPZ models of plankton dynamics: Their construction, coupling to physics, and application, *J. Oceanogr.*, *58*, 379–387.

Franks, P. J. S. (2009), Planktonic ecosystem models: Perplexing parameterizations and a failure to fail, *J. Plankton Res.*, *31*(11), 1299–1306.

Franks, P. J. (2014), Has Sverdrup's critical depth hypothesis been tested? Mixed layers vs. turbulent layers, *ICES J. Mar. Sci.*, fu175.

Friedrichs, M. A. M., et al. (2007), Assessment of skill and portability in regional marine biogeochemical models: Role of multiple planktonic groups, *J. Geophys. Res.*, *112*, C08001, doi:10.1029/2006JC003852.

Gilks, W. R., S. Richardson, and D. J. Spiegelhalter (1996), *Markov Chain Monte Carlo in Practice*, Chapman and Hall, Boca Raton, Fla.

Hahn-Woernle, L., H. Dijkstra, and H. J. Van der Woerd (2014), Sensitivity of phytoplankton distributions to vertical mixing along a North Atlantic transect, *Ocean Sci.*, *10*, 993–1011.

Hemmings, J. C. P., and P. G. Challenor (2012), Addressing the impact of environmental uncertainty in plankton model calibration with a dedicated software system: The marine model optimization testbed (MarMOT 1.1 alpha), *Geosci. Model Dev.*, *5*, 471–498.

Huisman, J., P. van Oostveen, and F. Weissing (1999), Critical depth and critical turbulence: Two different mechanisms for the development of phytoplankton blooms, *Limnol. Oceanogr.*, *44*, 1781–1787.

Jones, E., J. Parslow, and L. Murray (2010), A Bayesian approach to state and parameter estimation in a phytoplankton-zooplankton model, *Aust. Meteorol. Oceanogr. J.*, *59*, 7–16.

Koeller, P., et al. (2009), Basin-scale coherence in phenology of shrimps and phytoplankton in the North Atlantic Ocean, *Science*, *324*, 791–793.

Kuhn, A. M., K. Fennel, and J. P. Mattern (2015), Model investigations of the North Atlantic spring bloom initiation, *Prog. Oceanogr.*, doi:10.1016/j.pocean.2015.07.004.

Lévy, M. (2015), Exploration of the critical depth hypothesis with a simple NPZ model, *ICES J. Mar. Sci.*, *72*(6), 1916–1925, doi:10.1093/icesjms/fsv016, in press.

Lindemann, C., and M. A. St. John (2014), A seasonal diary of phytoplankton in the North Atlantic, *Frontiers Mar. Sci.*, *1*, 1–7.

Liu, Y., and D. Scavia (2010), Analysis of the Chesapeake Bay hypoxia regime shift: Insights from two simple mechanistic models, *Estuaries Coasts*, *33*, 629–639.

Mahadevan, A., E. D'Asaro, C. Lee, and M. J. Perry (2012), Eddy-driven stratification initiates North Atlantic spring phytoplankton blooms, *Science*, *337*, 54–58.

Matear, R. J. (1995), Parameter optimization and analysis of ecosystem models using simulated annealing: A case study at station P, *J. Mar. Res.*, *53*, 571–607.

Morel, A., and J.-F. Berthon (1989), Surface pigments, algal biomass profiles, and potential production of the euphotic layer: Relationships reinvestigated in view of remote-sensing applications, *Limnol. Oceanogr.*, *34*, 1545–1562.

Palamara, G. M., D. Z. Childs, C. F. Clements, O. L. Petchey, M. Plebani, and M. J. Smith (2014), Inferring the temperature dependence of population parameters: The effects of experimental design and inference algorithm, *Ecol. Evol.*, *4*, 4736–4750.

Platt, T., G. White III, L. Zhai, S. Sathyendranath, and S. Roy (2009), The phenology of phytoplankton blooms: Ecosystem indicators from remote sensing, *Ecol. Modell.*, *220*, 3057–3069, doi:10.1016/j.ecolmodel.2008.11.022.

Redfield, A. C. (1934), On the proportions of organic derivations in sea water and their relation to the composition of plankton, in *James Johnstone Memorial Volume*, edited by R. J. Daniel, pp. 177–192, Univ. Press of Liverpool, Liverpool, U. K.

Séférian, R., L. Bopp, M. Gehlen, D. Swingedouw, J. Mignot, E. Guilyardi, and J. Servonnat (2014), Multiyear predictability of tropical marine productivity, *Proceedings of the National Academy of Sciences*, 201315855.

Siegel, D. A., K. O. Buesseler, S. C. Doney, S. F. Sailley, M. J. Behrenfeld, and P. W. Boyd (2014), Global assessment of ocean carbon export by combining satellite observations and food-web models, *Global Biogeochem. Cycles*, *28*, 181–196, doi:10.1002/2013GB004743.

Smayda, T. J. (1997), What is a bloom? A commentary, *Limnol. Oceanogr.*, *42*, 1132–1136.

Sverdrup, H. U. (1953), On conditions for the vernal blooming of phytoplankton, *J. Conseil*, *18*, 287–295.

Taylor, J., and R. Ferrari (2011), Shutdown of turbulent convection as a new criterion for the onset of spring phytoplankton blooms, *Limnol. Oceanogr. Methods*, *56*, 2293–2307, doi:10.4319/lo.2011.56.6.229.

Truscott, J. E. (1995), Environmental forcing of simple plankton models, *J. Plankton Res.*, *17*, 2207–2232.

Weisheimer, A., and T. N. Palmer (2014), On the reliability of seasonal climate forecasts, *J. R. Soc. Interface*, *11*(96), 20131162.

Westberry, T., M. J. Behrenfeld, D. A. Siegel, and E. Boss (2008), Carbon-based primary productivity modeling with vertically resolved photoacclimation, *Global Biogeochem. Cy.*, *22*(2).