An empirically based minimalistic model for 1 representing seasonal phytoplankton dynamics

Piltz, Sofia Helena; Hjorth, Poul G.; Varpe, Øystein

Published in:
Marine Ecology - Progress Series

Link to article, DOI:
10.3354/meps13237

Publication date:
2020

Document Version
Peer reviewed version

Citation (APA):
An empirically based minimalistic model for representing seasonal phytoplankton dynamics

Sofia H. Piltz (piltz@umich.edu) *, 1, 2, 3, Poul G. Hjorth2, and Øystein Varpe4, 5

1 National Institute of Aquatic Resources, Technical University of Denmark, Kemitorvet, Bygning 202, 2800 Kongens Lyngby, Denmark
2 Department of Applied Mathematics and Computer Science, Technical University of Denmark, Asmussens allé, Bygning 303B, 2800 Kongens Lyngby, Denmark
3 Department of Mathematics, University of Michigan, East Hall room 2074, 530 Church Street, Ann Arbor, Michigan, 48109-1043, USA
4 Department of Arctic Biology, University Centre in Svalbard, 9171 Longyearbyen, Norway
5 Akvaplan-niva, Fram Centre, 9296 Tromsø, Norway

4th December 2019

Short title: Minimalistic model for bloom dynamics

Abstract

Supported by chl-a satellite data in the North Atlantic (and phytoplankton division rate computed from therein), the disturbance-recovery hypothesis for the initiation of phytoplankton blooms posits that the change in chl-a concentration is proportional to the relative change in the phytoplankton division rate. Here, we use this hypothesis introduced by Behrenfeld as a principal model assumption and construct a non-autonomous ordinary differential equation model for seasonally varying chl-a concentration. Our quantitative comparison between model simulations and in situ measurements of chl-a and primary production collected from a Swedish fjord is two-fold: First, by using approximate Bayesian computation, we find distributions of values for the three model parameters that best describe the chl-a data. Then, we validate our model by comparing the simulated (not fitted)
division rate to the data on division rate. Our minimalistic model is able to capture (1) the yearly trend in the chl-a concentration, (2) the pattern of growth and decline in the phytoplankton division rate, and (3) the decreasing trend in the relative change of the division rate exhibited in the data for several individual years. Moreover, the modeling efficiency (which is a measure of the goodness of fit of a (deterministic) model and nearly identical to $R^2$ for statistical models) is positive (between 0.3 and 0.9 with an average of 0.63) for all 11 years included in this study. Thus, we conclude that the change in chl-a concentration being proportional to the relative change in the division rate is a possible explanation for the bloom dynamics in the Gullmar fjord. In addition, our work provides a simple and empirically based differential equation for representing yearly dynamics of primary production, e.g., for generating ecological hypotheses using models of other trophic levels.

**Keywords:** non-autonomous ordinary differential equations, primary production, chlorophyll-a, division rate, Gullmar fjord, approximate Bayesian computation inference

### 1 Introduction

A shared view of the governing environmental mechanisms that initiate a phytoplankton bloom is lacking. On the one hand, a bloom is considered to initiate when the mixed layer depth shoals after deep winter mixing and exceeds a critical value, at which the mixed-layer-integrated phytoplankton respiration equals that of production (Sverdrup 1953). This *critical depth hypothesis* has been widely tested against data (Sathyendranath et al. 2015), with both consistent (Obata et al. 1996, Siegel et al. 2002) and inconsistent (Obata et al. 1996, Backhaus et al. 2003, Behrenfeld 2010) empirical evidence. To account for blooms that are observed before stratification of the water masses (Townsend et al. 1992, Ellertsen 1993), the *critical turbulence hypothesis* relates bloom initiation to a critical level of turbulent diffusivity (Huisman et al. 1999, Taylor & Ferrari 2011). On the other hand, analyses of satellite data (McClain 2009) and in situ measurements suggest that phytoplankton biomass integrated over the water column can start accumulating during the deepest mixing depth (Behrenfeld 2010, Boss & Behrenfeld 2010, Behrenfeld et al. 2013). These observations have led to the *disturbance-
recovery hypothesis (Behrenfeld et al. 2013) according to which a bloom initiates when deep
winter mixing has decoupled the ecological interactions by diluting the grazer population to
such a low concentration that allows for the phytoplankton concentration in the mixed layer to
increase (Behrenfeld 2014). In this work, we aim to increase the current understanding of the
initiation and shape of phytoplankton blooms by using previous data analyses (that led to the
disturbance-recovery hypothesis) to construct an empirically based mathematical representa-
tion for seasonally varying chlorophyll concentration.

Previous studies (Evans & Parslow 1985, Behrenfeld & Boss 2014) suggest that under
equilibrium conditions (i.e., when the mixed-layer-integrated phytoplankton production equals
with that of losses) the rate of change of biomass depends on the current division rate. How-
ever, as has been demonstrated by the in situ iron experiments, the increase in phytoplankton
concentration in these artificially generated blooms is due to a sudden increase (from a low
value) in phytoplankton division rate, and is rapidly caught up by increased encounter rates
with the grazers of phytoplankton [see, e.g., (Coale et al. 1996, Tsuda et al. 2003, Boyd et al.
2004, 2000)]. Consequently, the period of rapid growth of phytoplankton ends before nutrients
become limiting because of the ecological feedback from the predators of phytoplankton. In
contrast to such a strong top-down regulation, the continuous rise in phytoplankton division rate
due to gradually improving spring environmental conditions is a key factor in naturally occur-
ring phytoplankton blooms with high peak concentrations (Evans & Parslow 1985, Behrenfeld
& Boss 2014). Therefore, the increasing (decreasing) phytoplankton concentrations would
be better explained by accelerating (decelerating) division rates than by high division rates
(Behrenfeld 2014). Moreover, this principle would not be specific to a certain location but
apply to phytoplankton blooms globally (Behrenfeld 2014).

In addition to conceptual modeling frameworks, there exist contrasting conclusions within
the mathematical models for bloom dynamics as concerns the importance of bottom-up and
top-down (i.e., predation by herbivorous plankton) regulation mechanisms. Namely, mathem-
atical modeling work suggests that both bottom-up (Evans & Parslow 1985, Huppert et al.
mechanisms regulate the onset of a plankton bloom. These models
and plankton functional type models are built from first principles for the change of nutrients,
phytoplankton and zooplankton populations, and they reproduce qualitatively (e.g., when com-
pared to data by eyeballing) annual/interannual patterns seen in the bloom data, including the
latitudinal gradient [i.e., in temperate zones, there is a spring and an autumn bloom each year,
whereas only one bloom appears in spring in the Arctic (Cushing 1959)]. However, in such
models the parameter space becomes quickly very large which makes parameter fitting chal-
lenging, if not impossible. In addition, a survey of complex biogeochemical models shows
that there is a lot of variability in the mathematical representation of physiological processes
and characterizations of different types of plankton groups (Shimoda & Arhonditsis 2016). As
a result, including complexity in the models in small steps, the need for general theory, and
an interdisciplinary approach have been suggested as the next steps towards better models for
phytoplankton blooms (Fisheer et al. 2014, Behrenfeld & Boss 2014, Shimoda & Arhonditsis
2016).

Here, we contribute to the development of models for phytoplankton blooms with predict-
ive power by trading off detailed model construction from physiological principles for a parse
parameter space (i.e., 3-dimensional). We have chosen a parsimonious approach in order to
maximize our understanding of the model behavior, minimize the number of parameters, and
to be able to quantify the agreement between model simulations and data. A quantitative com-
parison to data has been previously done, e.g., by Ardyna et al. (Ardyna et al. 2014), where
a bloom was described with a Gaussian function. Similarly, we use a general function (the
sine-function) to represent seasonal changes in the phytoplankton division rate. However, in
contrast to (Ardyna et al. 2014), our model for the phytoplankton concentration incorporates a
previously shown dependency between the division rate and chl-a concentration. This principal
model assumption, known as the disturbance-recovery hypothesis and suggested by satellite
data from the North Atlantic (Behrenfeld et al. 2013, Behrenfeld 2014, Behrenfeld & Boss
2014), states that because of the ecological feedback (i.e., changes in grazing, viral attack,
and other losses proportional to changes in phytoplankton division rates), the rate of change
of phytoplankton concentration is not proportional to the cell division rate but to the relative
change in the cell division rate. To test the hypothesis as a suggested general mechanism for phytoplankton bloom initiation, we incorporate a quantitative comparison of the model simulations to chlorophyll and division rate (computed from the primary production) data collected in situ [and published previously in (Tiselius et al. 2015)] from the Swedish west coast.

2 Methods

In what follows, we use the terms ‘phytoplankton population’ (p) and ‘chlorophyll concentration’ (chl) interchangeably, because the data that we have access to contains measurements of the latter, while the mathematical model we construct is not limited and can also be used to describe the dynamics of the former. Similarly, we have access to measurements of primary production from which we compute the phytoplankton division rate.

2.1 The model

2.1.1 Empirically based, non-autonomous differential equation model for phytoplankton concentration $p$

Investigations of the satellite data (National Aeronautics and Space Administration Ocean Biology 2019) collected from the subartic Atlantic ocean suggest that the rate of change in the chlorophyll concentration, $\Delta \text{chl}/\Delta t$, changes in proportion to the relative change, $\Delta \mu_{\text{rel}}$, rather than the absolute change in division rate [see, e.g., Figure 3 in (Behrenfeld 2014)]. In addition, during the subarctic Atlantic bloom, the rate of biomass accumulation is uncorrelated with the division rate [see, e.g., Figure 3 in (Behrenfeld & Boss 2014)]. Considering change in chlorophyll as a change in the natural logarithm of measured chlorophyll [see caption for Figure 3 in (Behrenfeld 2014)] and assuming that the relative change in the division rate can be computed as the average of two consecutive measurements [see Equation (1) in (Behrenfeld 2014)], the conclusion of the data analysis in (Behrenfeld 2014) can be written as:

$$\frac{\Delta \ln \text{chl}}{\Delta t} \propto \Delta \mu_{\text{rel}} = \frac{2(\mu(t) - \mu(t - 8))}{\mu(t) + \mu(t - 8)},$$

(1)
where $\Delta t = 8$ days is the resolution of the satellite data and $\mu(t)$ is the division rate at time $t$. In these studies, the chlorophyll data is obtained using satellite-based data sensor to detect surface chlorophyll [National Aeronautics and Space Administration Ocean Biology 2019]. The division rate is then computed from the chlorophyll measurements using the Vertically Generalized Productivity Model [Behrenfeld & Falkowski 1997].

To investigate the hypothesis that the rate of change of phytoplankton concentration is proportional to the relative change in the cell division rate, we use the disturbance-recovery hypothesis as our model assumption and construct a non-autonomous ordinary differential equation describing the dynamics of phytoplankton blooms. In continuous time, the empirical observation that the rate of change in the logarithm of chlorophyll concentration, $d\ln p/dt$, is proportional to the relative change in phytoplankton division rate, $\Delta \mu(t)$, reported in [Behrenfeld 2014] and shown in Equation (1), is written as follows:

\[
\frac{d\ln p}{dt} \propto \Delta \mu(t) .
\]  

(2)

Assuming a proportionality constant $\alpha$, this yields the following equation:

\[
\frac{d\ln p}{dt} = \alpha \Delta \mu(t) .
\]  

(3)

Using logarithmic differentiation (that is, $d\ln p/dt = (1/p)(dp/dt)$) and rearranging, we obtain our non-autonomous differential equation model for chlorophyll concentration $p$:

\[
\frac{dp}{dt} = \dot{p} = \alpha \Delta \mu(t) p .
\]  

(4)

2.1.2 Division rate, $\mu(t)$, and the relative change of the division rate, $\Delta \mu(t)$

The novelty and inner workings of the model in Equation (4) are found in the phytoplankton division rate $\mu(t)$ in the following way: to account for seasonally changing environmental conditions, we represent $\mu(t)$ with a sine-function

\[
\mu(t) = A (1 + \sin[f(t + \phi)]) ,
\]  

(5)
where \( A \) is the difference between a maximum and minimum division rate within a year (i.e.,
the amplitude of the seasonal forcing). We assume seasonal changes in the division rate, and
therefore, take \( f = 2\pi/(365 \text{ days}) \), i.e., the period of the oscillation in the division rate is
exactly one year. Parameter \( \phi \) in Equation (5) represents the day of the year when the division
rate exceeds (from below) half of its maximum value during the year (i.e., the phase shift of
the oscillation). Thus, the peak of the division rate occurs a quarter of a year later than the \( \phi \)'th
day of the year.

We follow the work done in (Behrenfeld 2014) and assume that the relative change in the
phytoplankton division rate can be computed as an average of the division rate at time \( t \) and
an earlier time point. In (Behrenfeld 2014), the time difference between these points is 8 days,
which is the time difference between consecutive satellite measurements. However, here we
consider a continuous-time system based on the results reported in (Behrenfeld 2014), and
therefore, we denote the time difference between two consecutive points with a time offset, \( \tau \).
As a result, the relative change in the division rate, \( \Delta \mu(t) \), is computed as

\[
\Delta \mu(t, \tau, \phi) = \left( \frac{\mu(t) - \mu(t - \tau)}{\mu(t) + \mu(t - \tau)} \right) = \frac{\mu(t)}{2} \left( 2 \sin[f(t + \phi)] - \sin[f(t - \tau + \phi)] \right) \left( \frac{\sin[f(t + \phi)] + \sin[f(t - \tau + \phi)]}{2} \right).
\]

(6)

Note that the relative change in the division rate given in Equation (6) does not depend on the
difference between the maximum and minimum division rate within a year (i.e., \( A \)). In addition,
when the time offset is equal to zero (i.e., for \( \tau = 0 \)), \( \Delta \mu(t) = 0/(2 + 2\sin[f(t + \phi)] = 0 \) and our
minimalistic model (4) predicts an equilibrium chlorophyll density throughout the year at the
chlorophyll density in the beginning of the year [i.e., at the initial value \( p(0) \)]. Furthermore,
for \( f = 2\pi/365 \) and a given \( \phi \) and \( \tau \), the relative change in the division rate in Equation (6) has
a period of 365 days (i.e., \( \Delta \mu \) repeats itself every 365 days). Therefore, the model prediction
for the chlorophyll concentration which is a solution to Equation (4) is also 365-periodic. We
collect our model variables and parameters in Table 1.

We have not found a closed solution \( p(t) \) to the differential equation in Equation (4). There-
fore, in what follows, we will simulate the model in (4) numerically. More specifically, we
compare a model simulation of chl-a concentration at a given day of the year to in situ meas-
urements of chl-a concentration at the same day and quantify the difference (by computing the
straight-line distance) between the two. We then use in situ data for primary production (col-
lected at the same time as the chlorophyll data) to first compute the division rate, and then to
investigate how well our model, which is fitted to the chl-a data (but not to the division rate
data), reproduces the patterns exhibited in the division rate data.
Table 1: Model variables and parameters with their values, units, and biological explanation.

<table>
<thead>
<tr>
<th>Name</th>
<th>Description/Value</th>
<th>Units (d=day)</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t$</td>
<td>Independent variable</td>
<td>d</td>
<td>Time</td>
</tr>
<tr>
<td>$p(t)$</td>
<td>Compared to the data</td>
<td>mg m$^{-3}$</td>
<td>Chlorophyll concentration at time $t$</td>
</tr>
<tr>
<td>$p(0)$</td>
<td>Fitted to the data</td>
<td>mg m$^{-3}$</td>
<td>Chlorophyll concentration at time $t = 0$</td>
</tr>
<tr>
<td>$\mu(t)$</td>
<td>Validated with the data</td>
<td>mg C m$^{-2}$ d$^{-1}$</td>
<td>Phytoplankton division rate at time $t$</td>
</tr>
<tr>
<td>$\Delta \mu(t)$</td>
<td>Validated with the data</td>
<td>Dimensionless</td>
<td>Relative change in the phytoplankton division rate</td>
</tr>
<tr>
<td>$\phi$</td>
<td>Fitted to the data</td>
<td>d</td>
<td>Number of days after the first day of the year when half of the maximum division rate is exceeded from below</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>Fitted to the data</td>
<td>d$^{-1}$</td>
<td>Coefficient of proportionality, i.e., denoting the amount to which chlorophyll concentration is proportional to the accelerating/decelerating division rates</td>
</tr>
<tr>
<td>$\tau$</td>
<td>Fitted to the data</td>
<td>d</td>
<td>Amount of time before $t$ that is used in the computation of the relative change in division rate [see Equation (6)], i.e., determines the width of the time offset to which the rate of change of chlorophyll concentration is proportional</td>
</tr>
<tr>
<td>$f$</td>
<td>$2\pi/365$</td>
<td>d$^{-1}$</td>
<td>Period of the oscillation in phytoplankton division rate</td>
</tr>
</tbody>
</table>
2.1.3 Preliminary model simulations

As an initial test of our minimalistic model, we solve Equation (4) numerically to simulate the model with an arbitrarily chosen parameter set, and thereby, to compare model predictions qualitatively to the results of the data analysis reported in (Behrenfeld 2014). Indeed, an example simulation of our phytoplankton model reproduces characteristics of phytoplankton blooms, such as,

I the beginning of the bloom and an increase in phytoplankton population do not necessarily coincide,

II the phytoplankton population can increase despite the fact that the division rate is decreasing,

III the phytoplankton concentration remains low even though the relative change in division rate is strongly negative,

all observed in the satellite data from the subarctic Atlantic and described in (Behrenfeld 2014) (see Figure 1). To investigate whether these results agree quantitatively with a different dataset, and to infer model parameters, we use the algorithm developed by Beaumont et al. in (Beaumont et al. 2009). In doing so, we quantify the difference between simulations of our model and the Gullmar fjord data.
Figure 1: Preliminary model simulations. (Solid) Numerical solution to the model equation $\dot{p} = \alpha \Delta \mu p$ [Equation (4)] for the chl-a concentration (in mg m$^{-3}$), $p$, (dashed) the phytoplankton division rate [in mg C m$^{-2}$ d$^{-1}$], $\mu$($t$) [see also Equation (5)], and (dash-dotted) the relative change in the phytoplankton division rate (dimensionless), $\Delta \mu$($t$) [see Equation (6)], using an example parameter set (left) $(\tau, \alpha, \phi, p(0)) = (330, 0.015, -90, 1)$ and (right) $(\tau, \alpha, \phi, p(0)) = (71, 0.15, -180, 1)$. The empirically based differential equation model for phytoplankton population reproduces patterns observed in the data such as, (I) the beginning of the bloom and an increase in phytoplankton population do not necessarily coincide, (II) the phytoplankton population can increase despite the fact that the division rate is decreasing, and (III) the phytoplankton concentration remains low even though the relative change in the division rate is strongly negative. These empirical observations are reported in (Behrenfeld 2014).
2.2 Gullmar fjord data

To test our model, which principal assumption is based on the satellite observations from the subartic Atlantic (Behrenfeld 2014), we compare model simulations to in situ measurements collected from the Gullmar fjord (in the Släggö time series) on the west coast of Sweden and previously reported in (Tiselius et al. 2015). The Gullmar fjord is located on the southwest coast of Sweden where the Skagerrak strait connects the North sea and the sea areas leading to the Baltic Sea. The fjord has a surface area of 536 km² and a maximum depth of 120m. The surface water in the mouth of the fjord is a mixture of local runoff and water currents in the area such as low-saline water from the Baltic sea and saline water from the North Sea. As a result, the salinity of the fjord may vary significantly. We have access to Släggö time series that consists of biweekly in situ measurements of primary production collected at the mouth of the fjord (58°N, 11°E) since 1985. The fjord is a marine reserve and the mouth of the Gullmar fjord is considered as a good representative of the open-sea ecosystems in the area. The fjord has no major local sources of sewage, however, it receives large-scale pollution from the seas in northern Europe and Sweden (Lindahl et al. 2009). Similarly to the data set from the subartic Atlantic (Behrenfeld & Boss 2014), the highest primary production shows no relation to the phytoplankton biomass in the Gullmar fjord (Tiselius et al. 2015). In addition, the coastal ecosystem in the fjord (which is categorized as an ecosystem with intermediate level of productivity) was shown to display top-down regulation by zooplankton on a seasonal scale (Tiselius et al. 2015).

As concerns the Gullmar fjord data in our comparison between model simulations in Section 3, we will consider the data for primary production [units of mg C m⁻² d⁻¹] and chl-a (units mg m⁻³) that were reported previously in (Tiselius et al. 2015). Primary production was measured using the ¹⁴C-technique for each sample that is collected at depths 0, 1, 2, 3, 4, 6, 8, and 10m. Then, these samples were incubated in glass bottles for 4 hours and carbon uptake was transformed into daily production by the light factor method (Tiselius et al. 2015). The data are the average of these 8 bottles. Similarly, we will consider the amount of chl-a extracted from a 100ml filtered water sample and the average of these measurements from depths 1-10m. To compute the division rate from the primary production data, we divide each primary production
measurement by the same-day measurement for chl-a, which we have first multiplied by 40mg
(using the 1:40 chlorophyll to carbon ratio) and 10m (because the data represents the average
for depths 0–10m). Here, we limit our study to 11 years between 1993 and 2012 included in
the Gullmar fjord data that have no missing data between consecutive measurements and when
both chl-a and primary production have been measured during the same day.

2.3 Parameter fitting using approximate Bayesian computation

Approximate Bayesian computation is a set of methods for approximate Bayesian inference
which can be used whenever sampling from the model is possible. We investigate the problem
of fitting the parameters of model (4) to Gullmar fjord data (see Section 2.2) through Bayesian
inference. More specifically, we have made an implementation in Matlab of an existing algo-

11

Let us consider a model simulation for the chlorophyll concentration at a given time $t$, that
is, the solution to the differential equation in Equation (4) with an initial value $p(0)$, which
we can denote by $x(t; \alpha, \tau, \phi)$. In addition, we represent the available measurement data on
chlorophyll concentration with a vector $x$. As described in Section 2.2, the data are measured
approximately biweekly, that is, at time instances $t_i$, where $i$ denotes the day of the measure-
ment. Thus, excluding any measurement errors, the underlying assumption in our parameter
fitting procedure is that the data would be represented by a set of model simulations for time
instances \( t_i \), i.e., by \( x_i = x(t_i; \alpha, \tau, \phi) \), for some unknown, true parameter values. To account for
the presence of measurement errors, we incorporate normally distributed noise in our model
simulation. We denote the standard deviation of the measurement noise with \( \sigma \) (and because
we do not know the variance of measurement errors in advance, we incorporate the estimation
of \( \sigma \) in our parameter fitting process). Thus, for given parameter values \((\alpha, \tau, \phi, \sigma)\), the model
simulation (i.e., \( x^* \)) is described for each datapoint (i.e., element-wise) as normally distributed
with a mean equal to the solution of the differential-equation model in (4) and variance \( \sigma^2 \) [i.e.,
\( x^*_i \sim \mathcal{N}(x(t_i; \alpha, \tau, \phi), \sigma^2) \)].

2.3.1 Prior distributions for \( \sigma, \alpha, \phi, p(0) \), and \( \tau \)

We assume that the estimated parameters are mutually independent and have known, finite
lower and upper bounds. In the Bayesian framework, this information is described by inde-
pendent uniform probability densities. Consequently, we choose the upper bound for the stand-
ard deviation of the measurement noise to be \( \sigma = 0.1 \) [i.e., \( \sigma \sim \mathcal{U}(0, 0.1) \)]. As concerns the
proportionality coefficient \( \alpha \) (which can be considered as the intrinsic phytoplankton growth
rate, and which we therefore assume to be always positive), we choose 5 d\(^{-1}\) as the upper
bound for \( \alpha \) [i.e., \( \alpha \sim \mathcal{U}(0, 5) \)]. We choose this upper bound for \( \alpha \) by running several test
simulations which show that larger values of \( \alpha \) do not decrease the discrepancy between the
model simulation and chl-a data.

In the case of \( \phi \), we study the division rate in the Gullmar fjord data set in order to determine
the lower and upper bounds for the prior distribution of \( \phi \). Thus, we choose the upper and lower
bounds for the date when half of the maximum division rate is exceeded from below (i.e., the
division rate continues increasing) for each year individually. After investigating when the peak
division rate has been measured in each year, we assume that this day occurs before or on the
60–200th day after the first measurement [e.g., \( \phi \sim \mathcal{U}(0, -120) \)]. Similarly, we use the first
measurement of chl-a in each year as a guideline when choosing an upper and lower limit for
the initial value \( p(0) \) [e.g., \( p(0) \sim \mathcal{U}(0, 5) \)]. For upper and lower limits of the uniform prior
distributions for \( \phi \) and \( p(0) \) in the case of each individual year included in the data comparison
in Figure 2, see maximum and minimum values of the horizontal axis in the third (for $\phi$) and fourth [for $p(0)$] subpanels in Figure A1. In the nature of the biweekly data collection, we assume that time offset $\tau$, which is used in the computation for the relative change in the division rate in Equation (6), can be of any length between 0 and 14 days [i.e., $\tau \sim U(0, 14)$]. We choose such an upper bound for $\tau$ in order to restrict ourselves to biologically relevant values for the time offset. However, in order to illustrate how the results of the PMC ABC fitting change with prior distributions of $\tau$, we include in Figure 2 two example cases (years 1998 and 2008) where the upper and lower bounds for $\tau$ are different [i.e., $\tau \sim U(200, 365)$ and $\tau \sim U(0, 365)$].

### 2.3.2 Distance between a model simulation and the data

As a measure of discrepancy between a model simulation and the data for the chlorophyll concentration, we employ the Euclidean (i.e., the straight-line) distance between a model trajectory ($x^*$) and data ($x$):

$$d(x^*, x) = \frac{1}{N} \left( \frac{x^*}{|x^*|} - \frac{x}{|x|} \right)^2.$$  (7)

To follow the PMC ABC algorithm, we then determine a decreasing sequence of tolerance thresholds by setting the threshold of the subsequent iteration to be either (1) the distance between the data and the model simulation of the best 20% quantile of the current step or (2) equal to the tolerance threshold of the current step (if the distance of the 20% quantile is larger than the current tolerance threshold). Based on several test runs, we choose 20000 as the initial tolerance level. Finally, to get an approximation of the posterior, we iterate the PMC ABC algorithm for 10 times to collect 10000 candidate parameters (i.e., values for $\alpha$, $\tau$, $\phi$, $p(0)$, and $\sigma$) at each iteration that yield a distance between the perturbed model simulation and the data that is smaller than a given tolerance threshold. (For simulations with broad priors for $\tau$, we collect 30000 candidate parameters.)
3 Results

3.1 Parameter fitting and comparison between model simulations and chl-a data

In Figure [2], we compare the Gullmar fjord data on chl-a concentration with the chl-a concentration that we obtain as a numerical solution to the model in Equation [4] and using results of the PMC ABC parameter fitting. To assess the agreement between the data and the model, we report the distance $d$ [see Equation [7]] between the data points and the model simulation for them in each panel. In addition, we evaluate the root mean square distance (RMSE) and modeling efficiency (MEF) in order to understand how well our model predicts the chl-a concentration relative to the average of the observations (Stow et al. 2009). MEF is an indicator for how good a match there is between a simulation of a (deterministic) model and data. It is nearly identical to the coefficient of determination, which is often denoted by $R^2$, and measures the goodness of a fit of a statistical model. While $R^2$ has a lower limit of 0 (corresponding to the conclusion that the best prediction for the data is the observation average), MEF can also obtain negative values. This is because the sum of squares of the error between the model and the data may be greater than the total sum of squares. Consequently, a negative MEF value indicates that the (e.g., yearly) observation average is in a better agreement with the data than the model simulation. While an MEF value greater than 0 means that the model is a better model for the data than the observation average, MEF value close to 1 means that there is a good match between our model and the Gullmar fjord data.

Our model successfully captures the increasing (e.g., in year 2000) or decreasing (e.g., in years 2004, 2009, 2011, and 2012) trend in the chl-a concentration towards the end of the year. Based on the MEF value (between 0.34 and 0.9 with an average of 0.63, see Figure [2]), our model has better performance than the observation average for each of the 11 years included in this study. In particular in years 1998, 2004, and 2009, there is a close match (MEF above 0.7) between our model and Gullmar fjord data. In addition, for several years included in the comparison shown in Figure [2] (such as years 1995 and 2007) there are data points located both above and below the simulated chl-a concentration which further supports our model as an
indicator of the yearly trend in the chl-a concentration. Interestingly, for some years, e.g., 2005, the dynamics of the chl-a concentration are best described with a nearly flat model simulation.

For several years (e.g., 1997, 1998, 2004, 2005, 2008, 2009–2012), the highest chl-a concentration is reached during spring. This is a time of the year when according to the disturbance-recovery hypothesis the ecological coupling between primary producers and higher levels of the food chain is expected to be weak due to deep winter mixing (Behrenfeld 2014).

In year 1998, and assuming that the time offset $\tau$ can take any value between 200 and 365 days, our model successfully reproduces a high, late spring peak in chl-a concentration (see top right panel in Figure 2). However, if we limit the prior distribution for $\tau$ between 0 to 14 days, the model predicts a peak in the yearly trend during the same time (see panel on the second row and third column in Figure 2). For other years with a peak chl-a concentration above 10 mg m$^{-3}$, such as 1997, 2004, 2008, 2009, and 2012, there is a discrepancy between the model simulation and the data during the peak chl-a concentration. However, even though by eyeballing there seems to be a larger discrepancy between the model and the data in year 2012 than in year 1998, the computed distance between the model simulation and the data is in fact smaller in year 2012 than in year 1998, emphasizing the importance of quantitative rather than qualitative comparison between a model simulation and the data.

As concerns the results of the PMC ABC parameter fitting, the maximum likelihood estimates for $\tau$ that we use in our model simulations in Figure 2 (see red bars in top subpanels in Figure A1) are concentrated on small values (except for the years when the prior distribution for $\tau$ has a greater upper bound than 14, i.e., 1998 and 2008). Similarly, our parameter fitting suggests that a small value (i.e., $< 1$) of the coefficient of proportionality $\alpha$ would best fit the data (see red bars in second subpanels in Figure A1).
Figure 2: Comparison of the Gullmar fjord data and the model [i.e., numerical solution to the model equation \( \dot{p} = \alpha \Delta \mu p \) in Equation (4)]. The blue asterisks give the data for chl-a concentration (average from depths 1–10m) in the Slåggö time series of the Gullmar fjord data. The red circles denote the simulated chl-a concentration \([p(t)]\) for these same dates. We give the distance, \(d\), [see Equation (7)] root mean square error (RMSE), and modeling efficiency (MEF) [see (Stow et al. 2009)] between the asterisks (data) and circles (model) in each sub-panel. We simulate the model with parameter values [i.e., for \(\tau\), \(\alpha\), \(\phi\), and \(p(0)\)] and assuming a measurement error of \(\sigma\) obtained from using the approximate Bayesian computation (ABC) combined with population Monte Carlo (PMC) fitting method (see Section 2.3). For the fitted parameter values, we use the maximum likelihood estimate (for years 1995, 1997, 1998 (small \(\tau\)), 2000-2007, and 2009–2012) and the estimate in the posterior distribution that yields the minimum distance between the model and the data (for year 1998 (large \(\tau\)) and 2008). For precise parameter values and their location in the posterior distribution, see numerical values and red bars in the corresponding panel in Figure A1. To guide the eye, we show simulation results in red between the circles and we plot blue lines between the data points. These Gullmar fjord data were previously reported in (Tiselius et al. 2015).
3.2 Model validation with the division rate data

In addition to evaluating the goodness of a fit by computing the distance between the simulated and measured chl-a concentration for a given parameter set (see Figure 2), we assess the goodness of our model in two additional ways: By computing the distance between a model simulation (obtained with the same parameter values as in Figure 2) (1) for the phytoplankton division rate and the data for the division rate and (see Figure 3) (2) for the relative change in the division rate computed from the data and that computed from the model simulation for the division rate (see Figure 4). Thus, comparing our model simulation to the data on division rate allows us to analyze our model further as we carry out a validation similar to model forecast skill assessment [see, e.g., (Olsen et al. 2016)], because the distance between the simulated $\mu(t)$ and the measured division rate is not included in our parameter fitting in any way. We note that the model simulation for $\mu(t)$ is always between 0 and 2 as a result of how we have formulated the equation for $\mu$ using a sine-function (see Equation 5). Interestingly, the data for the division rate (computed from the data for the primary production and chl-a concentration by taking into account 1:40 chlorophyll-to-carbon ratio and that the chl-a data is an average for depths 0–10m, see Section 2.2) exceeds 2 d$^{-1}$ only in four of the 11 years included in our comparison shown in Figure 3.

When fitted to the data on chl-a concentration, although the modeling efficiency is below average (MEF negative) for all of the 11 years, our model successfully mimics the seasonal trend in the phytoplankton division rate exhibited in the data for a majority of the years shown in Figure 3. For example, the model reproduces the decreasing trend in the division rate towards the end of the year in years 1995, 2004, and 2005–2012. Furthermore, there is an agreement between the model and data as concerns a peak in the division rate occurring towards the middle/end of summer in years 1997 and 2000 (see Figure 3). For these two years, the relative change in the division rate is predicted (by the model) to change from negative to positive in spring (see years 1997 and 2000 in Figure 4). As concerns other years, the model successfully reproduces the decreasing trend in the relative change in the division rate towards the end of the year (see Figure 4). Similarly to the analysis that constitutes the basis for the disturbance-recovery hypothesis, a decreasing trend is observed in the relative change in the division rate...
computed from the satellite data for chl-a collected from the North Atlantic [see Figure 3b in Behrenfeld 2014].
Figure 3: Model validation I: Comparison of the Gullmar fjord data and the simulated division rate [i.e. $\mu(t) = A (1 + \sin[f(t + \phi)])$ in Equation (5)]. The blue asterisks give the data for the division rate that we compute from the data for the chl-a concentration and primary production (averages from depths 1–10m) in the Slåggö time series of the Gullmar fjord data with $\mu_{data} = \text{prim.prod}_{data}/(\text{chl-a}_{data} \cdot 40\text{mg} \cdot 10\text{m})$. The red circles denote the simulated division rate [$\mu(t)$] for these same dates. We give the distance, $d$, [see Equation (7)], root mean square error (RMSE), and modeling efficiency (MEF) [see (Stow et al. 2009)] between the asterisks (data) and circles (model) in each subpanel. In the model simulations, we use parameter values [i.e., for $\tau$, $\alpha$, $\phi$, and $\rho(0)$ and assuming a measurement error of $\sigma$] obtained from using the approximate Bayesian computation (ABC) combined with population Monte Carlo (PMC) method fitted to the chl-a data in Figure 2. For more details of the PMC ABC method, see Section 2.3. For the fitted parameter values and their location in the posterior distribution, see numerical values and red bars in the corresponding panel in Figure A1. To guide the eye, we show simulation results in red between the circles and we plot blue lines between the points for the division rate computed from the data. These Gullmar fjord data were previously reported in (Tiselius et al. 2015).
Figure 4: Model validation II: Comparison of the Gullmar fjord data and the simulated relative change in the division rate. The blue asterisks give the data for the relative change in the division rate that we compute from the data for the division rate in the Släggö time series of the Gullmar fjord data. That is, we first compute the division rate data from the primary production data using $\mu_{\text{data}} = \text{prim.prod}_{\text{data}}/(\text{chl-a}_{\text{data}} \cdot 40\text{mg} \cdot 10\text{m})$. We then compute the relative change in the division rate as $\Delta\mu_{\text{rel}} = (\mu_{t_2} - \mu_{t_1})/(\mu_{t_1} + \mu_{t_2}/2)$, where $t_1$ and $t_2$ are consecutive measurement dates in the data and $\mu_{t_1}$ and $\mu_{t_2}$ are the data for the division rate [or the model simulation for the division rate, $\mu(t)$] (see Figure 3) for these dates. The red circles denote the simulated relative change in the division rate for these same dates. We give the distance, $d$, [see Equation (7)] root mean square error (RMSE), and modeling efficiency (MEF) [see (Stow et al. 2009)] between the asterisks (data) and circles (model) in each subpanel. In the model simulations, we use parameter values [i.e., for $\tau$, $\alpha$, $\phi$, and $p(0)$ and assuming a measurement error of $\sigma$] obtained from using the approximate Bayesian computation (ABC) combined with population Monte Carlo (PMC) method fitted to the chl-a data in Figure 2. For more details of the PMC ABC method, see Section 2.3. For the fitted parameter values and their location in the posterior distribution, see numerical values and red bars in the corresponding panel in Figure A1. To guide the eye, we show simulation results in red between the circles and we plot blue lines between the points for the division rate computed from the data. These Gullmar fjord data are obtained by combining the prim.prod and F-prim.a data.
4 Discussion

To investigate if previous conclusions can be used to explain the dynamics of primary production, we constructed a non-autonomous ordinary differential equation model for the seasonal phytoplankton concentration. In contrast to comparing our model qualitatively to data (e.g., determining by eyeballing whether both a model simulation and the data exhibit simultaneous growth/decline), we quantified the agreement between model simulations and data by computing the straight-line distance between them. These quantitative comparisons between simulations of our model and in situ data (collected from a different location from the location of the data on which the principal model assumption is based) show that although the model is not able to capture the peak level of the yearly chl-a concentration, there is a good agreement between the model and the data as concerns the seasonal trend in the chl-a concentration.

More specifically, for each year included in this study the modeling efficiency is positive (average of 11 years is 0.63) indicating model performance from above average (0.34) to a close match (0.9) between the model and the chl-a data. Furthermore, for several individual years considered in this study, our model reproduces the timing of the peak in the division rate and the decreasing trend in the relative change in the division rate exhibited in the data. We note that neither of these quantities were considered in the fitting algorithm. That is, the distance between the model simulation and the data for these two quantities was not considered when finding the parameter values that yield a good agreement between a model simulation and the data for the chl-a concentration.

By taking a minimalistic modeling approach (in which we base our model on the results of previous data analyses instead of, e.g., constructing equations for each predator-prey interaction pair) and assuming that there is only one bloom per year, the parameter space of our model is parse. Indeed, it has one less parameter to be fitted than previous mathematical representations for the phytoplankton growth used in, e.g., (Evans & Parslow 1985) or (Huppert et al. 2005) and it does not include depth of the water column as an independent variable as, e.g., (Huisman et al. 1999). As a result, the model is fast to compute, it can be incorporated into a model with time as the only independent variable, and interannual noise can be easily added to the model for generating a representation of the phytoplankton biomass for several years in a row.
Thereby, our empirically based model that reproduces patterns seen in the Gullmar fjord data can be easily incorporated into other models, where a more realistic than a step- or a Gaussian-function representation for the dynamics at the lower levels of the food web is required for generating reliable ecological hypotheses, e.g., for changes in populations at the higher levels of the food web.

The primary model assumption that the increasing (decreasing) phytoplankton concentrations are better explained by accelerating (decelerating) division rates than by high division rates is suggested to apply to phytoplankton blooms globally (Behrenfeld 2014). Indeed, our work here shows that although this model assumption comes from studies of the North Atlantic satellite data, the model is in agreement with in situ-data collected from a fjord in the north temperate zone. In addition, we have preliminary simulations showing that the model agrees well with satellite measurements collected from the Southern Ocean. Therefore, we have a reason to believe that the model would work in not only the hypothesis’ original location, i.e., the North Atlantic, but also other oceanic regions. We note that the Gullmar fjord data provides an excellent test for the disturbance-recovery hypothesis for two main reasons: First, its location is different from the location of the data used in the formulation of the hypothesis. Second, the Gullmar fjord dataset includes both chl-a measurements collected in situ (and not via satellite) and primary production measurements collected in situ (and not computed from the chlorophyll measurements). Thus, the Gullmar fjord dataset is a strong test for the disturbance-recovery hypothesis. Although more work with data from different locations and with the other two prevailing hypotheses is needed, we conclude that the results of our work here provide some support for the disturbance-recovery hypothesis as a global mechanism for the initiation and development of phytoplankton blooms.

As concerns the other two prevailing (i.e., the critical depth (Sverdrup 1953) and turbulence (Huisman et al. 1999)) hypotheses for the initiation and development of phytoplankton blooms, we have left a comparison between these three for future work. First, this is because the scope of our work here is to test the disturbance-recovery hypothesis as a suggested mechanistic explanation for phytoplankton blooms globally. The critical turbulence hypothesis has been compared to the critical depth hypothesis and also qualitatively (but not quantitatively) to
data by Huisman et al. in (Huisman et al. 1999). Second, it is not clear if there exists yet a data set that would allow for a proper test of the Sverdrups critical depth hypothesis (Franks 2014).

Therefore, our current and future work revolves around determining whether there exist data that would allow for a fair comparison between the three different hypotheses. Such a comparison could be carried out, for example, as a Bayesian model selection [see, e.g., Toni et al. (2009)] and could help understand, e.g., whether the mechanism by which a bloom is induced and develops in a given location varies from year to year.

For simplicity, we consider seasonal variation in the phytoplankton division rate to have a period of 365 days. Consequently, a model simulation for the chl-a concentration cannot exhibit more than one peak chl-concentration per year. Autumn peaks of chlorophyll-a concentration occur in the Gullmar fjord data (Tiselius et al. 2015) and the ability of exhibiting more than one bloom per year is a generalization that we have left as future work. Indeed, bimodality is common in open oceans (Winder & Cloern 2010) and in the Arctic Ocean as ice algae bloom is followed by a pelagic bloom (Ji et al. 2013). In the case of several years considered in our study, our model does not capture the strong spring blooms in the Gullmar fjord during which the chl-a concentration is between 12 and 25 mg m$^{-3}$ in February–March. If one considers the principal assumption of our model (i.e., rate of change of chl-a concentration is proportional to the relative change in the division rate) as a representative for the forcing of the primary producers due to interaction between the grazers, winter–early spring is the time of the year when this interaction is expected to be small (Behrenfeld 2014). Thereby, one can suspect that the time of the year when there is a poor agreement between the model and the data coincides with the time of the year when other factors, such as rapid and favorable changes in temperature, light, or the mixed-layer depth, have a stronger effect on the phytoplankton dynamics than the interaction with grazers.

Furthermore, if the relative change in the division rate is considered as a representative of the interaction with grazers, our study suggests that the highest levels of the chl-concentration coincide with the time when this forcing term has reached its yearly minimum and is about to change sign from negative to positive. For example, in years 1997 and 2000 the relative change in the division rate changes sign in March–April, and as a result, the seasonal trend in
the chl-concentration is high during November. This is different from a pattern seen in year 2011, when the relative change in the division rate changes from negative to positive during December–January and the seasonal trend in the chl-a concentration is decreasing in November.

We used a previously developed method based on approximate Bayesian computation (Beau- mont et al. 2009) to find model parameters that give a small distance between a model simu-
lation and the data on chl-a concentration. We determined the prior distributions (for endpoints of these distributions, see endpoints of the x-axis in Figure A1) for \( \phi \) and \( p(0) \) by studying the data for each individual year included in our study. Our choice for the endpoints of the prior distributions for \( \sigma \) and \( \alpha \) are arbitrary. As concerns \( \sigma \), the prior distribution can reflect the error estimates for the chl-a measurements (which we do not have). As concerns \( \alpha \), our choice for the maximum possible value (\( \alpha = 5 \)) seems to be large enough (for our other choices of priors), because the posterior distributions for \( \alpha \) are concentrated on smaller values for each year (see Figure A1).

Our model’s sensitivity to parameters can be studied by looking at the posterior distributions (see Figure A1). Given that we use a uniform prior (i.e., flat distribution between the maximum and minimum values shown on the horizontal axis in each subpanel in Figure A1) for each of the parameters, we can conclude that the model is not very sensitive to the phytoplankton concentration at the beginning of the year (\( p(0) \)), measurement noise (\( \sigma \)), or the time of the year when half of the maximum division rate is exceeded (\( \phi \)). It is not surprising that the model is sensitive to the time offset parameter (\( \tau \)) (i.e., the posterior distributions for \( \tau \) have high peaks, see Figure A1) which can be considered as the main parameter characterizing the shape of the bloom (see discussion below). In addition, the model exhibits qualitatively different behavior for different values of the proportionality constant (\( \alpha \)), and our work here suggests that this constant is less than 1 for each of the years included in this study.

As is shown in the preliminary simulations in Figure 1, our model can generate qualitatively different bloom dynamics, for example, where \( \Delta \mu \) exhibits either an increasing (left) or a decreasing (right) yearly trend. This is because of differences in \( \tau \) that defines the time window during which the relative change of the division rate is computed. Consequently, \( \tau \) can also be interpreted as a representative of the events during \( \tau \) days in the past that have an effect on
the current division rate. Apart from two example years 1998 and 2008, we assumed in the
comparisons between model and data that such a “memory” in the division rate would not be
longer than 14 days. However, we note that this choice of a prior distribution for $\tau$ has an effect
on the simulation results. For example, a model simulation with a large $\tau$ and $\phi$ generates an
excellent agreement between the model and the data as concerns the high level of chl-a during
late spring (see Figure 2, top right panel). However, the division rate predicted by the model
reaches its maximum in early winter (see Figure 3, top right panel). While this example simu-
lation demonstrates the wide range of possible bloom dynamics generated by our minimalistic
model, it also suggests that more work is required for determining a biologically meaningful
upper and lower bounds for the prior distribution of $\tau$. Here, as concerns the lower bound, we
choose it to be 0 because we have not found a good biological justification for negative $\tau$ (i.e.,
for taking into account the effect that future conditions have on the present chl-a concentration).
As concerns the upper bound, based on our simulations with the model, a large $\tau$-value often
results in a “pointy” bloom or suggests a peak chl-a concentration in the middle of the winter
(e.g., see Figure 2, panels c) and i), respectively). Therefore, we suggest choosing a high upper
bound for $\tau$ if the dataset in question exhibits a peak chl-a concentration and if one can think
of physical or biological factors contributing to “long” memory in the ecosystem in question.

5 Conclusions

In this work, we combine deterministic differential equation modeling (which in general can
be used to test suggested hypotheses for the inner workings of biological processes) with in
situ data on chl-a concentration and primary production to contribute towards an understand-
ing of a general representation for the dynamics of primary production. Here, the hypothesis
that we test emphasizes the influence of grazing on phytoplankton population dynamics. Our
model reproduces the yearly trend in the chl-a concentration exhibited in the Släggö time series
for several different years collected from the Gullmar fjord. In addition, the model is capable
of generating bloom dynamics with a high level of chl-a concentration during one month and
followed by low levels of chl-a concentration during the rest of the year (as exhibited in the
Furthermore, comparisons between model simulations and in situ data for primary production show that in the case of several years considered here, the model not only reproduces the seasonal pattern and timing of the peak division rate but also captures the decreasing trend in the relative change in the division rate. As a result, our model, which assumes that the relative change in division rate is proportional to the rate of change in the phytoplankton concentration [as is suggested by the disturbance-recovery hypothesis (Behrenfeld et al. [2013])], suggests a possible empirically based explanation for the bloom dynamics. As a minimalistic mathematical representation for the dynamics of the seasonally chl-a concentration, our model can be easily fitted to other datasets, and thereby, considered as a simple (yet more realistic than a step-, sine-, or Gaussian function) representation of the dynamics at the lower levels of the food web. We argue that our model is important for the scientific community focusing on the dynamics of phytoplankton blooms because it can be used i) to test whether the disturbance-recovery hypothesis can explain the initiation and development of phytoplankton blooms globally and ii) in future modeling studies in which a simple and accurate representation of the seasonal unimodal primary production patterns is needed, for example, for simulating and predicting the dynamics and phenology of the predators of the primary producers.

Acknowledgements

We thank Peter Tiselius for several useful discussions and sending us the Släggö time series in the Gullmar Fjord data, which were obtained with support from the Swedish Environmental Protection Agency, the County Administrative Board of Västra Götaland and Bohuskustens Vattenvårdsförbund. We are grateful to Frank Schilder for numerous discussions on the implementation of the parameter fitting algorithm and we thank Chris Klausmeier and Philip Maybank for helpful discussions. S. H. P. was supported by People Programme (Marie Curie Actions) of the European Union’s Seventh Framework Programme (FP7/2007-2013) under REA grant agreement #609405 (COFUNDPostdocDTU) and by the Research Council of Norway through Yggdrasil 2014 mobility programme (IS-MOBIL, project number 227487). Ø. V. was supported by the Research Council of Norway through project 227046. We are grateful to the
three anonymous referees for their helpful comments.

References


zooplankton system. Nonlinear Anal Real World 10:1643–1650

different mechanisms for the development of phytoplankton blooms. Limnol Oceanogr
44(7):1781–1787

J Theor Biol 236:276–290

–171

Ji R, Jin M, Varpe Ø (2013) Sea ice phenology and timing of primary production pulses in the

Klausmeier CA (2010) Successional state dynamics: A novel approach to modeling nonequi-
librium foodweb dynamics. J Theor Biol 262:584–595

models of the pelagic microbial food web. Limnol and Oceanogr 58(1):301–313

Lindahl O, Andersson L, Belgrano A (2009) Primary phytoplankton productivity in the Gull-


National Aeronautics and Space Administration Ocean Biology (2019) Sea-viewing Wide
Field-of-view Sensor (SeaWiFS) Ocean Color Data. https://oceancolor.gsfc.nasa.gov/
(accessed Jun 4 2019)

Obata A, Ishizaka J, Endoh M (1996) Global verification of critical depth theory for phyto-
plankton bloom with climatological in situ temperature and satellite ocean color data. J
Geophys Res 101(C9):20657–20667


Appendix A: Posterior and prior distributions

For the posterior (and the minimum and maximum values of the uniform prior) distributions of the parameters $\tau$, $\alpha$, $\phi$, the simulated chl-a concentration at the beginning of the year, $p(0)$, and the simulated measurement error, $\sigma$, we use in our implementation of the PMC ABC parameter fitting algorithm (Beaumont et al. 2009), see A1. For parameter values used in the model simulations in Figures 2–4, see red bars and their numerical value given in each subpanel in Figure A1.
Figure A1: Posterior distributions of the model parameters $\tau$, $\alpha$, and $\phi$, the simulated chl-a concentration at the beginning of the year, $p(0)$, and the simulated measurement error, $\sigma$, accepted at the strictest tolerance level in the approximate Bayesian computation (ABC) combined with population Monte Carlo (PMC) method (Beaumont et al. 2009) for each year in the Gullmar fjord data we use to infer parameters of our model [see Equation 4 in Figure 2]. For each year, we give the numerical value of the parameter within each subpanel, and the red bar denotes the location of this value in the posterior distribution. We use this parameter value in model simulations in the corresponding panel in Figures 2–4. We note that the value for $\tau$ in year 1998 (large $\tau$) and $p(0)$ in year 2000 are included in the posterior. In these cases, the posterior has a long tail that is not visible when represented with a histogram of 100 bins as in this figure. For the prior distributions in the PMC ABC method, we use uniform distributions between the maximum and minimum values shown on the horizontal axis in each subpanel.