Empirically based minimalistic model for representing seasonal phytoplankton dynamics

Sofia H. Piltz[1,2,3,*], Poul G. Hjorth[2], Øystein Varpe[4,5,6]

[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**

[6]*Present address:* **Department of Biological Sciences, University of Bergen, 5020 Bergen, Norway**

ABSTRACT: Supported by chl *a* satellite data in the North Atlantic (and phytoplankton division rates computed from that data), 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. We used this hypothesis, introduced by Behrenfeld, as a principal model assumption and constructed a non-autonomous ordinary differential equation model for seasonally varying chl *a* concentrations. Our quantitative comparison between model simulations and *in situ* measurements of chl *a* and primary production collected from a Swedish fjord was 2-fold: first, using approximate Bayesian computations, we found distributions of values for the 3 model parameters that best described the chl *a* data. Then, we validated our model by comparing the simulated (not fitted) division rate to the division rate determined from the data. Our minimalistic model was 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 was positive (between 0.3 and 0.9 with an average of 0.63) for all 11 yr included in this study. 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.

KEY WORDS: 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 respira-

tion 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), which suggests that 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 representation for seasonally varying chlorophyll concentrations.

Previous studies (Evans & Parslow 1985, Behrenfeld & Boss 2014) suggest that under equilibrium conditions (i.e. when the mixed-layer-integrated phytoplankton production equals that of losses), the rate of change of biomass depends on the current division rate. However, as has been demonstrated by *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, which is rapidly caught by increased encounter rates with the grazers of phytoplankton (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 strong top-down regulation, the continuous rise in phytoplankton division rate due to gradually improving spring environmental conditions is a key factor in naturally occurring 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 rather would apply to phytoplankton blooms globally (Behrenfeld 2014).

In addition to conceptual modeling frameworks, contrasting conclusions exist within the mathematical models for bloom dynamics as concerns the impor-

tance of bottom-up and top-down (i.e. predation by herbivorous plankton) regulation mechanisms. Namely, mathematical modeling work suggests that both bottom-up (Evans & Parslow 1985, Huppert et al. 2002, 2005, Platt et al. 2009) and top-down (Evans & Parslow 1985, Truscott & Brindley 1994, Scheffer et al. 1997, van Nes & Scheffer 2004, Freund et al. 2006, Dakos et al. 2009, Gao et al. 2009, Klausmeier 2010) 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 qualitatively reproduce (e.g. when compared 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 quickly becomes very large, which makes parameter fitting challenging, if not impossible. In addition, a survey of complex biogeochemical models shows that there is a great deal 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 next steps towards better models for phytoplankton blooms (Behrenfeld & Boss 2014, Fischer et al. 2014, Shimoda & Arhonditsis 2016).

Here, we contribute to the development of models for phytoplankton blooms with predictive power by trading off detailed model construction from physiological principles for a sparse 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 comparison to data has been previously done by 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 phytoplankton concentration incorporates the disturbance–recovery hypothesis as the principal model assumption for the dependency between the division rate and chlorophyll *a* (chl *a*) concentration. Thus, and as suggested by the North Atlantic satellite data (Behrenfeld et al. 2013, Behrenfeld 2014, Behrenfeld

& Boss 2014), we assume 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 rather to the relative change in the cell division rate. To test the hypothesis as a suggested general mechanism for phytoplankton bloom initiation, we incorporated 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. (2016) from the Swedish west coast.

2. METHODS

In the following, we use the terms 'phytoplankton population' (*p*) and 'chlorophyll concentration' (chl) interchangeably, because the data we had access to contains measurements of the latter, while the mathematical model we constructed is not limited and can also be used to describe the dynamics of the former. We also had access to measurements of primary production from which we computed the phytoplankton division rate.

2.1. The model

2.1.1. Empirically based, non-autonomous differential equation model for *p*

Investigations of satellite data collected from the subarctic Atlantic Ocean (SeaWiFS; https://oceancolor. gsfc.nasa.gov/) suggest that the rate of change in the chlorophyll concentration, $\Delta \text{chl}/\Delta t$, changes in proportion to the relative change, $\Delta \mu_{rel}$, rather than the absolute change in division rate (see Fig. 3 in Behrenfeld 2014). In addition, during the subarctic Atlantic bloom, the rate of biomass accumulation is uncorrelated with the division rate (see Fig. 3 in Behrenfeld & Boss 2014). Considering change in chlorophyll as a change in the natural logarithm of measured chlorophyll (see caption for Fig. 3 in Behrenfeld 2014), and assuming that the relative change in the division rate can be computed as the average of 2 consecutive measurements (see Eq. 1 in Behrenfeld 2014), the conclusion of the data analysis in Behrenfeld (2014) can be written as:

$$\frac{\Delta \ln \text{chl}}{\Delta t} \propto \Delta \mu_{rel} = \frac{2[\mu(t) - \mu(t-8)]}{\mu(t) + \mu(t-8)} \quad (1)$$

where $\Delta t = 8$ d is the resolution of the satellite data and $\mu(t)$ is the division rate at time t. In these studies, the chlorophyll data was obtained using satellite-based data sensor to detect surface chlorophyll (SeaWiFS; https://oceancolor.gsfc.nasa.gov/). 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 used the disturbance–recovery hypothesis as our model assumption and constructed 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, $\text{d} \ln p/\text{d}t$, is proportional to the relative change in phytoplankton division rate, $\Delta \mu(t)$, reported in Behrenfeld (2014) and shown in Eq. (1), is written as follows:

$$\frac{\text{d} \ln p}{\text{d}t} \propto \Delta \mu(t) \quad (2)$$

Assuming a proportionality constant α, this yields the following equation:

$$\frac{\text{d} \ln p}{\text{d}t} = \alpha \Delta \mu(t) \quad (3)$$

Using logarithmic differentiation; that is, $\text{d} \ln p/\text{d}t = (1/p)(\text{d}p/\text{d}t)$, and rearranging, we obtain our non-autonomous differential equation model for p:

$$\frac{\text{d}p}{\text{d}t} = \dot{p} = \alpha \Delta \mu(t) p \quad (4)$$

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

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

$$\mu(t) = A\{1 + \sin[f(t + \phi)]\} \quad (5)$$

where A is the difference between maximum and minimum division rate (i.e. the amplitude of the seasonal forcing) and f is the period of the oscillation. We assumed seasonal changes in the division rate, and therefore, used $f = 2\pi / (365 \text{ d})$, i.e. the period of the oscillation in the division rate is exactly 1 yr. Parameter ϕ in Eq. (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 ϕ'th day of the year.

We followed Behrenfeld (2014) and assumed that $\Delta\mu(t)$ 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 d, which is the time difference between consecutive satellite measurements. However, here we considered a continuous-time system based on the results reported in Behrenfeld (2014), and therefore, we denoted the time difference between 2 consecutive points with a time offset, τ. As a result, $\Delta\mu(t)$ is computed as:

$$\Delta\mu(t,\tau,\phi) = \frac{\mu(t)-\mu(t-\tau)}{\frac{\mu(t)+\mu(t-\tau)}{2}} = \frac{2\{\sin[f(t+\phi)]-\sin[f(t-\tau+\phi)]\}}{2+\sin[f(t+\phi)]+\sin[f(t-\tau+\phi)]} \quad (6)$$

Note that $\Delta\mu(t)$ given in Eq. (6) does not depend on the difference between the maximum and minimum division rate within a year (i.e. A). In addition, when $\tau = 0$, $\Delta\mu(t) = 0/\{2 + 2\sin[f(t + \phi)]\} = 0$ and our minimalistic model (Eq. 4) predicts an equilibrium chl 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 ϕ and τ, $\Delta\mu(t)$ in Eq. (6) has a period of 365 d (i.e. $\Delta\mu$ repeats itself every 365 d). Therefore, the model prediction for chl which is a solution to Eq. (4) is also 365-periodic. Model variables and parameters are shown in Table 1.

We did not find a closed solution $p(t)$ to the differential equation in Eq. (4). In the following section, Eq. (4) was simulated numerically. More specifically, we compared a model simulation of chl at a given day of the year to *in situ* measurements of chl *a* concentration at the same day and quantified the difference (by computing the straight-line distance) between the 2. We then used *in situ* data for primary production (collected at the same time as the chlorophyll data) to first compute the division rate, and then to investigate how well our model, which was fitted to the chl *a* data (but not to the division rate data), reproduced the patterns exhibited in the division rate data.

2.1.3. Preliminary model simulations

As an initial test of our minimalistic model, we solved Eq. (4) numerically to simulate the model with an arbitrarily chosen parameter set and thereby compare model predictions qualitatively to the results of the data analysis reported in Behrenfeld (2014). Indeed, an example simulation of our phytoplankton model reproduced characteristics of phytoplankton blooms, such as: (1) the beginning of the bloom and an increase in phytoplankton population do not necessarily coincide; (2) the phytoplankton population can increase despite the fact that the division rate is decreasing; and (3) the phytoplankton concentration remains low even though $\Delta\mu(t)$ is strongly negative; all observed in the satellite data from the subarctic Atlantic and described in (Behrenfeld 2014) (Fig. 1). To investigate whether these results agreed quantitatively with a different data set, and to infer model parameters, we used the algorithm developed by Beaumont et al. (2009). In doing so, we quantified the difference between simulations of our model and data from a Swedish fjord.

Table 1. Model variables and parameters with their values, units, and biological explanation

| Name	Description/value	Units	Explanation
t	Independent variable	d	Time
$p(t)$	Compared to the data	mg m^{-3}	Chlorophyll concentration at time t
$p(0)$	Fitted to the data	mg m^{-3}	Chlorophyll concentration at time $t = 0$
$\mu(t)$	Validated with the data	mg C m^{-2}d^{-1}	Phytoplankton division rate at time t
$\Delta\mu(t)$	Validated with the data	Dimensionless	Relative change in the phytoplankton division rate
ϕ	Fitted to the data	d	Number of days after the first day of the year when half of the maximum division rate is exceeded from below
α	Fitted to the data	d^{-1}	Coefficient of proportionality, i.e. denoting the amount to which chlorophyll concentration is proportional to the accelerating/decelerating division rates
τ	Fitted to the data	d	Amount of time before t that is used in the computation of the relative change in division rate (see Eq. 6), i.e. determines the width of the time offset to which the rate of change of chlorophyll concentration is proportional
f	$2\pi/365$	d^{-1}	Period of the oscillation in phytoplankton division rate
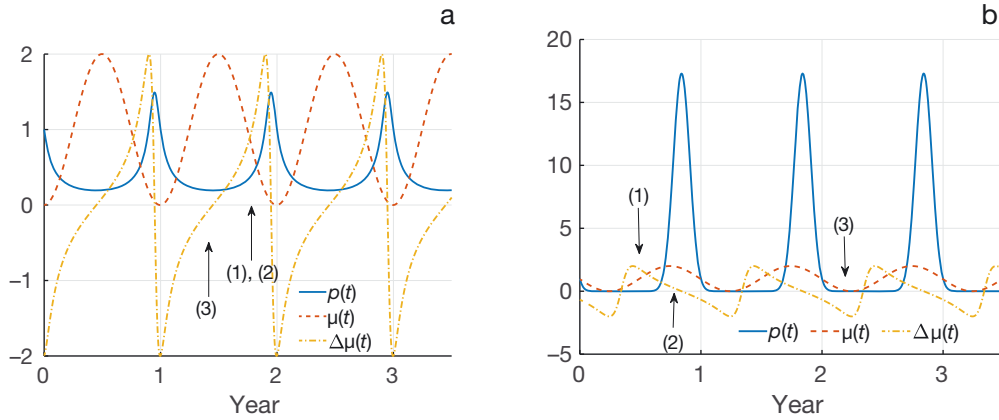
Fig. 1. Preliminary model simulations. Solid lines: numerical solution to the model equation $\dot{p} = \alpha\Delta\mu p$ (Eq. 4) for the chl *a* concentration p (in mg m^{-3}); dashed lines: phytoplankton division rate $\mu(t)$ (in mg C m^{-2} d^{-1}) (Eq. 5); dash-dotted lines: relative change in the phytoplankton division rate $\Delta\mu(t)$ (dimensionless) (Eq. 6), using an example parameter set (a) $(\tau,\alpha,\phi,p[0]) = (330, 0.015,-90,1)$ and (b) $(\tau,\alpha,\phi,p[0]) = (71, 0.15,-180,1)$. See section 2.1.3 for model behavior at arrow locations 1, 2 and 3

2.2. Gullmar fjord data

To test our model, the principal assumption of which is based on the satellite observations from the subarctic Atlantic (Behrenfeld 2014), we compared model simulations to *in situ* measurements collected from the Gullmar fjord (in the Släggö time series) and previously reported in (Tiselius et al. 2016). 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 km^2 and a maximum depth of 120 m. 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 had access to time series, which consisted of biweekly *in situ* measurements of primary production collected at the mouth of the fjord (58° N, 11° E) since 1985. Gullmar fjord is a marine reserve and the mouth is considered a good representative of the open-sea ecosystems in the area. It has no major local sources of sewage, but receives large-scale pollution from the seas in northern Europe and Sweden (Lindahl et al. 2009). Similar to the data set from the subarctic Atlantic (Behrenfeld & Boss 2014), the highest primary production rate in the Gullmar fjord shows no relation to the phytoplankton biomass (Tiselius et al. 2016). In addition, the coastal ecosystem in the fjord (which is categorized as having an intermediate level of productivity) was shown to display top-down regulation by zooplankton on a seasonal scale (Tiselius et al. 2016).

In our comparison between model simulations in Section 3, we used the data for primary production (mg C m^{-2} d^{-1}) and chl *a* (mg m^{-3}) reported previously in Tiselius et al. (2016). Primary production was measured using the ^{14}C-technique for each sample that was collected at depths 0, 1, 2, 3, 4, 6, 8, and 10 m. These samples were then incubated in glass bottles for 4 h, and carbon uptake was transformed into daily production by the light factor method (Tiselius et al. 2016). The data are the average of these 8 bottles. We also included the amount of chl *a* extracted from a 100 ml filtered water sample and the average of these measurements from depths of 1–10 m. To compute the division rate from the primary production data, we divided each primary production measurement by the same-day measurement for chl *a*, which we first multiplied by 40 mg (using the 1:40 chlorophyll to carbon ratio) and 10 m (because the data represent the average for depths 0–10 m). We limited our study to 11 yr between 1993 and 2012, to years where the Gullmar fjord data set had no missing data between consecutive measurements and both chl *a* and primary production were measured on the same day (i.e. 1995, 1997, 1998, 2000, 2004, 2005, 2007, 2008, 2009, 2011, 2012).

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 investigated the problem of fitting the

parameters of Eq. (4) to the Gullmar fjord data (see Section 2.2) through Bayesian inference. More specifically, we made an implementation in MATLAB of an existing algorithm for approximate Bayesian computation (ABC) combined with a population Monte Carlo (PMC) method introduced by Beaumont et al. (2009). In contrast to, e.g. least-squares fitting, Bayesian inference allows one to study the results from the posterior parameter distribution rather than just a single value that gives the best fit as a result of an optimization method. Bayesian inference is well-suited for complex problems in ecology (Clark 2005, Beaumont 2010), and it has been previously used in aquatic ecology to estimate parameters and the accuracy of models describing open sea microbial food webs (Lignell et al. 2013) or to predict the impact of biogeochemical cycles and excess nutrients on water quality (Zhang & Arhonditsis 2008). For an introduction to Bayesian approaches and their use in evolution and ecology, seeBeaumont (2010). For a detailed description of the PMC ABC method that we used to fit parameters to data, see Beaumont et al. (2009, p. 987).

Let us consider a model simulation for the chlorophyll concentration at a given time t, that is, the solution to Eq. (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 $\mathbf{x}$. As described in Section 2.2, the data were measured approximately biweekly, that is, at time instances t_i, where i denotes the day of the measurement. Thus, excluding any measurement errors, the underlying assumption in our parameter-fitting procedure was that the data would be represented by a set of model simulations for time instances t_i, i.e. by $\mathbf{x}_i = x(t_i;\alpha,\tau,\phi)$, for some unknown, true parameter values. To account for the presence of measurement errors, we incorporated normally distributed noise in our model simulation. We denoted the standard deviation of the measurement noise with σ (and because we did not know the variance of measurement errors in advance, we incorporated the estimation of σ in the parameter-fitting procedure). Thus, for given parameter values $(\alpha,\tau,\phi,\sigma)$, the model simulation (i.e. $\mathbf{x}^{\star}$) 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 Eq. (4) and variance σ^2; i.e. $\mathbf{x}_i^{\star} \sim \mathcal{N}[x(t_i;\alpha,\tau,\phi),\sigma^2)]$.

2.3.1. Prior distributions for σ, α, ϕ, $p(0)$, and τ

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

In the case of ϕ, we studied the division rate in the Gullmar fjord data set in order to determine the lower and upper bounds for the prior distribution of ϕ. Thus, we chose 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 was measured in each year, we assumed that this day occured before or on the 60–200th day after the first measurement; e.g. $\phi \sim \mathcal{U}(0,-120)$. Similarly, we used 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 ϕ and $p(0)$ in the case of each individual year included in the data comparison (see Fig. 2), see maximum and minimum values of the horizontal axis in the third (for ϕ) and fourth (for $p[0]$) subpanels of each panel in Fig. A1 in the Appendix. We assumed that τ, which is used in the computation for $\Delta\mu(t)$ in Eq. (6), can be of any length between 0 and 14 d; i.e. $\tau \sim \mathcal{U}(0,14)$. We chose such an upper bound for τ in order to restrict ourselves to biologically relevant values. However, in order to illustrate how the results of the PMC ABC fitting changed with prior distributions of τ, we included 2 example cases (see Fig. 2, years 1998 and 2008) where the upper and lower bounds for τ were different, i.e. $\tau \sim \mathcal{U}(200,365)$ and $\tau \sim \mathcal{U}(0,365)$.

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 employed the Euclidian (i.e. straight-line) distance (d) between a model trajectory ($\mathbf{x}^{\star}$) and data ($\mathbf{x}$):

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

To follow the PMC ABC algorithm, we then determined 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 was larger than the current tolerance threshold). Based on several test runs, we chose 20 000 as the initial tolerance level. Finally, to get an approximation of the posterior, we iterated the PMC ABC algorithm 10 times to collect 10 000 candidate parameters (i.e. values for α, τ, ϕ, $p[0]$, and σ) at each iteration that yielded a distance between the perturbed model simulation and the data that was smaller than a given tolerance threshold. For simulations with broad priors for τ, we collected 30 000 candidate parameters.

3. RESULTS

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

In Fig. 2, we compare the Gullmar fjord data on chl *a* concentration with the chl *a* concentration that we obtained as a numerical solution to the model in Eq. (4) using results of the PMC ABC parameter-fitting. To assess the agreement between the data and the model, we report the distance (d in Eq. 7) between the data points and the model simulation in each panel. In addition, we evaluated the root mean square error (RMSE) and modeling efficiency (MEF) in order to understand how well our model predicted 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, R^2, which 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 better agreement with the data than the model simulation. While an MEF value > 0 means that the model is a better model for the data than the observation average, an MEF value close to 1 means that there is a close match between our model and the Gullmar fjord data.

Our model successfully captured the increasing (e.g. in 2000) or decreasing (e.g. in 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 Fig. 2), our model has better performance than the observation average for each of the 11 yr included in this study. In particular in 1998, 2004, and 2009, there is a close match (MEF above 0.7) between our model and Gullmar fjord data. In addition, for several of the years in Fig. 2 (such as 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 chl *a* concentration. Interestingly, for some years, e.g. 2005, the dynamics of the chl *a* concentration are best described with a nearly flat curve.

For most years (i.e. 1997, 1998, 2004, 2005, 2008, 2009, 2011, 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 1998, and assuming that τ can take any value between 200 and 365 d, our model successfully reproduces a high, late spring peak in chl *a* concentration (Fig. 2c). However, if we limit the prior distribution for τ to between 0 and 14 d, the model predicts a peak in the yearly trend during the same time (Fig. 2f). For other years with a peak chl *a* concentration above 10 mg m^{-3}, including 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 2012 than in 1998, the computed distance between the model simulation and the data is in fact smaller in 2012 than in 1998, emphasizing the importance of quantitative rather than qualitative comparison between a model simulation and the data.

Regarding the results of the PMC ABC parameter-fitting, the maximum likelihood estimates for τ that we used in our model simulations in Fig. 2 (see red bars in top subpanels in each panel in Fig. A1) are concentrated on small values (except for the years when the prior distribution for τ 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 α would best fit the data (see red bars in second subpanels in each panel in Fig. A1).
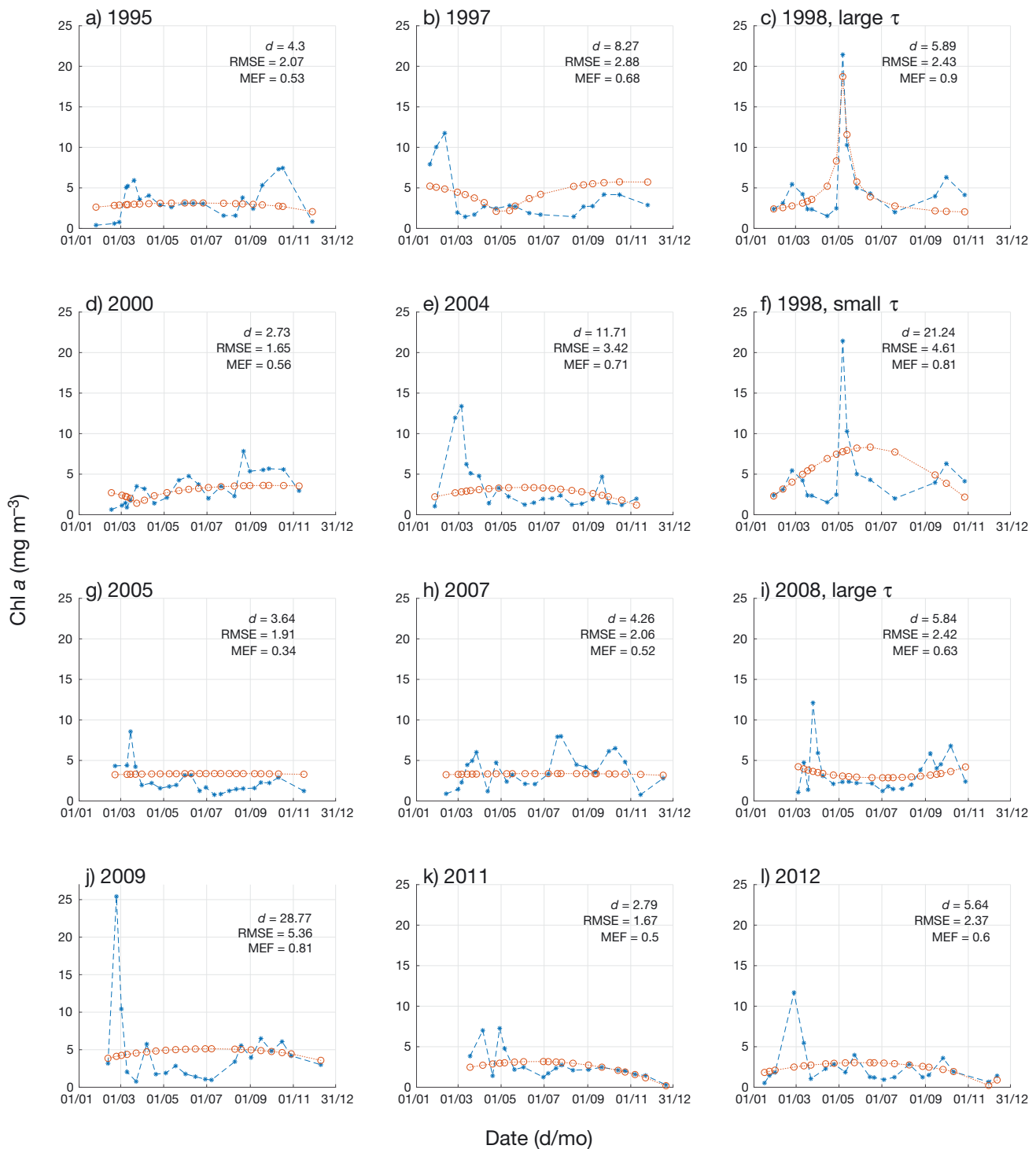
Fig. 2. Comparison between Gullmar fjord data (Tiselius et al. 2016) and the model (i.e. numerical solution to the model in Eq. 4). Blue asterisks: chl *a* concentration (average from depths 1–10 m) from time series data; red circles: simulated chl *a* concentration, $p(t)$, for these same dates. Values given for distance, d (see Eq. 7), RMSE, and modeling efficiency (MEF) between the data and model for each year. See Section 2.3 for methods. For the fitted parameter values, we used the maximum likelihood estimate (for 1995, 1997, 1998 [small τ], 2000, 2004, 2005, 2007, 2009, 2011, 2012) and the estimate in the posterior distribution that yields the minimum distance between the model and the data (for 1998 [large τ] and 2008). For precise parameter values and their location in the posterior distribution, see numerical values and red bars in the corresponding panel in Fig. A1. Data points are joined with dots/dashes for clarity

3.2. Model validation with the division rate data

In addition to evaluating the goodness-of-fit by computing the distance between the simulated and measured chl *a* concentration for a given parameter set (see Fig. 2), we assessed the performance of our model in 2 additional ways: by computing the distance between a model simulation (obtained with the same parameter values as in Fig. 2) (1) for the phytoplankton division rate and the data for the division rate (Fig. 3) and (2) for $\Delta\mu(t)$ computed from the data and that computed from the model simulation for the division rate (Fig. 4). We compute $\mu(t)$ from the data as: $\mu_{data} = $ prim.prod$_{data}$ /(chl a_{data} · 40 mg · 10 m). Then, the relative change in the division rate data is: $\Delta\mu_{rel}^{\star} = (\mu_t - \mu_{t_1})/[(\mu_{t_1} + \mu_{t_2})/2]$, where t_1 and t_2 are consecutive measurement dates in the data and μ_{t_1} and μ_{t_2} are the data for the division rate (for the model simulation for the division rate, $\mu(t)$) for these dates. 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 Olsen et al. 2016), because the distance between the simulated and measured $\mu(t)$ 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 $\mu(t)$ using a sine function (see Eq. 5). Interestingly, the data for $\mu(t)$ (computed from the data for the primary production and chl *a* concentration by taking into account the 1:40 chlorophyll-to-carbon ratio and that the chl *a* data is an average for depths 0–10 m; see Section 2.2) exceeds 2 d^{-1} only in 4 of the 11 years (Fig. 3).

When fitted to the data on chl *a* concentration, although the MEF is below average (negative) for all 11 yr, our model successfully mimics the seasonal trend in the phytoplankton division rate exhibited in the data for a majority of the years shown in Fig. 3. For example, the model reproduces the decreasing trend in the division rate towards the end of the year in 1995, 2004, 2005, 2007, 2008, 2009, 2011, and 2012. Furthermore, there is agreement between the model and data as concerns a peak in the division rate occurring towards the middle/end of summer in 1997 and 2000 (see Fig. 3). For these 2 years, $\Delta\mu(t)$ is predicted (by the model) to change from negative to positive in spring (i.e. 1997 and 2000 in Fig. 4). In other years, the model successfully reproduces the decreasing trend in $\Delta\mu(t)$ towards the end of the year (Fig. 4) as has also been observed in the $\Delta\mu(t)$ computed from the North Atlantic chl *a* data (see Fig. 3b in Behrenfeld 2014).

4. DISCUSSION

To investigate if the disturbance–recovery hypothesis can be used to explain the dynamics of primary production, we constructed a non-autonomous ordinary differential equation model for 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 good agreement between the model and the data in the seasonal trend in chl *a* concentration. More specifically, for each year included in this study, the MEF is positive (average of 11 yr 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 of the years included in this study, our model reproduces the timing of the peak in the division rate and the decreasing trend in $\Delta\mu(t)$ exhibited in the data. 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 2 quantities was not considered when finding the parameter values that yielded good agreement between a model simulation and the data for the chl *a* concentration.

By taking a minimalistic modeling approach (in which we based 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 1 bloom yr^{-1}, the parameter space of our model is sparse. Indeed, it has one less parameter to be fitted than previous mathematical representations for the phytoplankton growth used in Evans & Parslow (1985) or Huppert et al. (2005), and it does not include depth of the water column as an independent variable as in 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 for generating a representation of the phytoplankton biomass for several years in a row. Thus, 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-
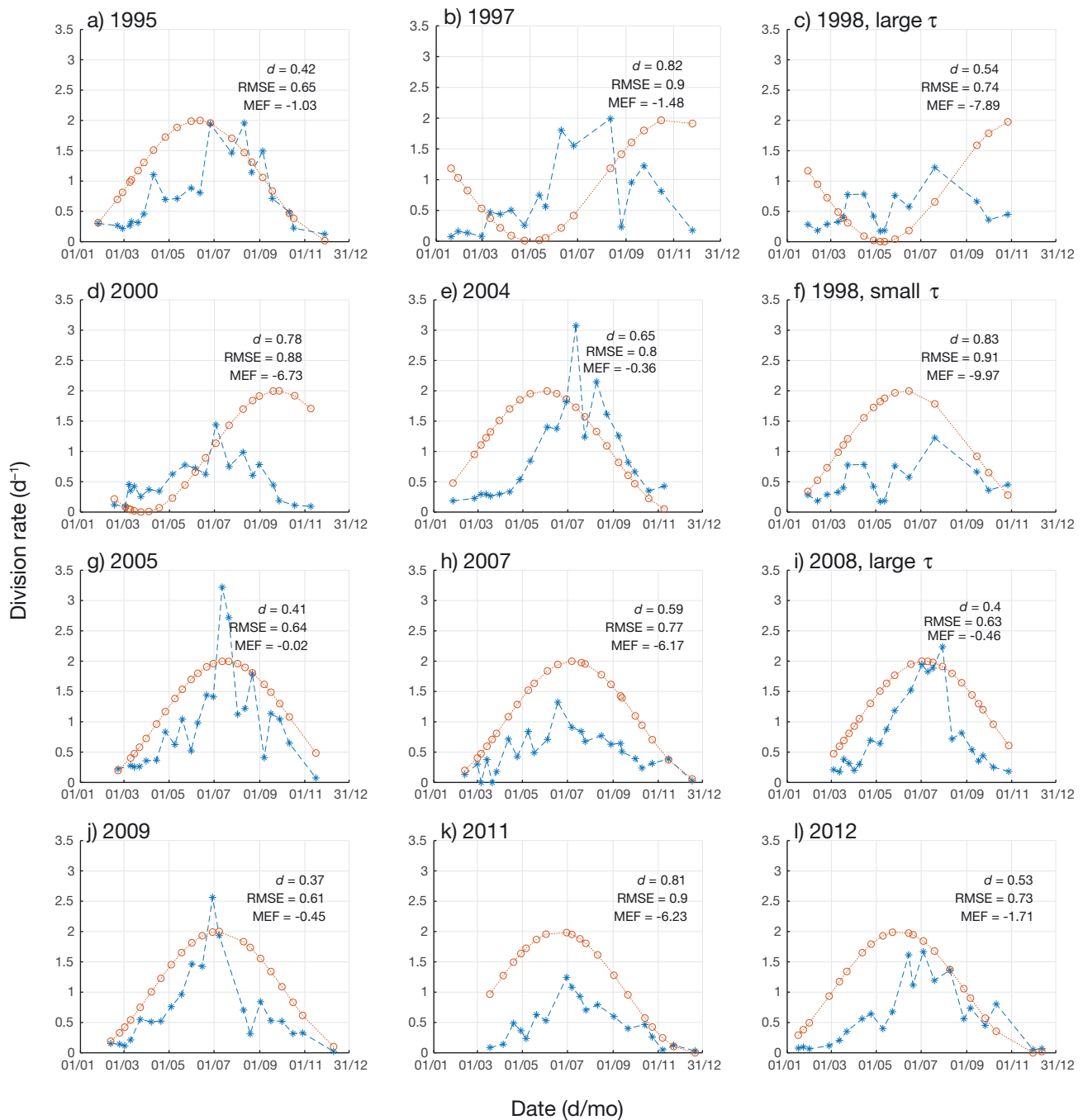
Fig. 3. Model validation I: comparison between Gullmar fjord data (Tiselius et al. 2016) and the simulated division rate in Eq. (5). Blue asterisks: division rate computed from the data for the chl *a* concentration and primary production (averages from depths 1–10 m) in the time series data (see Section 3.2). Red circles: simulated division rate for these same dates. See Fig. 2 for further details

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 based on Behrenfeld (2014). Indeed, our study shows that although this model assumption comes from studies of North Atlantic satellite data, the model is in an agreement with *in situ* data collected from a fjord in the north
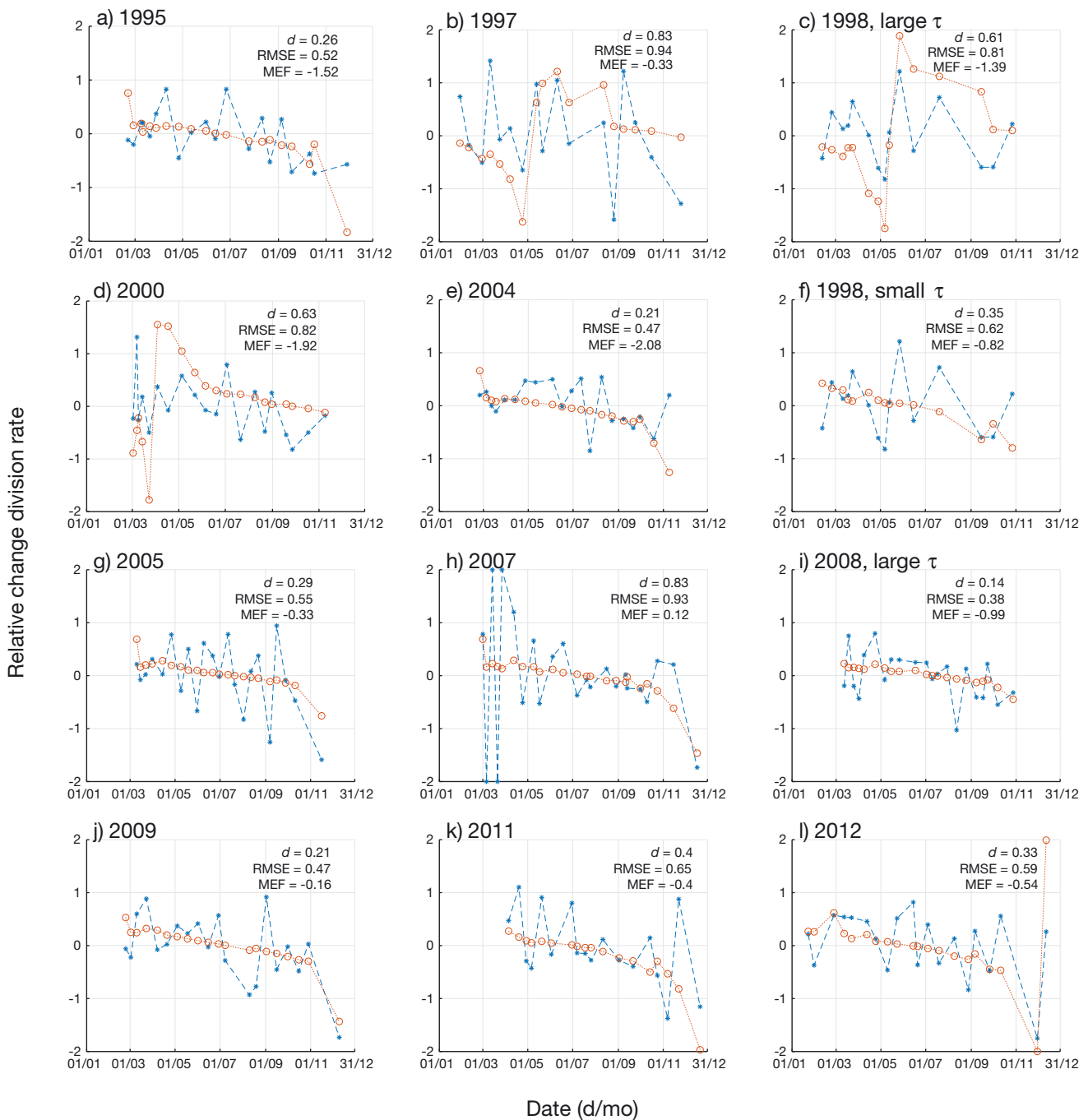
Fig. 4. Model validation II: comparison between Gullmar fjord data (Tiselius et al. 2016) and the simulated relative change in division rate in the time series data (see Section 3.2). Blue asterisks: relative change in division rate computed from the data for the division rate in the time series data. Red circles: simulated relative change in the division rate for the same dates. See Fig. 2 for further details

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 reason to believe that the model would work in not only the hypothesis' original location, i.e. the North Atlantic, but also in other

oceanic regions. We note that the Gullmar fjord data provides an excellent test for the disturbance–recovery hypothesis for 2 main reasons: first, its location is different from the location of the data used in the formulation of the hypothesis. Second, the Gullmar fjord data set 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 data set is a strong test for the disturbance–recovery hypothesis. Although more work with data from different locations and with the other 2 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.

Regarding the other 2 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 them and the disturbance–recovery hypothesis for future work. First, this is because the scope of our work here was to test the disturbance–recovery hypothesis as a suggested mechanistic explanation for phytoplankton blooms. The critical turbulence hypothesis has been compared to the critical depth hypothesis and also qualitatively (but not quantitatively) to data by Huisman et al. (1999). Second, there are questions about the suitability of existing data sets for a proper test of Sverdrup's critical depth hypothesis (Franks 2015). A comparison between the 3 different hypotheses could help elucidate whether the mechanism by which a bloom is induced and develops in a given location varies from year to year.

For simplicity, we considered seasonal variation in the phytoplankton division rate to have a period of 365 d. Consequently, a model simulation for the chl *a* concentration cannot exhibit more than one peak chl *a* concentration per year. Autumn peaks occur in the Gullmar fjord data (Tiselius et al. 2016) and the ability to exhibit more than 1 bloom per year is a generalization that we have left for future studies. 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 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 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 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 $\Delta\mu(t)$ is considered as a representative of the interaction with grazers, our study suggests that the highest levels of the chl *a* 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 1997 and 2000, $\Delta\mu(t)$ changes sign from negative to positive in March–April, and as a result, the chl *a* concentration is high during November. This is different from the pattern observed in 2011, when $\Delta\mu(t)$ changed from negative to positive during December–January and the chl *a* concentration decreased in November.

We used a previously developed method based on approximate Bayesian computation (Beaumont et al. 2009) to find model parameters that give a small distance between a model simulation and the data on chl *a* concentration. We determined the prior distributions (for endpoints of these distributions, see endpoints of the *x*-axis in Fig. A1) for ϕ 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 σ and α are arbitrary. Regarding σ, the prior distribution can reflect the error estimates for the chl *a* measurements (which we did not have). For α, 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 α are concentrated on smaller values for each year (Fig. A1).

Our model's sensitivity to parameters can be studied by looking at the posterior distributions (Fig. 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 panel in Fig. A1) for each of the parameters, we can conclude that the model is not very sensitive to the phytoplankton concentration at $p(0)$, σ, or ϕ. It is not surprising that the model is sensitive to τ (i.e. the posterior distributions for τ have high peaks; Fig. A1) which can be considered as the main parameter characterizing the shape of the bloom (see below). In addition, the model exhibits qualitatively different behavior for different values of α, and our work here suggests that $\alpha < 1$ for each of the years included in this study.

As is shown in the preliminary simulations in Fig. 1, our model can generate qualitatively different bloom dynamics, for example, where $\Delta\mu$ exhibits either an increasing (Fig. 1a) or a decreasing (Fig. 1b) yearly trend. This is because of differences in τ that defines

the time window during which $\Delta\mu(t)$ is computed. Consequently, τ can also be interpreted as a representative of the events during τ days in the past that have an effect on the current division rate. Apart from 2 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 d. However, we note that this choice of a prior distribution for τ has an effect on the simulation results. For example, a model simulation with a large τ and ϕ generates an excellent agreement between the model and the data for the high level of chl *a* during late spring (Fig. 2c). However, the division rate predicted by the model reaches its maximum in early winter (Fig. 3c). While this example demonstrates the wide range of possible bloom dynamics generated by our minimalistic model, it also suggests that more work is required to determine biologically realistic upper and lower bounds for the prior distribution of τ. Here, we used 0 for the lower bound because we did not find a good biological justification for negative τ (i.e. taking into account the effect that future conditions have on the present chl *a* concentration). For the upper bound, based on our simulations with the model, a large τ value often resulted in a high peak bloom or suggested a peak chl *a* concentration in the middle of the winter (e.g. Fig. 2c,i). Therefore, we suggest choosing a high upper bound for τ if the data set 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 combined 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 understanding of a general representation for the dynamics of primary production. Here, the hypothesis that we tested emphasized the influence of grazing on phytoplankton population dynamics. Our model reproduces the yearly trend in the chl *a* concentration exhibited in the time series for several 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 1 mo, followed by low levels of chl *a* during the rest of the year (as exhibited in the chl *a* data for 1998). 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 $\Delta\mu(t)$. As a result, our model, which assumes that $\Delta\mu(t)$ 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 bloom dynamics. As a minimalistic mathematical representation for the dynamics of the seasonal chl *a* concentration, our model can be easily fitted to other data sets, 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 to test whether the disturbance-recovery hypothesis can explain the initiation and development of phytoplankton blooms and in future modeling studies in which a minimalistic representation of the seasonal unimodal primary production patterns is needed, for example, for simulating and predicting the dynamics of the predators of the primary producers.

LITERATURE CITED

Ardyna M, Babin M, Gosselin M, Devred E, Rainville L, Tremblay JÉ (2014) Recent Arctic Ocean sea ice loss triggers novel fall phytoplankton blooms. Geophys Res Lett 41:6207–6212

Backhaus JO, Hegseth EN, Wehde H, Irigoien X, Hatten K, Logemann K (2003) Convection and primary production in winter. Mar Ecol Prog Ser 251:1–14

Beaumont MA (2010) Approximate Bayesian computation in evolution and ecology. Annu Rev Ecol Evol Syst 41: 379–406

Beaumont MA, Cornuet JM, Marin JM, Robert CP (2009) Adaptive approximate Bayesian computation. Biometrika 96:983–990

Behrenfeld MJ (2010) Abandoning Sverdrup's critical depth hypothesis on phytoplankton blooms. Ecology 91: 977–989

Behrenfeld MJ (2014) Climate-mediated dance of the plankton. Nat Clim Chang 4:880–887

Behrenfeld MJ, Boss ES (2014) Resurrecting the ecological underpinnings of ocean plankton blooms. Annu Rev Mar Sci 6:167–194

Behrenfeld MJ, Falkowski PG (1997) Photosynthetic rates derived from satellite-based chlorophyll concentration. Limnol Oceanogr 42:1–20

Behrenfeld MJ, Doney SC, Lima I, Boss ES, Siegel DA (2013) Annual cycles of ecological disturbance and recovery underlying the subarctic Atlantic spring plankton bloom. Global Biogeochem Cycles 27:526–540

Boss E, Behrenfeld MJ (2010) In situ evaluation of the initiation of the north Atlantic phytoplankton bloom. Geophys Res Lett 37:L18603

Boyd PW, Watson AJ, Law CS, Abraham ER and others (2000) A mesoscale phytoplankton bloom in the polar southern ocean stimulated by iron fertilization. Nature 407:695–702

Boyd PW, Law CS, Wong CS, Nojiri Y and others (2004) The decline and fate of an iron-induced subarctic phytoplankton bloom. Nature 428:549–553

Clark JS (2005) Why environmental scientists are becoming Bayesians. Ecol Lett 8:2–14

Coale KH, Johnson KS, Fitzwater SE, Gordon RM and others (1996) A massive phytoplankton bloom induced by an ecosystem-scale iron fertilization experiment in the equatorial Pacific Ocean. Nature 383:495–501

Cushing DH (1959) The seasonal variation in oceanic production as a problem in population dynamics. ICES J Mar Sci 24:455–464

Dakos V, Beninca E, van Nes EH, Philippart CJ, Scheffer M, Huisman J (2009) Interannual variability in species composition explained as seasonally entrained chaos. Proc R Soc B 276:2871–2880

Ellertsen HC (1993) Spring blooms and stratification. Nature 363:24

Evans GT, Parslow JS (1985) A model of annual plankton cycles. Biol Oceanogr 3:327–347

Fischer AD, Moberg EA, Alexander H, Brownlee EF and others (2014) Sixty years of Sverdrup. Oceanography (Wash DC) 27:222

Franks PJS (2015) Has Sverdrup's critical depth hypothesis been tested? Mixed layers vs. turbulent layers. ICES J Mar Sci 72:1897–1907

Freund JA, Mieruch S, Scholze B, Wiltshire K, Feudel U (2006) Bloom dynamics in a seasonally forced phytoplankton–zooplankton model: trigger mechanisms and timing effects. Ecol Complex 3:129–139

Gao M, Shi H, Li Z (2009) Chaos in a seasonally and periodically forced phytoplankton–zooplankton system. Nonlinear Anal Real World Appl 10:1643–1650

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

Huppert A, Blasius B, Stone L (2002) A model of phytoplankton blooms. Am Nat 159:156–171

Huppert A, Blasius B, Olinky R, Stone L (2005) A model for seasonal phytoplankton blooms. J Theor Biol 236:276–290

Ji R, Jin M, Varpe Ø (2013) Sea ice phenology and timing of primary production pulses in the arctic ocean. Glob Change Biol 19:734–741

Klausmeier CA (2010) Successional state dynamics: a novel approach to modeling nonequilibrium foodweb dynamics. J Theor Biol 262:584–595

Lignell R, Haario H, Laine M, Thingstad TF (2013) Getting the right parameter values for models of the pelagic microbial food web. Limnol Oceanogr 58:301–313

Lindahl O, Andersson L, Belgrano A (2009) Primary phytoplankton productivity in the Gullmar Fjord, Sweden. An evaluation of the 1985–2008 time series. Report No. 6306. Swedish Environmental Protection Agency, Stockholm

McClain CR (2009) A decade of satellite ocean color observations. Annu Rev Mar Sci 1:19–42

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

Olsen E, Fay G, Gaichas S, Gamble R, Lucey S, Link J (2016) Ecosystem model skill assessment. Yes we can! PLOS ONE 11:e0146467

Platt T, White GN III, Zhai L, Sathyendranath S, Roy S (2009) The phenology of phytoplankton blooms: ecosystem indicators from remote sensing. Ecol Modell 220:3057–3069

Sathyendranath S, Ji R, Browman HI (2015) Revisiting Sverdrup's critical depth hypothesis. ICES J Mar Sci 72: 1892–1896

Scheffer M, Rinaldi S, Kuznetsov Y, van Nes EH (1997) Seasonal dynamics of *Daphnia* and algae explained as a periodically forced predator-prey systems. Oikos 80: 519–532

Shimoda Y, Arhonditsis GB (2016) Phytoplankton functional type modelling: Running before we can walk? A critical evaluation of the current state of knowledge. Ecol Modell 320:29–43

Siegel DA, Doney SC, Yoder JA (2002) The North Atlantic spring phytoplankton bloom and Sverdrup's critical depth hypothesis. Science 296:730–733

Stow CA, Jolliff J, McGillicuddy DJ Jr, Doney SC and others (2009) Skill assessment for coupled biological/physical models of marine systems. J Mar Syst 76:4–15

Sverdrup H (1953) On conditions for the vernal blooming of phytoplankton. ICES J Mar Sci 18:287–295

Taylor JR, Ferrari R (2011) Shutdown of turbulent convection as a new criterion for the onset of spring phytoplankton blooms. Limnol Oceanogr 56:2293–2307

Tiselius P, Belgrano A, Andersson L, Lindahl O (2016) Primary productivity in a coastal ecosystem: a trophic perspective on a long-term time series. J Plankton Res 38:1092–1102

Townsend DW, Keller MD, Sieracki ME, Ackleson SG (1992) Spring phytoplankton blooms in the absence of vertical water column stratification. Nature 360:59–62

Truscott J, Brindley J (1994) Ocean plankton populations as excitable media. Bull Math Biol 56:981–998

Tsuda A, Takeda S, Saito H, Nishioka J and others (2003) A mesoscale iron enrichment in the western subarctic Pacific induces a large centric diatom bloom. Science 300:958–961

van Nes EH, Scheffer M (2004) What minimal models cannot tell: a comment on a model of phytoplankton blooms. Am Nat 163:924–926

Winder M, Cloern JE (2010) The annual cycles of phytoplankton biomass. Philos Trans R Soc B 365:3215–3226

Zhang W, Arhonditsis GB (2008) Predicting the frequency of water quality standard violations using Bayesian calibration of eutrophication models. J Gt Lakes Res 34: 698–720

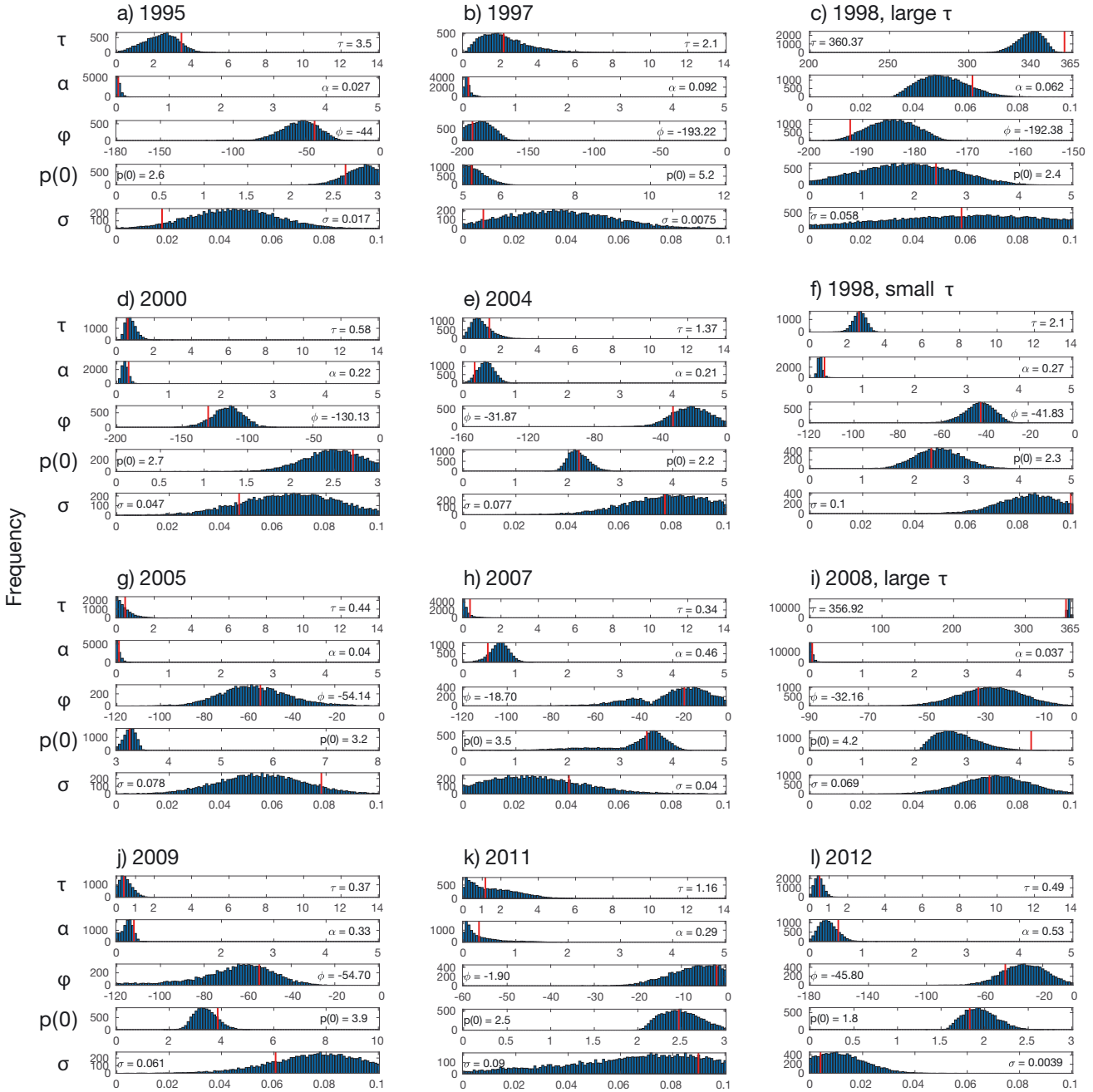**Appendix.** Posterior and prior distributions of the model parameters

Fig. A1. Posterior distributions of the model parameters for time offset, τ, proportionality constant, α, and the day of the year when the division rate exceeds (from below) half of its maximum value during the year, ϕ, the simulated chl a concentration at the beginning of the year, $p(0)$, and the simulated measurement error, σ, 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 used to infer parameters of our model (see Eq. 4) in Fig. 2. For each year, we give the numerical value of the parameter within each panel, 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 Figs. 2–4. Note that the value for τ in 1998 (large τ) and $p(0)$ in 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