



Bayesian inference to partition determinants of community dynamics from observational time series

C. M. Mutshinda^{1,4}, Z. V. Finkel², C. E. Widdicombe³ and A. J. Irwin¹

¹*Department of Mathematics and Statistics, Dalhousie University, Halifax, NS, Canada*

²*Department of Oceanography, Dalhousie University, Halifax, NS, Canada*

³*Plymouth Marine Laboratory, Prospect Place, Plymouth, PL1 3DH, UK*

⁴*Corresponding author; Email: crispin.mutshinda@dal.ca*

Keywords: Bayesian inference; Biotic interactions; Community dynamics; Environmental forcing; Markov chain Monte Carlo; Spike-and-slab prior.

Abstract: Ecological communities are shaped by a complex interplay between abiotic forcing, biotic regulation and demographic stochasticity. However, community dynamics modelers tend to focus on abiotic forcing overlooking **biotic interactions**, due to notorious challenges involved in modeling and quantifying inter-specific interactions, particularly for species-rich systems such as planktonic assemblages. Nevertheless, inclusive models with regard to the full range of plausible drivers are essential to characterizing and predicting community response to environmental changes. Here we develop a Bayesian model for identifying, from *in-situ* time series, the biotic, abiotic and stochastic factors underlying the dynamics of species-rich communities, focusing on the joint biomass dynamics of biologically meaningful groups. We parameterize a multivariate model of population co-variation with an explicit account for demographic stochasticity, density-dependent feedbacks, pairwise interactions, and abiotic stress mediated by changing environmental conditions and resource availability, and work out explicit formulae for partitioning the temporal variance of each group in its biotic, abiotic and stochastic components. We illustrate the methodology by analyzing the joint biomass dynamics of four major phytoplankton functional types namely, diatoms, dinoflagellates, coccolithophores and phytoflagellates at Station L4 in the Western English Channel using weekly biomass records and coincident measurements of environmental covariates describing water conditions and potentially limiting resources. Abiotic and biotic factors explain comparable amounts of temporal variance in log-biomass growth across functional types. Our results demonstrate that effective modelling of resource limitation and inter-specific interactions is critical for quantifying the relative importance of abiotic and biotic factors.

Abbreviations: AR model–Autoregressive model; BUGS–Bayesian inference Using Gibbs Sampling; LASSO–Least Absolute Shrinkage and Selection Operator; MCMC–Markov Chain Monte Carlo; PSU–Practical Salinity Unit; VAR model–Vector Autoregressive model.

1. Introduction

Fluctuations in population abundance or biomass within ecological communities result from a complex interplay of abiotic and biotic factors combined with demographic stochasticity or ecological drift (Lande et al. 2003, Loreau and de Mazancourt 2008, Mutshinda 2009, 2010). The abiotic factors involve environmental forcing mediated by changing physical conditions (e.g., temperature) and/or fluctuations in the availability of non-living resources affecting the growth, reproduction and maintenance of organisms. The biotic forces include density-dependent feedbacks and species-species interactions within and across trophic levels. Negative density dependent regulation is a pervasive feature of population growth processes (Nicholson 1933, Cooper 2003, Brook and Bradshaw 2006), and may result from competition for limiting resources among conspecifics (intra-specific competition) or from other mechanisms operating in a density-dependent fashion such as predation and diseases. **Demographic sto-**

chasticity refers to the variability in population growth resulting from randomness in outcomes of demographic events (births, deaths, dispersal) across individuals in a finite population (Lande et al. 2003).

Community dynamics models that integrate these various sources of variability enable better understanding of the causes of shifts in community structure and allow sensible predictions of community responses to environmental fluctuations. So far, models concerned with ecological responses to environmental changes have mostly focused on abiotic forcing with less regard to biotic interactions (Gilman et al. 2010). However, there is an emerging consensus among ecologists on the major role of biotic interactions in structuring ecological communities and modulating population responses to climate change (e.g., Davis et al. 1998a,b, Araujo and Luoto 2007, Heikkinen et al. 2007, Gilman et al. 2010, Götzenberger et al. 2012). **Yet, the modelling and quantification of biotic interactions is challenging**, particularly for species-rich systems such as microbial or planktonic assemblages due to the huge

number of potential interactions between species (Ovaskainen et al. 2017), raising the curse of dimensionality as the number of pairwise interactions increases exponentially with the number of populations under consideration. As a result, community dynamics models, typically vector autoregressive models (VAR) models, are often restricted to a few populations. VAR models, which extend univariate autoregressive (AR) models to multivariate time series, provide a flexible framework for analyzing the joint dynamics of co-occurring populations and evaluating, from observational time series, the biotic, abiotic, and stochastic factors drivers of population dynamics and community structure (Ives et al. 2003, Mutshinda et al. 2009, 2011, Hampton et al. 2013). A VAR model of order p , denoted as $VAR(p)$, treats each endogenous variable as a linear function of its own p lags and the lags of all other endogenous variables in the system. VAR models of any order $p > 1$ can be rewritten in the form, which is more convenient for analytical operations. Underlying autoregressive models of order 1 is the assumption that the future is independent of the past given the present, known as the *Markov property*.

VAR models of species-rich assemblages typically require some kind of complexity reduction to overcome the curse of dimensionality resulting from the huge number of parameters, particularly the inter-specific interactions coefficients (e.g., Ovaskainen et al. 2010, Pollock et al. 2014, Clark et al. 2018). Kissling et al. (2012) review complexity reduction techniques for modelling biotic interactions in species-rich assemblages, separating them into statistical and ecological methods. Statistical methods for complexity reduction hinge essentially on the assumption that weak interactions prevail in natural communities (Mutshinda et al. 2009, 2011) and accordingly, attempt to produce a sparse interaction matrix by resorting to sparsity-inducing methods. In the Bayesian framework (McCarthy 2007, Gelman et al. 2013), sparsity-inducing methods can be categorized into indicator-based variable selection methods and regularization or shrinkage techniques. Indicator-based variable selection techniques (e.g., George and McCulloch 1993) achieve sparsity by excluding the redundant predictors from the model. On the other hand, regularization methods (e.g., Park and Casella. 2008, Mutshinda and Sillanpää 2010, 2011, 2012) keep all potential predictors in the model, but impose a penalty constraining the magnitudes of their coefficients to shrink towards zero, causing the effects of redundant predictors to be zero or nearly so and virtually excluding such predictors from the model. The most popular regularization methods are the LASSO (Tibshirani 1996) and ridge regression (Hoerl and Kennard 1970) which constrain the L_1 and L_2 norm of the coefficient vectors, respectively. Because of its ability to set some of the coefficients exactly to zero, the LASSO is increasingly popular as a variable selection tool. We refer to O'Hara and Sillanpää (2009) for an extensive review of Bayesian variable selection methods. The rationale of ecological methods for complexity reduction is to aggregate the plethora of species under consideration in a few biologically meaningful groups such as organismal taxonomic units or functional types, and focus on group-level data. This is pragmatic since it is much easier to deal with a handful of groups than many individual

species. In addition, species-level data are usually fraught with missing values. The aggregation of species into a few biologically meaningful groups may also be required by the question being addressed. For instance, when designing models for predicting phytoplankton biomass or characterizing traits, it is useful to aggregate the host of species into functional types based on distinctive ecological functionality or chemical roles, which are sensible proxies for ecological and biogeochemical functions (Le Quéré et al. 2005, Irwin and Finkel 2018).

In this paper, we develop a Bayesian model for evaluating, from long-term monitoring data, the factors and mechanisms underlying the structure and dynamics of species-rich communities, focusing on the joint biomass dynamics of biologically meaningful groups. We parameterize a multivariate model of group biomass co-variation integrating demographic stochasticity, density-dependent effects, pairwise interactions, and abiotic stress mediated by changing environmental conditions and limiting resources. We devise formulae for partitioning the temporal variance in log-biomass growth of each group in its biotic, abiotic and stochastic components. We illustrate the method by analyzing the joint biomass dynamics of four major phytoplankton functional types namely, diatoms, dinoflagellates, coccolithophores and phytoflagellates using long-term data recorded at Station L4 in the Western English Channel as part of the Western Channel Observatory oceanographic time series (Widdicombe et al. 2010a,b). The data involve weekly biomass concentrations of the four functional types and coincident measurements of five environmental covariates. Two of the environmental covariates (temperature and salinity) describe environmental conditions whereas the other three (irradiance, nitrogen, and silicate) represent potentially and limiting resources.

2. Materials and methods

In the sequel, boldface lower-case letters and boldface capital letters represent vectors and matrices, respectively.

2.1. Model specification

We describe the biomass dynamics of each taxonomic group by a stochastic Gompertz model extended to incorporate pairwise interactions and external forcing mediated by changing environmental conditions (temperature or salinity) and potentially limiting abiotic resources (irradiance, nitrogen, and silicate). The separation of environmental covariates into environmental conditions and limiting resources is important since they enter the model differently. Henceforth, we refer to the covariates describing environmental conditions as environmental conditions. Let $Y_{i,t}$ and $X_{q,t}$ denote respectively, the biomass of the i th taxonomic group ($i = 1, \dots, n$) and the Z-score of the q th environmental condition at time t ($t \geq 2$). We assume that

$$Y_{i,t} = Y_{i,t-1} \exp \left\{ r_{i,t} \left[1 + \sum_{j=1}^n \frac{\alpha_{ij}}{k_i} \ln(Y_{j,t-1}) \right] + \sum_{q=1}^Q X_{q,t} \beta_{iq} + \varepsilon_{i,t} \right\} \quad (1)$$

where $r_{i,t} > 0$ denotes the density-independent growth rate of the i th group from time $t-1$ to t at average environmental conditions, $\beta_{i,q}$ represents the effect of the q th environmental condition (temperature or salinity) on the i th group, and $\varepsilon_{i,t}$ is a zero-mean random disturbance affecting the biomass dynamics of the i th group at t . The interaction coefficient $\alpha_{i,j}$ ($i \neq j$) represents the effect of the j th group's biomass on the growth rate of the i th group, relative to the effect of the i th group's biomass on its own growth rate. The scaling factor $k_i > 0$ is the carrying capacity of the i th group's log-biomass. We set all intraspecific interaction coefficients $\alpha_{i,i}$ to -1 for identifiability purposes. In its typical form, the Gompertz model has a minus sign in front of the density-dependent term

$$\sum_{j=1}^n \frac{\alpha_{ij}}{k_i} \ln(Y_{j,t-1})$$

appearing in the exponent of equation (1), with all intraspecific interaction coefficients $\alpha_{i,i}$ set to 1 (Mutshinda et al. 2009, 2011). Here we use a positive sign instead so that negative and positive coefficients portray detrimental and beneficial impacts respectively, hence the choice of -1 for the intraspecific coefficients. Depending on its sign and magnitude, the inter-specific interaction can be detrimental to both participants (competition), detrimental to one and beneficial to the other (exploitation or antagonism), detrimental to one without cost or benefit to the other (amensalism), of no effect to either participant (neutralism), or beneficial to at least one of the participants with no harm to either (facilitation). Facilitation can take the form of mutualism if both participants benefit or commensalism when it only benefits one of them (Stachowicz 2001).

For purposes of estimating the model parameters, it is convenient to work on a logarithmic scale. Re-writing (1) on the natural logarithmic scale with $y_{i,t}$ representing the natural logarithm of $Y_{i,t}$ yields

$$y_{i,t} = y_{i,t-1} + r_{i,t} \left[1 + \sum_{j=1}^n \frac{\alpha_{ij}}{k_i} y_{j,t-1} \right] + \sum_{q=1}^Q X_{q,t} \beta_{i,q} + \varepsilon_{i,t} \quad (2)$$

Denoting the change (growth) $y_{i,t} - y_{i,t-1}$ in the log biomass of the i th group from time $t-1$ to t by $\Delta y_{i,t}$, it follows from (2) by direct differentiation and re-arranging of terms that

$$\alpha_{i,j} = \left[\frac{\partial \Delta y_{i,t}}{\partial y_{j,t-1}} \right] \left[\frac{\partial \Delta y_{i,t}}{\partial y_{i,t-1}} \right]^{-1} \quad (3)$$

Consequently, the inter-specific interaction coefficient $\alpha_{i,j}$ quantifies the effect of group j on the growth rate of group i , relative to the (negative) feedback $\partial \Delta y_{i,t} / \partial y_{i,t-1}$ of the i th group's biomass on its own growth rate from time $t-1$ to t , whose magnitude is $h_{i,t} = r_{i,t} |\alpha_{i,i}| / k_i$ or simply $r_{i,t} / k_i$ since $|\alpha_{i,i}| = 1$ by design.

In compact matrix notation, the joint dynamics model (2) for a n -group system reads

$$\mathbf{y}_t = \mathbf{y}_{t-1} + \mathbf{R} (\mathbf{1}_n + \mathbf{A} \mathbf{y}_{t-1}) + \boldsymbol{\beta} \mathbf{x}_t + \boldsymbol{\varepsilon}_t \quad (4)$$

where $\mathbf{y}_t = (y_{1,t}, \dots, y_{n,t})^T$ and $\mathbf{x}_t = (X_{1,t}, \dots, X_{Q,t})^T$ denote respectively the n -vector of group-level log-biomasses and the Q -vector of Z-scores of the environmental conditions at time t , \mathbf{R} is the $n \times n$ diagonal matrix with $\mathbf{R}(i,i) = r_{i,t}$, $\mathbf{1}_n$ is the n -vector of ones, \mathbf{A} is the $n \times n$ matrix with (i,j) th entry $A_{ij} = \alpha_{ij} / k_i$, $\boldsymbol{\beta}$ is the $n \times Q$ matrix with i th row $(\beta_{i,1}, \dots, \beta_{i,Q})$, and $\boldsymbol{\varepsilon}_t = (\varepsilon_{1,t}, \dots, \varepsilon_{n,t})$ is the n -vector of zero-mean random disturbances affecting the biomass dynamics of the n groups at time t . We assume that the error vectors $\boldsymbol{\varepsilon}_t$ ($t \geq 2$) are serially independent, but allow the contemporaneous components $\varepsilon_{1,t}, \dots, \varepsilon_{n,t}$ of $\boldsymbol{\varepsilon}_t$ to covary, so that we can evaluate potential dependencies among groups that the involved covariates may not capture. More specifically, we assume that each $\boldsymbol{\varepsilon}_t$ is multi-normally distributed around the n -dimensional zero vector with a potentially non-diagonal covariance matrix \mathbf{V}_t . The residual variances (the diagonal entries of \mathbf{V}_t) lump together the group-specific variances due to potentially important covariates that may be omitted from the model with demographic stochasticity and sampling error. Demographic stochasticity can be extracted by capitalizing on its inverse scaling with the population size or biomass. If we assume that sampling variation is negligible, then we can decompose the residual covariance matrix as

$$\mathbf{V}_t = \mathbf{D}_t + \mathbf{C} \quad (5)$$

where \mathbf{D}_t is a $n \times n$ time-dependent diagonal matrix, with $D_t(i,i) = \delta_i / Y_{i,t-1}$ representing the demographic stochasticity affecting the biomass dynamics of the i th group from time $t-1$ to t , and $\delta_i > 0$ denoting the demographic variance parameter specific to the i th group. Due to its inverse scaling with the population size or biomass, demographic stochasticity is only relevant for small populations (e.g., Saether et al. 2000, Lande et al. 2003, Mutshinda et al. 2009, 2011). The second component of \mathbf{V}_t is the $n \times n$ time-independent residual environmental covariance matrix \mathbf{C} whose diagonal elements $C_{i,i}$ and off diagonal elements $C_{i,j}$ ($i \neq j$) represent respectively the group-specific and joint responses to un-modelled environmental factors.

Resource limitation to the growth of individual groups is not explicit in the model described by equations (1) and (2), but is potentially important. To facilitate the variance decomposition analysis, we consider two different versions of the model namely, a reduced model represented by equations (1) and (2) with $r_{i,t} = r_i$ (independently of time), and a full model involving an explicit account of resource limitation. In the full model, we incorporate resource limitation by letting $r_{i,t} = \mu_i f_i(t)$, where $\mu_i > 0$ is the maximum net growth rate of the i th group from one time to the next at average environmental conditions (temperature and salinity) and optimal resource combination, and $f_i(t)$ is a suitable saturation function with $0 < f_i(t) \leq 1$. The most common form of growth limiting term is the Michaelis-Menten saturation function also called the Monod or hyperbolic equation (Irwin and Finkel 2018). When considering under this form of growth limitation a single limiting resource P taking the value P_t at time t , $f_i(t)$ is simply the Michaelis-Menten term $P_t / (KP_i + P_t)$, where $KP_i > 0$ denotes the resource half-saturation constant representing the resource level at which $r_{i,t} = \mu_i / 2$. When dealing with more than one limiting resource, there are different ways of modelling co-limitation. Here we combine Michaelis-Menten

factors of the resources under consideration with the minimum function so that each functional type grows at the rate permitted by the scarcest resource (the resource with the lowest Michaelis-Menten term), according to Liebig's law (de Baar 1994, van der Ploeg and Kirkham 1999).

2.2. Model fitting

Since the joint posterior distribution of the model parameters is not available in closed-form, numerical algorithms such as MCMC samplers (Gilks et al. 1996) are required to simulate from it. MCMC allow one to simulate the parameter(s) of interest from the correct posterior distribution without the need to know its exact mathematical form. In practice, MCMC simulation is conveniently implemented using freely available software packages such as WinBUGS/OpenBUGS (Lunn et al. 2000, Thomas et al. 2006), JAGS (Plummer 2003), or Stan (Stan Development Team 2018). These software packages require the analyst to specify the model in R-like code that the program uses to produce posterior MCMC samples according to the prescribed prior distribution and likelihood or sampling distribution of the data at hand.

2.3. Variance partitioning

Based on the reduced model (2), the growth rate $\Delta y_{i,t}$ of the i th group can be expressed as a linear function of the lagged log-biomasses $y_{j,t-1}$ of all populations and coincident values $\mathbf{x}_t = (X_{1,t}, \dots, X_{Q,t})$ of the environmental conditions under consideration at time t , with Gaussian additive noise. This facilitates the decomposition of the temporal variance in log-biomass growth of each group into additive contributions from different components. Accordingly, we use the reduced model to partition the temporal variance in the log-biomass dynamics of individual types into contributions from biotic factors, abiotic forcing and demographic stochasticity. We then introduce resource limitation and evaluate its contribution to temporal variation in the log-biomasses of individual groups by the subsequent reduction of the diagonal entries of the residual environmental covariance matrix \mathbf{C} (the group-specific residual environmental variances).

Based on the reduced model where $r_{i,t} = r_i$, we decompose the total variance T_i in the log-biomass growth of the i th group over the study period as $T_i = G_i + E_i + D_i$, where G_i , E_i and D_i , and denote respectively the contributions of negative density feedback, environmental variability (including inter-specific interactions), and demographic stochasticity. Letting \bar{y}_i represent the average log-biomass of the i th type over the study period, an estimate of demographic stochasticity in the log-biomass dynamics of the i th type at average log-biomass is

$$D_i = \delta_i / \exp(\bar{y}_i) \quad (6)$$

If we denote by n_i the temporal variance in the log-biomass of the i th type over the study period, and by $\mathbf{x}_q = (X_{q,1}, \dots, X_{q,T})$, then

$$G_i = \left(\frac{r_i}{k_i}\right)^2 v_i \quad (7)$$

$$E_i = \sum_q \beta_{iq}^2 \text{var}(\mathbf{x}_q) + \left(\frac{r_i}{k_i}\right)^2 \sum_{j \neq i} \alpha_{ij}^2 v_j + C_{ii} \quad (8)$$

Since the forcing variables $\mathbf{x}_q = (q = 1, \dots, Q)$ are mean-centered and standardized to unit variance, the formula for E_i simplifies as

$$E_i = \sum_q \beta_{iq}^2 + \left(\frac{r_i}{k_i}\right)^2 \sum_{j \neq i} \alpha_{ij}^2 v_j + C_{ii} \quad (9)$$

From the above variance partitioning, it follows that the proportion of temporal variation in the log-biomass growth of the i th group attributable to biotic factors (density-dependent regulation and pairwise interactions) is

$$B_i = [G_i + \left(\frac{r_i}{k_i}\right)^2 \sum_{j \neq i} \alpha_{ij}^2 v_j] / T_i \quad (10)$$

The proportion of temporal variance in the log-biomass growth of the i th group due to environmental variability (including inter-specific interactions) is E_i/T_i . Similarly, the proportion of temporal variance in the log-biomass growth of the i th group due to density-dependent regulation is G_i/T_i and, at the average log-biomass of the i th group, the proportion of temporal variance in the log-biomass growth of the i th type due to demographic stochasticity is D_i/T_i . Alternatively, we can also separate the temporal variance in the log-biomass of the i th group excluding demographic stochasticity, $(T_i - D_i)$ in its abiotic and biotic components given by

$$\sum_q \beta_{iq}^2 + C_{ii} \text{ and } G_i + \left(\frac{r_i}{k_i}\right)^2 \sum_{j \neq i} \alpha_{ij}^2 v_j,$$

respectively. We can also decompose the biotic component of temporal variation in the log-biomass of the i th group into contributions G_i from density-dependence and

$$\left(\frac{r_i}{k_i}\right)^2 \sum_{j \neq i} \alpha_{ij}^2 v_j$$

from inter-specific interaction among types.

3. Application of the Bayesian model to the L4 phytoplankton community

3.1. Background

Phytoplankton is a generic name for a highly diverse group of photosynthetic organisms found in the euphotic zone of aquatic environments. As the base of the marine food web, phytoplankton sustain most of the aquatic life from zooplankton through fish to marine mammals (e.g., Doney 2006). They also play a critical role in global biogeochemical cycles, particularly in the global carbon cycle through the "biological pump", which accounts for roughly half of the total carbon fixation on Earth (Field et al. 1998). Therefore, identifying the factors and mechanisms underlying phytoplankton biomass dynamics and community structure and evaluating their relative importance is essential if we are to understand the

functioning of marine ecosystems and predict the structure of planktonic communities in a changing ocean (Irwin and Finkel 2018). Here we analyze the joint biomass dynamics of four major phytoplankton functional types namely, diatoms, dinoflagellates, coccolithophores and phytoflagellates using long-term monitoring data from a well-studied coastal station in the Western English Channel.

3.2. Description of data

We consider weekly observations of taxonomically resolved phytoplankton biomass (mg C m^{-3}), along with co-extensive measurements of five environmental covariates recorded at 10 m depth in the upper mixed layer at Station L4 ($50^\circ 15.00'N$, $4^\circ 13.02'W$) as part of the long-term Western Channel Observatory oceanographic time series (Widdicombe et al. 2010a,b). The data cover the period spanning 14 April 2003 through 31 December 2009, a period over which all the required data are available. Two of the five environmental variables under consideration namely, temperature ($Temp$, $^\circ\text{C}$) and salinity (Sal , PSU) describe water conditions. The temperature values range from 7.5°C to 18.9°C with a mean value of 13°C , whereas all salinity values fall in a narrow range from 34.20 to 35.54 PSU with a mean value of 35.10 PSU (Fig. 1). The other three variables, photosynthetic active radiation (PAR ; $\text{mol m}^{-2} \text{d}^{-1}$), nitrogen concentration (Nit , $\mu\text{mol L}^{-1}$) and silicate concentration (Sil , $\mu\text{mol L}^{-1}$) characterize resource availability. Time-series plots of the five environmental variables at Station L4 over the study period reveal strong seasonal fluctuations (Fig. 1). Irradiance and temperature are highly correlated, but the two variables enter the model differently, which helps mitigate potential identifiability issues.

Phytoplankton biomass concentrations were calculated using taxa-specific biovolumes that were converted to carbon according to the equations of Menden-Deuer and Lessard (2000) (Widdicombe et al. 2010b). For the purpose of the present study, we aggregate the phytoplankton into four functional types namely, diatoms, dinoflagellates, coccolithophorids and phytoflagellates and show the temporal variability of the natural log of each type's biomass in Figure 2. Each group exhibits large variability both seasonally and inter-annually, with diatoms typically peaking in spring or early summer, while dinoflagellates are more common in summer. The coccolithophore biomass is highly variable and attributed to the dominance of the bloom-forming *Emiliania huxleyi*. The phytoflagellates dominated (more than 50% of the biomass) by unidentified flagellates less than $5\mu\text{m}$ in diameter show the least temporal variability relative to other functional types (Fig. 2).

3.3. Prior specification and model fitting

Before completing the model specification with explicit statements of priors on all unknown quantities, it is worth mentioning the postulated form of resource co-limitation to the growth rates of individual functional types in the full model. Following Mutshinda et al. (2017), we assume that the realized density-independent and resource-limited growth rate of the i th functional type from week $w-1$ to w is given by $r_{i,w} = \mu_i f_i(w)$, where $\mu_i > 0$ is the maximum net growth rate of the i th functional type from one week to the next at average environmental conditions and optimal resource combination, and

$$f_i(w) = \min \left[\frac{PAR_w}{PAR_w + KE_i}, \frac{Nit_w}{Nit_w + KN_i}, \frac{Sil_w}{Sil_w + KS_i} \right] \quad (11)$$

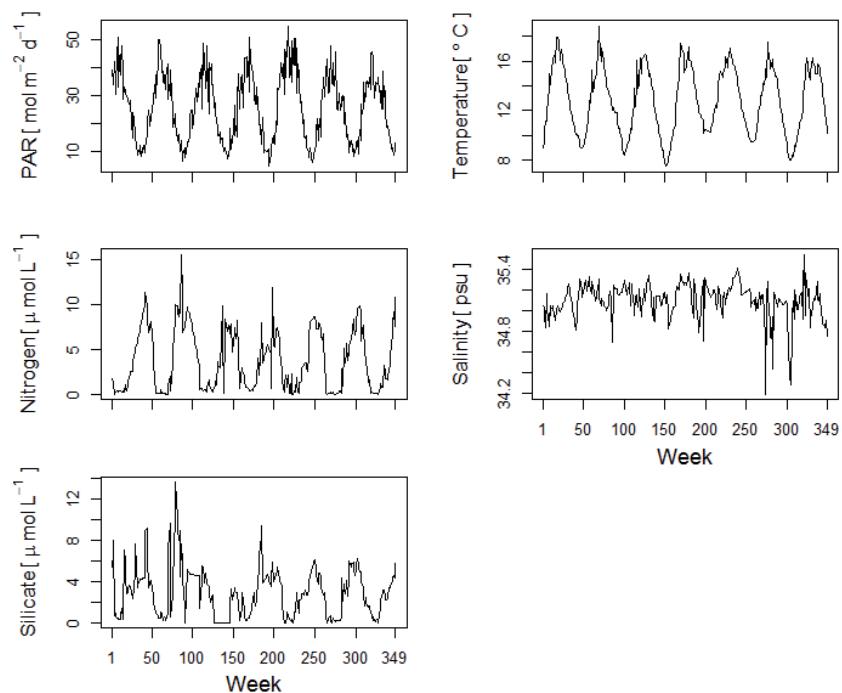


Figure 1. Timeplots of the abiotic variables at Station L4 between April 2003 and December 2009.

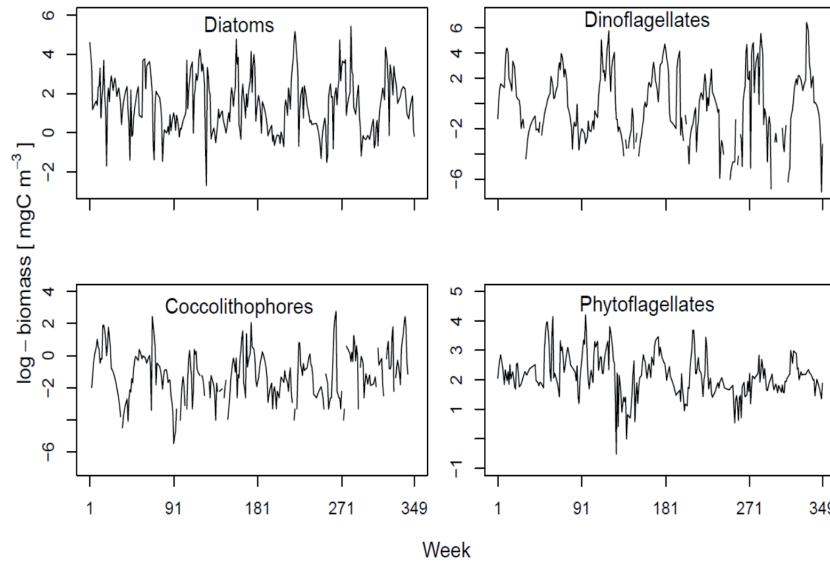


Figure 2. Timeplots of the diatom, dinoflagellate, coccolithophore, and phytoflagellate biomass (mg C m^{-3}) at Station L4 over the study period on the natural logarithmic scale.

where $KE_i > 0$, $KN_i > 0$ and $KS_i > 0$ denote respectively the irradiance, nitrogen and silicate half-saturation constants representing the level of each resource at which $r_i, w = \mu_i/2$, provided that none of the other resources is limiting.

It is good practice to incorporate available information into the analysis by specifying informative prior distributions for some of the model parameters, which can greatly improve the estimation precision (Delean et al. 2013). Based on our previous experience with the L4 data (Mutshinda et al. 2016, 2017), we assign $\text{Gamma}(6,4)$ priors to the intrinsic growth rates r_i and the carrying capacity k_i , and place $\text{inv-Gamma}(0.1, 0.1)$ priors on the group-specific demographic variance parameters δ_i independently for $i = 1, \dots, 4$. We assume the interspecific interaction coefficients $\alpha_{i,j}$ and the environmental effects $\beta_{i,q}$ to be *a priori* normally distributed around zero with respective variances $\phi_{i,j}$ and $\rho_{i,q}$, where $\phi_{i,j} \sim \text{Exp}(\lambda_\alpha)$ and $\rho_{i,q} \sim \text{Exp}(\lambda_\beta)$. A convenient prior for a covariance matrix, which guaranteeing the positive-definiteness of the posterior, is the inverse-Wishart distribution (Gelman et al. 2013). We assign to the residual covariance matrix C an inverse-Wishart prior distribution with scale matrix $S = I_4$, the 4×4 identity matrix, and 4 degrees of freedom select to be the smallest possible i.e., the rank of the scale matrix S , to convey a lack of relevant prior information. The specification of the reduced model is completed with prior statements for the hyper-parameters λ_α and λ_β , for which we assume the fairly uninformative $\text{Gamma}(0.1, 0.1)$ priors independently.

We complete the specification of the full model with prior assumptions about the maximum net growth rates μ_i and the half-saturation constants for irradiance, nitrogen and silicate. Again, drawing on our previous experience with the L4 data, we assign $\text{Gamma}(6,4)$ on the maximum net growth rates μ_i of each functional type and impose positively truncated normal priors $N(15,100)I(0,\infty)$ and $N(0,0.01)I(0,\infty)$ on the irradiance and nitrogen half-saturation constants for each functional type, respectively.

Silicate is an important limiting nutrient for diatoms, but it is unclear whether it is worth keeping as a covariate in

models describing the biomass dynamics of the three other functional types. To find this out, we impose on the silicate half-saturation constants positively truncated normal mixture prior with a spike-and-slab structure (George and McCulloch 1993) to implement variable selection. More specifically we assume that $KS_i \sim N(0, s_i)I(0, \infty)$, where $s_i = (1 - \xi_i) * 0.001 + \xi_i * 10$, where $\xi_i \in [0, 1]$ is the inclusion indicator taking the value one when silicate is required in the dynamics of the i th functional type and the value zero otherwise, whereas 0.001 and 10 are the variances of the spike and the slab components of our mixture prior, respectively. We assign $\text{Bernoulli}(0.25)$ priors independently on the silicate inclusion indicators ξ_i and estimate them from the data alongside the other model parameters. We use MCMC to simulate, through OpenBUGS, from the joint posterior of the unknown quantities and base our inferences on posterior MCMC samples. We ran 10,000 iterations of three parallel Markov chains following a burn-in period of 4,000 iterations and thinned the post burn-in samples by a factor of 10. We assessed convergence of the Markov chains through visual inspection of traceplots and posterior autocorrelation plots. After about 2,000 iterations, the chains were mixing well, jumping freely around the parameter space. Accordingly, we set the burn-in period to 4,000.

3.4. Results

We start by presenting the results of the reduced model without an explicit account for resource limitation. We then consider the full model to evaluate the proportion of temporal variance in the log-biomass dynamics of each phytoplankton functional type that is attributable to the incorporation of resource limitation.

3.4.1. Results of the reduced model. The intrinsic growth rate (r_i ; week^{-1}) represents the net density independent realized growth rate of the i th functional type at average environmental conditions over the time series (13°C and 35.10 PSU). Aggregated together, diatoms have the highest intrinsic growth rate with a doubling time of 6 days, followed by phyto-

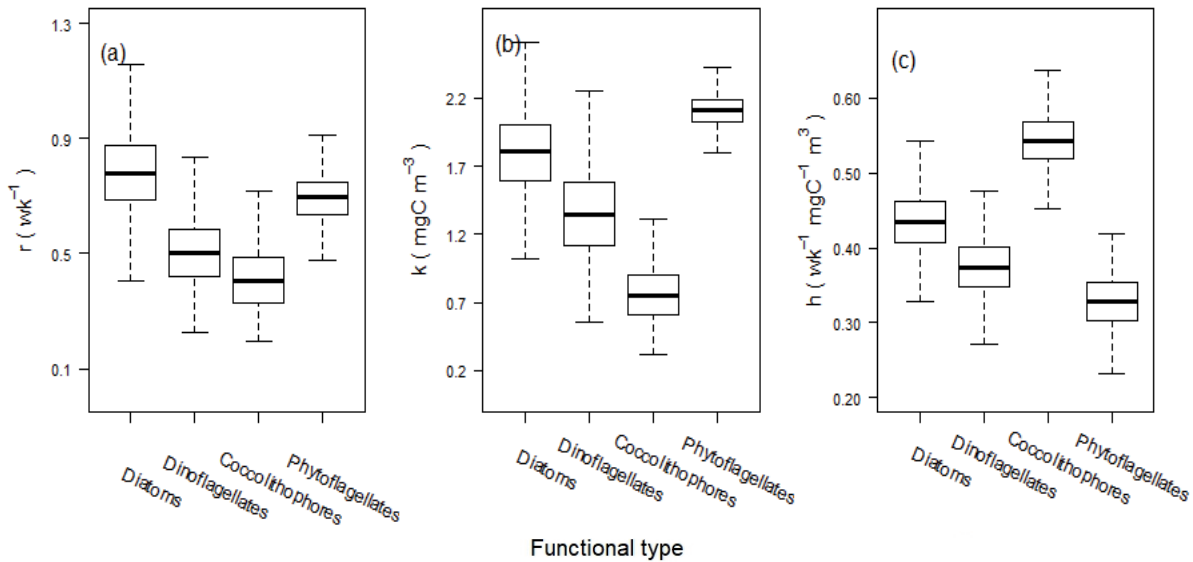


Figure 3. Boxplots summarizing the posterior distributions of intrinsic growth rates (a), carrying capacities (b), and subsequent density-dependent feedback (c) for diatoms, dinoflagellates, coccolithophores and phytoflagellates. Each boxplot shows the median (thick line), the inter-quartile range (box) and the full range of the distribution spanning 1.5 times the inter-quartile below the first quartile (Q_1) to 1.5 times inter-quartile range above the third quartile (Q_3), excluding outliers.

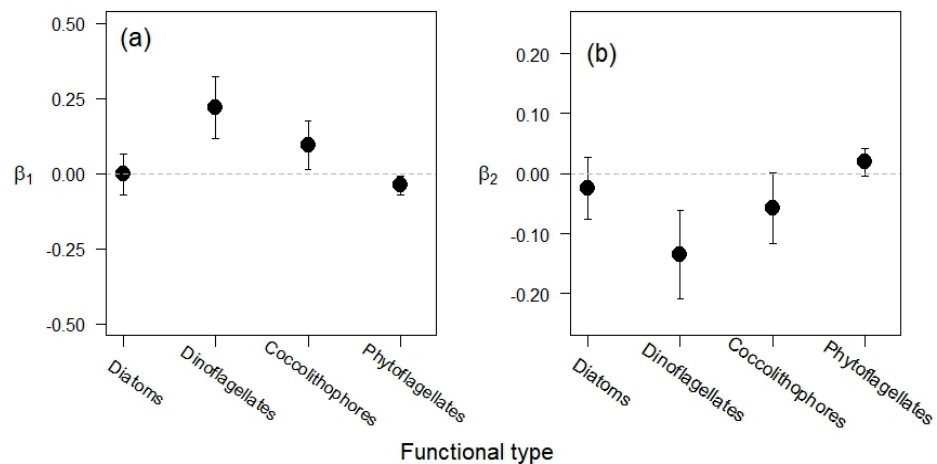


Figure 4. Error bars (posterior mean \pm 1 sd) of (a) the temperature and (b) salinity effects on growth rates of each functional type.

flagellates and dinoflagellates with respective doubling times of 7 and 10 days. Coccolithophores have the lowest intrinsic growth rate, which is half the diatom intrinsic growth rate, with a doubling time of 12 days (Fig. 3a). Phytoflagellates have the highest carrying capacity (k) followed by diatoms and dinoflagellates, while coccolithophores have the lowest carrying capacity (Fig. 3b). The magnitude of the negative density feedback on each functional type's growth rate, as measured by $h = r/k$, is strongest on coccolithophores and weakest on phytoflagellates (Fig. 3c). If we assume that the density-dependent losses are predominantly due to grazing, this result indicates a higher grazing vulnerability for coccolithophores as opposed to phytoflagellates, consistent with the finding of Mutshinda et al. (2017) that at Station L4, coccolithophores undergo a strong control by both specialist and generalist grazers, when essentially only specialist grazers control the other three functional types.

The growth rates of diatoms and phytoflagellates are essentially unresponsive to small changes in temperature and salinity around the average conditions (13°C and 35.10 PSU), whereas dinoflagellates and coccolithophores respond positively to increasing temperature and negatively to increasing salinity (Fig. 4). However, dinoflagellates are more responsive than coccolithophores to changing water conditions with magnitudes of temperature and salinity effects roughly twice those of coccolithophores, implying faster dinoflagellate biomass accumulation at warmer temperatures, in line with earlier results found at this site (Widdicombe et al. 2010a, Mutshinda et al. 2017). The tight coupling between temperature and dinoflagellate biomass at L4 may explain the clear seasonal patterns in dinoflagellate log-biomass (Fig. 2).

The posterior densities of the interaction coefficients for pairs of functional types summarize the interactions between functional types (Fig. 5).

Table 1. Summary of inferred functional type interactions derived from the signs and magnitudes of the interaction coefficients reported in Figure 5. The sign of the pairwise interaction on each type in the pair is shown in parentheses.

| Functional type pair | Interaction |
|---------------------------------|---------------------|
| Diatom-Dinoflagellate | Mutualism (+,+) |
| Diatom-Coccolithophore | Neutral (0, 0) |
| Diatom-Phytoflagellate | Commensalism (0, +) |
| Dinoflagellate-Coccolithophore | Mutualism (+, +) |
| Dinoflagellate-Phytoflagellate | Amensalism (-, +) |
| Coccolithophore-Phytoflagellate | Competition (-, -) |

Nine of the twelve interspecific interaction coefficients have posterior means at least one standard deviation away from zero. Considering these as statistically important, our results demonstrate a variety of interaction mechanisms among the study functional types at Station L4 with a prevalence of facilitation (Table 1).

The posterior means (standard deviations in parentheses) of the demographic variance parameter for diatoms, dinoflagellates, coccolithophores, and phytoflagellates are 0.47 (0.13), 0.06 (0.01), 0.07 (0.01), and 0.43 (0.10), respectively.

Overall, environmental variability, including the variability due to changing temperature and salinity, inter-specific interactions and residual environmental variances account on average for 33 to 62% of temporal variation in the dynamics

of individual functional types. Density-dependent feedbacks and demographic stochasticity account on average for 21 to 56% and 4 to 20% of temporal variation in the dynamics of individual functional types, respectively (Fig. 6).

If we exclude demographic stochasticity and consider the residual variation captured by the covariance matrix as essentially due to un-modelled abiotic factors, then both abiotic and biotic factors explain substantial amounts of temporal variation in log-biomass growth across functional types, with abiotic forces dominating in all functional types but coccolithophores (Fig. 7).

Most of the biotic component of temporal variance is due to density-dependence, which on average accounts for 84%, 94%, 92% and 83% of biotic variability in diatoms, dinoflagellates, coccolithophores and phytoflagellates respectively.

3.4.2. Results of the full model. The incorporation of resource limitation allowed us to estimate the maximum net growth rate, of each functional type at average conditions (13°C and 35.10 PSU) in nutrient- and light-replete waters. Diatoms have the highest maximum net growth rate followed by phytoflagellates, dinoflagellates and coccolithophores with mean doubling times 2.5, 3.9, 5, and 5.4 days respectively, which are much shorter than the 6-12 day doubling times reported from the reduced model. The posterior distributions of the carrying capacities are consistent between the full and the reduced model.

Near average conditions (13°C and 35.10 PSU), diatoms and phytoflagellates are, as types, essentially unaffected by

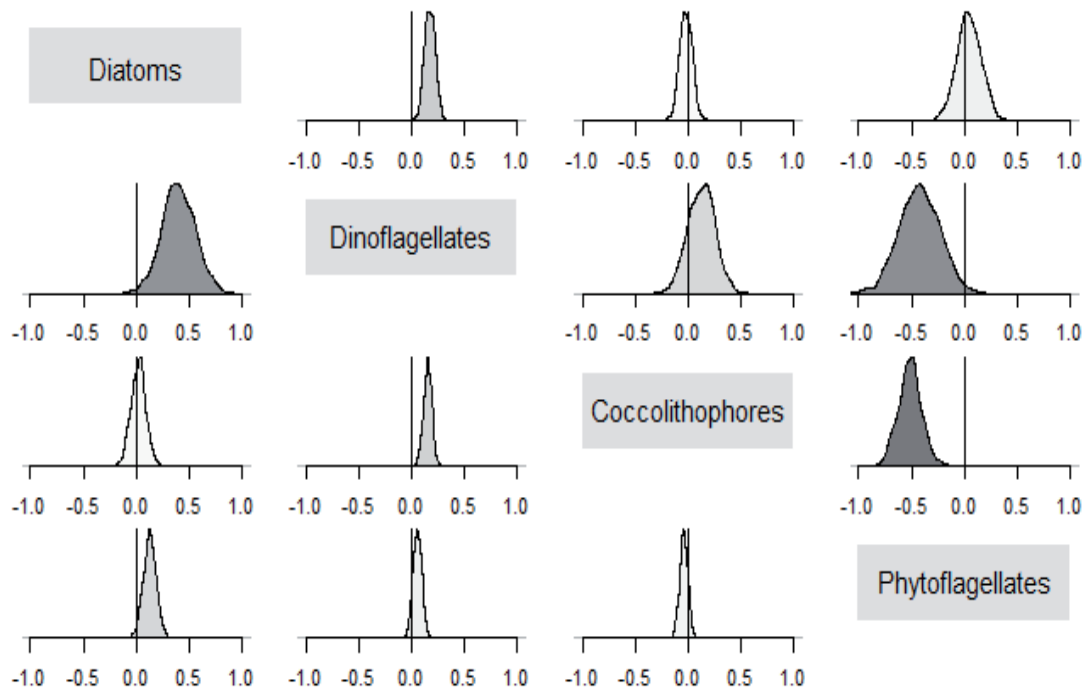


Figure 5. Posterior densities of pairwise interaction coefficients, $\alpha_{i,j}$, with all intra-type ($i = j$) coefficients set equal to -1 to serve as reference for interpreting the inter-type interactions. Panel (i,j) displays the posterior density of the interaction coefficient of the functional type whose name appears on the j -th column on the type whose name appears on the i -th row, with darker shading indicating a distribution with a posterior mean further from 0.

Figure 6. Barplots for the decomposition of total temporal variance in the dynamics of individual functional types based on the reduced model with no explicit account for resource limitation. The contributions of demographic stochasticity, density-dependent regulation and environmental variability are represented by (left), (middle), and (right) bars.

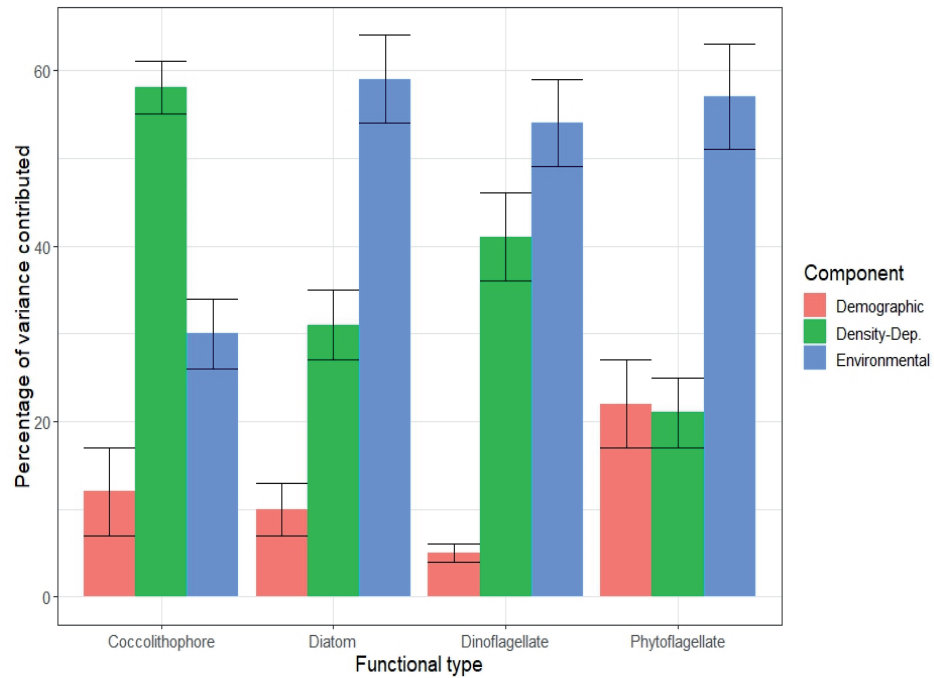
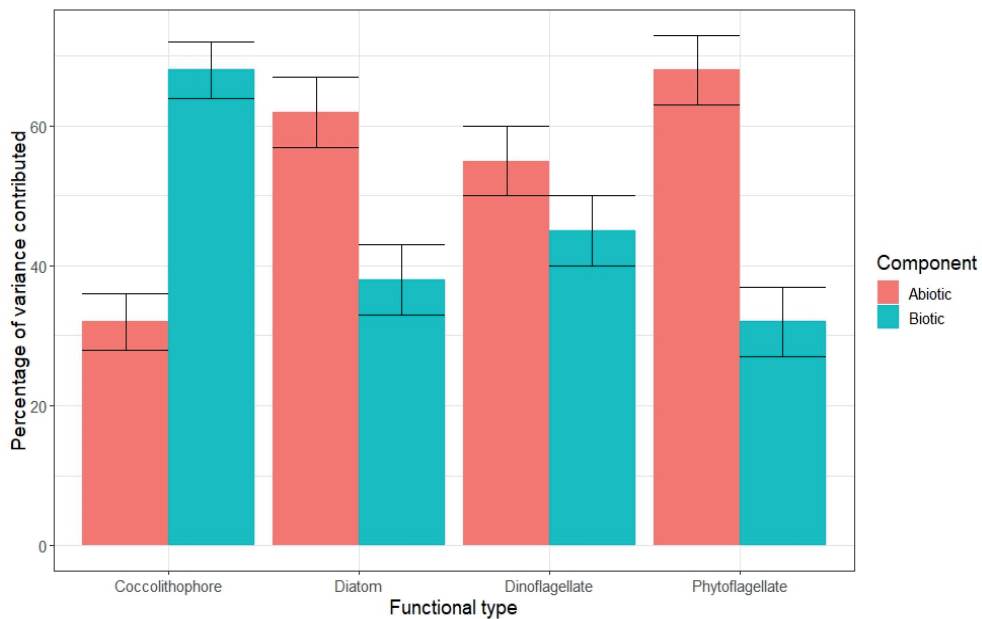


Figure 7. Decomposition of total temporal variance (excluding demographic stochasticity) into contributions from abiotic (left) and biotic (right) factors based on the reduced model with no explicit account for resource limitation.



small changes in temperature and salinity, whereas dinoflagellates and coccolithophores respond positively to increasing temperature and negatively to increasing salinity. The posterior distributions of the temperature and salinity effects are consistent between the full and the reduced model (Fig. 8a,b). For all functional types, the posterior estimates of half saturation constants for inorganic nitrogen are very low relative to the nitrogen concentrations ($0.1\text{--}16\ \mu\text{mol L}^{-1}$) observed in seawater at this site over the time series (Fig. 8c). Coccolithophores have the lowest half-saturation constant for nitrogen, supporting their ability to thrive better than other functional types at low nitrogen concentrations, in line with the findings of many previous studies including Litchman et al. (2006).

Phytoflagellates and coccolithophores have respectively the smallest and largest irradiance half-saturation constants, whereas diatoms and dinoflagellates fall in between with comparable irradiance half-saturation constants (Figure 8d). These results suggest that phytoflagellates experience saturating irradiance for most of the year, due partly to their distinctive small sizes, which provide a high light absorption per unit of pigment (Finkel and Irwin 2000, Finkel 2001, Finkel et al. 2004) and thereby, a competitive advantage over functional types with larger cells under low light conditions. The competitive edge over other functional types at low light levels allows phytoflagellates to maintain a high biomass with relatively little seasonal variability around the year (Fig. 2). The

high irradiance half-saturation constant of coccolithophores indicates that they require high irradiance levels to grow, consistent with the documented increase in coccolithophore biomass during summer months under thermal stratification and increased irradiance (e.g., Holligan and Harbour 1977, Boalch et al. 1978), and the patterns of taxonomic succession at Station L4.

Based on the premise that silicate may not be important for non-diatom phytoplankton, we independently assigned spike-and-slab mixture prior to the silicate half saturation constants to perform variable selection, assuming a prior inclusion probability of 0.25 or equivalently prior odds of 1:3 for including silicate in each functional type's biomass growth model. The posterior inclusion probabilities of silicate for diatoms, dinoflagellates, coccolithophores and phytoflagellates were 0.90, 0.03, 0.01, and 0.21 respectively, and t corresponding Bayes factors (the ratios of posterior odds to prior odds) are 18, 0.09, 0.03, and 0.96, respectively. Interpreted on the Jeffreys' scale of evidence (Jeffreys 1961), these Bayes factors indicate clearly that silicate is redundant for all functional types but diatoms, which is not surprising since diatoms are the only functional type requiring silicate for growth. The posterior median of the silicate half-saturation constant for diatoms, $0.65 \mu\text{mol L}^{-1}$, is slightly larger than the median silicate concentration in sea water over the study period $0.54 \mu\text{mol L}^{-1}$ and far below the 4th quartile $5.33 \mu\text{mol L}^{-1}$, implying that silicate is frequently limiting to diatom growth.

The explicit integration of resource limitation significantly altered the magnitudes of the interspecific interactions affecting each functional type. It broadly reduced the mag-

Table 2. Posterior means of pairwise interaction coefficients among the four functional types at Station L4. Cell (i,j) shows the posterior estimate of the interaction coefficient of the functional type j on functional type i under the reduced model (top) and the full model (bottom). Bold numbers indicate interaction coefficients considered statistically different from zero (at least one standard deviation away from zero).

| Functional type | Diatoms | Dino. | Cocco. | Phyto. |
|------------------|-------------|-------------|--------------|--------------|
| Diatoms | -1 | 0.18 | -0.02 | 0.04 |
| | -1 | 0.08 | -0.15 | 0.12 |
| Dinoflagellates | 0.41 | -1 | 0.13 | -0.43 |
| | 0.31 | -1 | 0.06 | -0.27 |
| Coccolithophores | 0.02 | 0.15 | -1 | -0.51 |
| | -0.06 | 0.08 | -1 | -0.37 |
| Phytoflagellates | 0.13 | 0.05 | -0.03 | -1 |
| | 0.10 | 0.02 | -0.03 | -1 |

nitudes of important interactions by 20 to 50%, except the negative effect of coccolithophores on diatoms and the positive effect of phytoflagellates on diatoms whose magnitudes increased by 7.5 and 2.4 fold respectively, shifting from statistically negligible to important (Table 2).

For all functional types, the decrease in residual variances consecutive to the integration of resource limitation was small (<10%) and much less than we expected. Our results suggest that resource limitation regulates the dynamics of our study community by constraining the realized growth rate and adjusting the strength of inter-specific interactions.

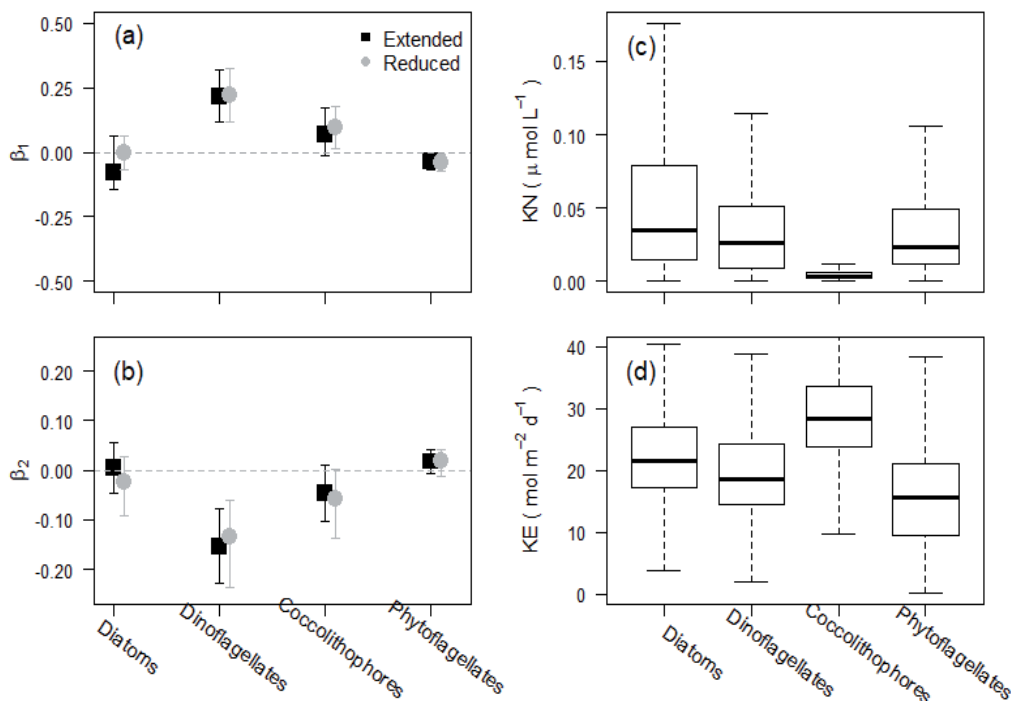


Figure 8. (a) and (b) posterior means (black filled squares) of temperature and salinity effects along with the error bars (posterior mean ± 1 sd) for all functional types based on the full model, with corresponding results based on the reduced model plotted in grey. (c) and (d): Boxplots summarizing the posterior distributions of nitrogen and PAR half-saturation constants.

4. Discussion

In this paper, we developed a Bayesian model and a methodology for quantifying from *in-situ* time-series data, the biotic, abiotic and stochastic drivers of community structure in species-rich assemblages. We parameterized a VAR model of population co-variation integrating demographic stochasticity, density-dependent regulation, pairwise interactions, and abiotic forcing mediated by changing environmental conditions and potentially limiting resources, and worked out explicit formulae for partitioning the temporal variance in each group's dynamics into contributions from biotic factors, abiotic factors and demographic stochasticity.

Although density-dependent regulation is a population-level mechanism, it has far-reaching community-level dynamical implications as a modulator of the among-populations interaction effects. The interaction coefficient $\alpha_{i,j}$ conveys the sign of the effect of population j on the growth of population i , but it does not completely determine the magnitude of the effect experienced by population i which, under the Gompertz model assumed here, is given by $h_i|\alpha_{i,j}|$. This establishes a direct link between the population- and community-level biotic regulation with the density feedback h_i acting as a regulator amplifying the impact of other populations on the growth rate of population i beyond the level assumed by the interaction coefficients $\alpha_{i,j}$ ($j \neq i$) when $h_i > 1$, and *vice-versa*. In our example, h_i lies between 0 and 1, implying that pairwise interaction coefficients generally overestimate the actual effects of one functional type on another (Figure 3). The intrinsic growth rates, r_i , and carrying capacities, k_i , determine, through the ensuing density-dependent feedback, the magnitude of inter-specific interaction effects experienced by each population. In a competitive community, a population with higher growth rates can be at competitive disadvantage if its carrying capacity is low. Conversely, a population with low intrinsic growth rates can enjoy a competitive advantage if its carrying capacity is high enough.

A number of studies, including Mutshinda et al. (2009, 2011), Porzig et al. (2016) and Lany et al. (2017) have used the multivariate Gompertz model to integrate biotic interactions, abiotic forcing and demographic stochasticity in a community context. To our knowledge, this study is the first to incorporate resource co-limitation, and our analysis of the L4 data demonstrates that the estimates of interaction coefficients might change in magnitude and/or in sign with the inclusion of resource limitation (e.g., Brooks and Tasman 2018). This presents a potential challenge to interpreting the modelling results since some of the identified interactions may be artifacts of mutual correlations with one or more unmeasured environmental variables (e.g., Hampton et al. 2006). This sort of caution is an inevitable consequence of statistical modelling since models and data are always incomplete. Additional support for inferences can be gathered by showing results that are robust to changes in model formulation and replication with additional data.

Our example analysis using phytoplankton data from Station L4 illustrated the ability of our modeling approach to derive trait values for key parameters determining growth of

functional types, to partition the temporal variance of community log biomass, and to quantify the magnitude and direction of biotic interactions among functional types. The four functional types examined differ broadly in most of the traits (Figs 3-4). Our results are consistent with observed patterns of taxonomic succession at Station L4 with coccolithophores following blooms of diatoms, as the euphotic zone becomes silicate depleted and starts to stratify (Giraudeau and Bailey 1995). The dinoflagellate biomass tends to increase in low salinity (negative salinity coefficient) and high temperature (positive temperature coefficient) conditions (Figure 4), which are tied to increased stratification and stability of the water column, representing conditions unfavorable to diatoms due to poor nutrient replenishment from the hypolimnion (e.g., Hartman et al. 2014).

Mutshinda et al. (2017) parameterized a Bayesian model to quantify trait values of these four functional types from the data considered here, describing the net growth rate of each functional type as a linear combination of a density-independent and a density-dependent component. This analysis was simpler than the method presented here in that pairwise interactions were not individually identified. Since Mutshinda et al. (2017) incorporated temperature in terms of absolute deviations from each functional type's optimal growth temperature rather than the z -scores considered here, the positive temperature effects found here for dinoflagellates and coccolithophores indicate that the optimum growth temperatures for these two functional types exceed the average temperature over the time series. The parameterization of biotic factors adopted here allows us to evaluate the mutual effects of the functional types under consideration on each other, which may yield valuable insights into the mechanisms of biomass co-variation and taxonomic succession at our study site.

The big difference (approximately a factor of 2) between the realized net growth rate and the maximum net growth rate, from the reduced and the full model respectively, suggests that resource limitation is a major control on biomass growth for all functional types. Ecologists often look at ecological control from a "top-down" or from a "bottom-up" perspective. The top-down view suggests that predators control the numbers of organisms, whereas the bottom-up view holds that the quality and availability of food dictates the numbers of organisms above in the food web. The strong density feedbacks and resource limitation suggest that top-down and bottom-up processes cooperate in regulating the L4 phytoplankton community.

The low variance explained by resource limitation may reflect a mismatch between the sampling resolution and the fast variability and uptake of nutrients. For example, the weekly bulk estimates of inorganic nitrogen concentrations used here may be relatively uninformative at physiological scales since nitrogen is taken up rapidly when available (Laws 2013). Moreover, organic sources of nitrogen not taken into account may be important for some functional types (e.g., phytoflagellates), and their omission may result in artificially low half saturation constants for inorganic nitrogen as found here. The fact that the inclusion of resources considerably altered the magnitude and even the direction of pairwise

interaction coefficients indicates that variation in resources mediates the pairwise interactions among functional types.

Mutshinda et al. (2013) used Bayesian variable selection to identify the environmental drivers of species abundance in a diverse phytoplankton community. Our treatment of silicate limitation herein illustrates the value of Bayesian variable selection in disclosing the dependence of population growth on potentially important resources that enter the model non-linearly, and the results revealed that diatoms are the only functional type subject to silicate limitation as expected.

Our method demonstrates a new approach to studying the dynamics of complex communities from time-series data. We incorporate abiotic factors including environmental conditions and resource availability, biotic interactions parameterized by density-dependent factors and pairwise interactions between functional types, and stochastic variation including demographic variability due to small population sizes. Our Bayesian approach allows general forms for these model components and robust quantifications through MCMC methods. We demonstrate that resource limitation can greatly affect the quantification and interpretation of traits. This new framework for analyzing traits and biotic interactions of functional groups is applicable to community dynamics in a broad range of ecological communities.

Acknowledgements: The authors would like to thank the editor and an anonymous referee for useful comments that considerably improved the article. Phytoplankton biomass and environmental data were provided by the Western Channel Observatory, which was funded as part of the UK Natural Environmental Research Council's National Capability. Data are available at www.westernchannelobservatory.org.uk. This work was supported by the Simons Collaboration on Computational Biogeochemical Modelling of Marine Ecosystems/CBIOMES (Grant ID: 549935, AJI).

References

- Araújo, M.B. and M. Luoto. 2007. The importance of biotic interactions for modelling species distributions under climate change. *Global Ecol. Biogeogr.* 16:743–753.
- Boalch, G.T., D.S. Harbour and E.I. Butler. 1978. Seasonal phytoplankton production in the western English Channel 1964–1974. *J. Mar. Biol. Assoc. UK* 58:943–953.
- Brook, B.W. and C.J.A. Bradshaw. 2006. Strength of evidence for density dependence in abundance time series of 1198 species. *Ecology* 87:1445–1451.
- Brooks, C. and C. Tasman. 2018. Density and biotic interactions modify the combined effects of global and local stressors. *Oikos* 127:1746–1758.
- Clark, N.J., K. Wells and O. Lindberg. 2018. Unravelling changing interspecific interactions across environmental gradients using Markov random fields. *Ecology* 99:1277–1283.
- Cooper, G.J. 2003. *The Science of the Struggle for Existence*. Cambridge University Press, Cambridge, UK.
- Davis, A.J., J.H. Lawton, B. Shorrocks and L.S. Jenkinson. 1998a. Individualistic species responses invalidate simple physiological models of community dynamics under global environmental change. *J. Anim. Ecol.* 67:600–612.
- Davis, A., L.S. Jenkinson, J.H. Lawton, B. Shorrocks and S. Wood. 1998b. Making mistakes when predicting shifts in species range in response to global warming. *Nature* 391:783–786.
- De Baar, H.J.W. 1994. Von Liebig law of the minimum and plankton ecology (1899–1991). *Prog. Oceanogr.* 33:347–386.
- Delean, S., B.W. Brook and C.J.A. Bradshaw. 2013. Ecologically realistic estimates of maximum population growth using informed Bayesian priors. *Methods Ecol. Evol.* 4:34–44.
- Doney, S.C., 2006. Phytoplankton in a warmer world. *Nature* 444: 695–696.
- Field, C.B., M.J. Behrenfeld, J.T. Randerson and P. Falkowski. 1998. Primary production of the biosphere: integrating terrestrial and oceanic components. *Science* 281:237–240.
- Finkel, Z.V., A.J. Irwin and O. Schofield. 2004. Resource limitation alters the 3/4 size scaling of metabolic rates in phytoplankton. *Mar. Ecol. Prog. Ser.* 273:269–279.
- Finkel, Z.V., 2001. Light absorption and size scaling of light-limited metabolism in marine diatoms. *Limnol. Oceanogr.* 46:86–94.
- Finkel, Z.V. and A.J. Irwin. 2000. Modelling size-dependent photosynthesis: light absorption and the allometric rule. *J. Theor. Biol.* 204:361–369.
- Gelman, A., J.B. Carlin, H.S. Stern, D.B. Dunson, A. Vehtari and D.B. Rubin. 2013. *Bayesian Data Analysis* 3rd Ed. Chapman and Hall, London, England.
- George, E.I. and R.E. McCulloch. 1993. Variable selection via Gibbs sampling. *J. Am. Stat. Assoc.* 85:398–409.
- Gilks, W.R., S. Richardson and D.J. Spiegelhalter (eds.), 1996. *Markov Chain Monte Carlo in Practice*. Chapman and Hall, London, UK.
- Gilman, S.E., M.C. Urban, J. Tewksbury, G.W. Gilchrist and R.D. Holt. 2010. A framework for community interactions under climate change. *Trends Ecol. Evol.* 25:325–331.
- Giraudeau, J. and G.W. Bailey. 1995. Spatial dynamics of coccolithophore communities during an upwelling event in the Southern Benguela system. *Cont. Shelf Res.* 15:1825–1852.
- Götzenberger, L., F. de Bello, K.A. Bråthen, J. Davison, A. Dubuis, A. Guisan, et al. 2012. **Ecological assembly rules in plant communities**—approaches, patterns and prospects. *Biol. Rev.* 87:111–127.
- Hampton, S.E., E.E. Holmes, L.P. Scheef, M.D. Scheuerell, S.L. Katz, D.E. Pendleton and E.J. Ward. 2013. Quantifying effects of abiotic and biotic drivers on community dynamics with multivariate autoregressive (MAR) models. *Ecology* 94:2663–2669.
- Hampton, S.E., M.D. Scheuerell and D.E. Schindler. 2006. Coalescence in the Lake Washington story: interaction strengths in a planktonic food web. *Limnol. Oceanogr.* 51:2042–2051.
- Hartman, S.E., M.C. Hartman, D.J. Hydes, D. Smythe-Wright, F. Gohin, F. and P. Lazure. 2014. The role of hydrographic parameters, measured from a ship of opportunity, in bloom formation of *Karenia mikimotoi* in the English Channel. *J. Mar. Syst.* 140: 39–49.
- Heikkinen, R.K., M. Luoto, R. Virkkala, R.G. Pearson and J.-H. Körber. 2007. Biotic interactions improve prediction of boreal bird distributions at macro-scales. *Glob. Ecol. Biogeogr.* 16:754–763.
- Hoerl, A. E. and R. W. Kennard. 1970. Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics* 12(1): 55–67.
- Holligan, P.M. and D.S. Harbour. 1977. The vertical distribution and succession of phytoplankton in the western English Channel in 1975 and 1976. *J. Mar. Biol. Assoc. U.K.* 57:1075–1093.

- Irwin, A.J. and Z.V. Finkel. 2018. Phytoplankton functional types: a trait perspective. In: Kirchman, D.M. and Gasol, J.M. (eds), *Microbial Ecology of the Ocean*. Wiley. Chapter 11, pp. 435–465.
- Ives, A.R., B. Dennis, K.L. Cottingham and S.R. Carpenter. 2003. Estimating community stability and ecological interactions from time-series data. *Ecol. Monogr.* 73:301–330.
- Jeffreys, H., 1961. *The Theory of Probability* (3rd ed.). Oxford University Press, Oxford, UK
- Kissling, W.D., C.F. Dormann, J. Groeneveld, T. Hickler, I. Kühn, G.J. McInerny, J.M. Montoya, C. Römermann, K. Schiffers, F.M. Schurr, A. Singer, J.-C. Svenning, E.Z. Niklaus, and R.B. O'Hara. 2012. Towards novel approaches to modelling biotic interactions in multispecies assemblages at large spatial extents. *J. Biogeogr.* 39:2163–2178.
- Lande, R., S. Engen and B-E. Saether. 2003. *Stochastic Population Dynamics in Ecology and Conservation*. Oxford University Press, Oxford.
- Lany, N.K., P.L. Zarnetske, T.C. Gouhier and B.A. Menge. 2017. Incorporating Context Dependency of Species Interactions in Species Distribution Models. *Integr. Comparat. Biol.* 57:159–167, <https://doi.org/10.1093/icb/ix057>.
- Laws, E.A. 2013. Evaluation of in situ phytoplankton growth rates: A synthesis of data from varied approaches. *Ann. Rev. Mar. Sci.* 5:247–268.
- Le Quéré, C., S.P. Harrison, P.I. Colin, E.T. Buitenhuis, et al. 2005. Ecosystem dynamics based on plankton functional types for global ocean biogeochemistry models. *Glob. Change Biol.* 11: 2016–2040.
- Litchman, E., C.A. Klausmeier, J.R. Miller, O.M. Schofield and P.G. Falkowski. 2006. Multi-nutrient, multi-group model of present and future oceanic phytoplankton communities. *Biogeosciences* 3:585–606.
- Loreau, M. and C. de Mazancourt. 2008. Species synchrony and its drivers: neutral and nonneutral community dynamics in fluctuating environments. *Am. Nat.* 172:E48–E6.
- Lunn, D.J., A. Thomas, N. Best and D. Spiegelhalter. 2000. WinBUGS - A Bayesian modelling framework: Concepts, structure, and extensibility. *Stat. Comput.* 10:325–337.
- McCarthy, M. 2007. *Bayesian Methods in Ecology*. Cambridge University Press, New York.
- Menden-Deuer, S. and E.J. Lessard. 2000. Carbon to volume relationships for dinoflagellates, diatoms, and other protist plankton. *Limnol. Oceanogr.* 45:569–579.
- Mutshinda, C.M., Z.V. Finkel, C.E. Widdicombe and A.J. Irwin. 2017. Phytoplankton traits from long-term oceanographic time-series. *Mar. Ecol. Prog. Ser.* 576:11–25.
- Mutshinda, C.M., Z.V. Finkel, C.E. Widdicombe and A.J. Irwin. 2016. Ecological equivalence of species within phytoplankton functional groups. *Funct. Ecol.* 30:1714–1722.
- Mutshinda, C.M., Z.V. Finkel and A.J. Irwin. 2013. Which environmental factors control phytoplankton populations? A Bayesian variable selection approach. *Ecol. Model.* 269:1–8.
- Mutshinda, C. M. and M. J. Sillanpää 2011. Bayesian shrinkage analysis of QTLs under shape-adaptive shrinkage priors, and accurate re-estimation of genetic effects. *Heredity* 107:405–412.
- Mutshinda, C.M. and M.J. Sillanpää. 2012. A decision rule for quantitative trait locus detection under the extended Bayesian LASSO model. *Genetics* 192:1483–1491.
- Mutshinda, C.M., R.B. O'Hara and I.P. Woiwod. 2011. A multispecies perspective on ecological impacts of climatic forcing. *J. Anim. Ecol.* 80:101–107.
- Mutshinda, C.M. 2010. *Bayesian Analysis of Community dynamics*. PhD thesis, University of Helsinki, Helsinki, Finland.
- Mutshinda, C.M. and M.J. Sillanpää. 2010. Extended Bayesian LASSO for multiple quantitative trait loci mapping and unobserved phenotype prediction. *Genetics* 186:1067–1075.
- Mutshinda, C.M., R.B. O'Hara and I.P. Woiwod. 2009. What drives community dynamics? *Proc. R. Soc. London, Ser. B* 276:2923–2929.
- Mutshinda C.M. 2009. Cutting across discipline boundaries: Statistical prospects in disclosing and handling the workings of natural biodiversity. *Intl. J. Biol.* 1:101–103.
- Nicholson, A.J. 1933. The balance of animal populations. *J. Anim. Ecol.* 2:131–178.
- O'Hara R.B. and M.J. Sillanpää. 2009. A review of Bayesian variable selection methods: what, how and which? *Bayesian Analysis* 4:85–115.
- Ovaskainen, O., J. Hottola and J. Siitonen. 2010. Modeling species co-occurrence by multivariate logistic regression generates new hypotheses on fungal interactions. *Ecology* 91:2514–2521.
- Ovaskainen, O, G. Tikhonov, D. Dunson, V. Grøtan, S. Engen, B.-E. Sæther and N. Abrego. 2017. How are species interactions structured in species-rich communities? A new method for analysing time-series data. *Proc. R. Soc. London, Ser. B* 284:20170768.
- Park, T. and G. Casella. 2008. The Bayesian LASSO. *J. Amer. Stat. Assoc.* 2008:103:681–686.
- Plummer, M., et al. 2003. Jags: A program for analysis of Bayesian graphical models using gibbs sampling. In Proceedings of the 3rd international workshop on distributed statistical computing, volume 124. Vienna, Austria.
- Pollock, L.J., R. Tingley, W.K. Morris, N. Golding, R.B. O'Hara, K.M. Parris, P.A. Vesik and M.A. McCarthy. 2014. Understanding co-occurrence by modelling species simultaneously with a Joint Species Distribution Model (JSDM). *Methods Ecol. Evol.* 5: 397–406.
- Porzig, E.L., N.E. Seavy, J.M. Eadie, D.L. Humple, G.R. Geupel and T. Gardali. 2016. Interspecific interactions, population variation, and environmental forcing in the context of the community. *Ecosphere* 7(6), e01349. [10.1002/ecs2.1349](https://doi.org/10.1002/ecs2.1349).
- Saether, B.-E., J. Tufto, S. Engen, K. Jerstad, O.W. Rostad and J.E. Skåtan. 2000. Population dynamical consequences of climate change for a small temperate songbird. *Science* 287:854–856.
- Stachowicz, J.J. 2001. Mutualism, facilitation, and the structure of ecological communities. *Bioscience* 51:235–246.
- Stan Development Team. 2018. Stan modelling language user's guide and reference manual, version 2.18.0. <http://mc-stan.org/>.
- Tibshirani, R. 1996. Regression shrinkage and selection via LASSO. *J. Roy. Stat. Soc. B* 58:267–288.
- Thomas, A., R.B. O'Hara, U. Ligges and S. Sturtz. 2006. Making BUGS Open. *R News* 6:12–17.
- van der Ploeg, R.R. and M. Kirkham. 1999. On the origin of the theory of mineral nutrition of plants and the law of the minimum. *Soil Sci. Soc. Am. J.* 63:1055–1062.
- Widdicombe, C.E., D. Eloire, D. Harbour, R.P. Harris and P.J. Somerfield. 2010a. Long-term phytoplankton community dynamics in the Western English Channel. *J. Plankton Res.* 32: 643–655.
- Widdicombe, C.E., D. Eloire, D. Harbour, R.P. Harris and P.J. Somerfield. 2010b. Time series of phytoplankton abundance and

composition at station L4 in the English Channel from 1988 to 2009. PANGAEA, <https://doi.org/10.1594/PANGAEA.754335>, In supplement to: Widdicombe, C.E. et al. (2010) Long-term phytoplankton community dynamics in the Western English Channel. *J. Plankton Res.* 32:643–655, <https://doi.org/10.1093/plankt/fbp127>.

Received July 30, 2019
Revised September 24, 2019
Accepted October 22, 2019

Supplementary information

OpenBUGS code for fitting the reduced model and the full model. The file may be downloaded from www.akademaii.com.