



ORIGINAL ARTICLE

Environmental control of the dominant phytoplankton in the Cariaco basin: a hierarchical Bayesian approach

CRISPIN M. MUTSHINDA^{1*}, LUIS TROCCOLI-GHINAGLIA², ZOE V. FINKEL³,
FRANK E. MÜLLER-KARGER⁴ & ANDREW J. IRWIN¹

¹Mathematics & Computer Science, Mount Allison University, Sackville, NB, Canada, ²Escuela de Ciencias Aplicadas del Mar, Universidad de Oriente, Boca de Río, Isla de Margarita, Venezuela, ³Environmental Science Program, Mount Allison University, Sackville, NB Canada, and ⁴Institute for Marine Remote Sensing/IMaRS, College of Marine Science, University of South Florida, St. Petersburg, Florida, USA

Abstract

We develop a hierarchical Bayesian model linking the abundance of individual phytoplankton species with over a decade (1995–2011) of environmental data from the Cariaco Ocean Time Series Program in the Cariaco Basin, Venezuela, to characterize how phytoplankton respond to environmental forcing. Temperature, salinity, irradiance, and macronutrient concentrations account for 39% of the variation in log cell abundance across 67 species. Individual phytoplankton taxa varied widely in their response to these environmental variables. A principal component analysis of the environmental response profiles clearly distinguishes the responses of diatoms and dinoflagellates to environmental forcing. Phytoplankton abundance primarily varied with temperature, pH, and irradiance, with salinity and macronutrient concentrations acting as secondary drivers. In the aggregate, our results demonstrate that environmental changes, whether short-term or a result of climate change, should be expected to have dramatic consequences on the taxonomic composition of phytoplankton communities.

Key words: Bayesian analysis, bottom-up control, community structure, environmental forcing, phytoplankton abundance

Introduction

Phytoplankton are a diverse group of photoautotrophs, comprising tens of thousands of species globally and hundreds of species in local communities. Some of their photosynthetically fixed organic carbon is sequestered in the deep ocean through the biological pump, effectively removing inorganic carbon from the atmosphere for long periods, including tens of thousands to millions of years when buried in sediments (Tedesco & Thunell 2003; Müller-Karger et al. 2005). Knowledge of the environmental factors regulating phytoplankton community composition can offer valuable clues as to how primary production and carbon biogeochemistry may respond to climate change. Irradiance, temperature, salinity, pH and the concentration of macronutrients such as nitrate, phosphate and silicic

acid are potential regulators of phytoplankton biomass, productivity and community structure. Because these factors change spatially and temporally, both naturally and as a consequence of anthropogenic activities, it is of interest to understand the consequences of such changes on phytoplankton community structure.

Phytoplankton biomass is strongly influenced by resource (nutrient and light) availability, but the presence or absence of a particular species or even the diversity of the community is considerably more difficult to predict. Given the diversity and variability of phytoplankton communities, they are a natural choice for examining existing theories of biodiversity. For example, are species abundances in a community changing independently of the environmental conditions, or are species selected according to competitive abilities or other factors? The resolution

*Correspondence: Crispin Mutshinda, Mathematics & Computer Science, Mount Allison University, 67 York Street, Sackville, New Brunswick, Canada E4L 1E6. E-mail: cmutshinda@mta.ca

Published in collaboration with the University of Bergen and the Institute of Marine Research, Norway, and the Marine Biological Laboratory, University of Copenhagen, Denmark

of this question is crucial to understanding how phytoplankton will respond to climate change. If communities are assembled at random and relative species abundances are essentially driven by demographic stochasticity (random drift) according to the neutral model (Hubbell 2001, 2005), then taxonomic composition may be unimportant. If, on the other hand, species are selected according to their traits, we will need detailed knowledge of these traits to predict the impact of climate change on phytoplankton communities and cascading consequences on food web structure and carbon biogeochemistry (Boyd et al. 2010; Finkel et al. 2010).

Statistical analyses of observational data can reveal the complex array of physiological and ecological interactions that shape species abundance patterns in natural communities, and the Cariaco Ocean Time Series Program in the Cariaco Basin (Venezuela) is well-suited to this purpose. The Cariaco Basin is a permanently anoxic zone with sediments recording hundreds to thousands of years of export production (Peterson & Haug 2006; Black et al. 2007). Because of its geochemical and paleoclimatological importance, the CARIACO (Carbon Retention In A Coloured Ocean) Ocean Time Series Programme was started in 1995 to study the primary production and export flux in the Cariaco Basin. Over 16 years of monthly observations on hydrography, water chemistry, macronutrient concentrations, settling sediment flux and phytoplankton species abundance and taxonomic composition make this an ideal site to examine the controls on phytoplankton community structure. Here we develop a hierarchical Bayesian model (Gelman et al. 2003) linking phytoplankton abundance patterns in the Cariaco Basin to potential environmental drivers to determine the degree to which species respond to environmental forcing.

Overview of the Bayesian modelling approach

Bayesian inference is an approach to statistical analysis where all uncertainty is expressed in terms of probability. The major difference between the classical (frequentist) and the Bayesian approach is that the classical approach treats parameters as fixed unknown quantities and seeks the parameter values that maximize the likelihood of the data at hand, whereas in the Bayesian paradigm, all unknown quantities are treated as random variables on which the investigator is required to specify a prior distribution to describe prior knowledge about the plausible values, before observing the data. For a parameter θ , the prior distribution is denoted as $p(\theta)$. After observing some data y , the likelihood, $p(y|\theta)$, of the data is used to update the prior

knowledge about θ encoded in the prior $p(\theta)$ into a posterior distribution $p(\theta|y)$ according to Bayes' rule as $p(\theta|y) = p(\theta)p(y|\theta)/p(y)$, where the marginal distribution of the data, $p(y) = \int_{\Theta} p(y|\theta)P(\theta)d\theta$, is nothing but a normalizing constant that ensures the posterior integrates to 1. The posterior distribution is the target of Bayesian inference as it represents the data-updated state of knowledge about the plausible values of unknown quantities, and allows a direct quantification of the uncertainty in the parameters. Bayesian conclusions arise in the form of probabilistic statements about the unknown quantities, based on their posterior distributions.

In models with many parameters, the computation of the normalizing constant $p(y)$ involves a high-dimensional integral with typically no analytical solution. For example, if the model involves k parameters denoted as $\theta_1, \dots, \theta_k$, then $p(y) = \int_{\Theta_1} \dots \int_{\Theta_k} p(y|\theta_1, \dots, \theta_k)p(\theta_1, \dots, \theta_k)d\theta_1 \dots d\theta_k$. Note that multidimensional integrals arise also in computing the marginal distributions of individual parameters. For example, if one's focus is on θ_1 only, then the rest of parameters are viewed as nuisance parameters, and integrated out. That is, $p(\theta_1|y) = \int_{\Theta_2} \dots \int_{\Theta_k} p(\theta_1, \dots, \theta_k|y)d\theta_2 \dots d\theta_k$. A similar argument applies when making predictions based on a fitted model, in which case all model parameters are integrated out as nuisance quantities. Hence, the predictive distribution of a future observation, \tilde{y} , based on a fitted model with parameter $\theta = (\theta_1, \dots, \theta_k)^T$ is defined by $p(\tilde{y}|y) = \int_{\Theta_1} \dots \int_{\Theta_k} p(\tilde{y}|\theta_1, \dots, \theta_k)p(\theta_1, \dots, \theta_k|y)d\theta_1 \dots d\theta_k$. Bayesian inference, therefore, allows one to fully take into account the uncertainty about the parameter estimates when making predictions with the model.

Markov chain Monte Carlo (MCMC) simulation methods (e.g. Gilks et al. 1996, p 1–19) are the most popular approach to evaluating posterior distributions in high-dimensional Bayesian models. The principle of MCMC is to generate a large sample from a distribution of interest (usually the joint posterior), and base inferences on the simulated sample. For example, the posterior mean and variance of a given quantity can be estimated by the mean and variance of an MCMC-generated posterior sample. Moreover, the bounds of the $100(1 - \alpha)\%$ credible interval (Bayesian analogue of confidence interval) are given by the $\alpha/2$ and $(1 - \alpha/2)$ quantiles of the simulated samples. Credible intervals can be used for hypothesis testing by considering the hypothesis that a parameter θ equals some fixed value θ_0 plausible if the $100(1 - \alpha)\%$ credible interval of θ covers θ_0 , and implausible otherwise.

Simulation-based methods such as MCMC are particularly important when dealing with hierarchical Bayesian (HB) models (e.g. Gelman et al. 2003,

p 117–56; Berlinier 1996). These are Bayesian models where prior distributions are specified in a hierarchical structure. That is, the parameters involved in the likelihood have priors, and the parameters of these priors may also have priors (so-called hyper-priors), whose parameters (hyper-parameters) may in turn have priors, and so on, with the process coming to an end when no more priors are introduced. The HB modelling provides a valuable framework for analysing data with complex structures by defining realistic priors that match the system being studied.

Materials and methods

Study site and environmental data

The Cariaco Basin is a large (~160 km long, 70 km wide and ~1400 m deep) structural depression on the northern continental shelf of Venezuela. The Cariaco Ocean Time Series programme has occupied a time series in the basin (10°30'N, 64°40'W) to collect observations on a monthly basis since November 1995. During the upwelling season (usually December–January to April–May), the Trade Winds intensify coastal upwelling that bring nutrients from deep waters to the surface, increasing primary productivity by a factor of 10 compared to typical productivity values observed in November. Some of the resulting organic material remains ungrazed and sinks. The decomposition of the sinking material leads to permanent anoxia below ~250 m depth (Richards 1975; Müller-Karger et al. 2001, 2004, 2010).

The CARIACO Ocean Time Series surveys have been conducted using the R/V *Hermano Ginés* of the Fundación La Salle de Ciencias Naturales de Venezuela. During each cruise, samples are collected from up to 20 depths between the surface and 1310 m. We used data from the seven depths closest to the surface where environmental data and phytoplankton samples were taken together at 1, 7, 15, 25, 35, 55 and 75 m. Analyses of bottle samples (temperature, salinity, pH, macronutrient concentrations) are conducted according to standard field methods with minor modifications (the operational field manual is available at: http://imars.usf.edu/pubs/CARIACO_Methods_Manual.pdf). Phytoplankton in discrete samples (50 ml for cruises 1–112 and 100 ml for cruises 113–167) at each depth were identified to species level, and counted (cells l⁻¹) using the Utermöhl technique (Hasle 1978). The phytoplankton and environmental data are archived online in the Cariaco Ocean Time Series database (<http://www.imars.usf.edu/CAR/> (accessed 24 November 2012)).

As a quality control step, we assembled the taxonomic identifications made across all cruises,

removed observations that were not clearly identified as phytoplankton, corrected spelling variations, and grouped synonyms using taxonomic information in AlgaeBase (algaebase.org). In total there were 525 phytoplankton species observed, of which the majority were identified to the species level. We analysed the abundance of the 67 most frequently observed phytoplankton species from six different functional groups: diatoms, dinoflagellates, coccolithophores, cyanobacteria, ciliates and silicoflagellates (Table I); 11 of these are genera-level classifications, and one (*Emiliania* + *Gephyrocapsa*) is a mixture of two common coccolithophore genera. When a species was not observed in a given cruise, abundance was set to half the minimum observable threshold ($5 \cdot 10^{-3}$ cells l⁻¹). Phytoplankton abundance data were available for 167 out of the 177 cruises spanning November 1995 to January 2011; they are missing for cruises 22, 30, 36, 57, 58, 95, 96, 99, 100 and 139.

Seven variables were considered in the analysis of possible relationships with phytoplankton abundance: water temperature (°C), salinity (psu), macronutrient concentration (nitrate, phosphate, silicic acid, $\mu\text{mol l}^{-1}$), pH and irradiance. Irradiance was estimated from Sea-viewing Wide-Field-of-View (SeaWiFS) satellite-derived monthly sea-surface PAR ($\text{mol m}^{-2} \text{day}^{-1}$) and was attenuated over depth using the diffuse attenuation coefficient k_{490} (m^{-1}) obtained from Giovanni (<http://disc.sci.gsfc.nasa.gov/giovanni>) from a $0.4^\circ \times 0.4^\circ$ box around the CARIACO Ocean Time Series station from October 1997 to December 2010. The monthly SeaWiFS average was used for months outside this period. Irradiance at depth z was attenuated according to the Beer–Lambert law, $E(z) = E(0)e^{-kz}$, and averaged over the mixed layer (0–25 m at our sampling resolution). Bio-optical properties at the CARIACO station are summarized by Lorenzoni et al. (2011).

Missing environmental data (4–5% of the macronutrients, 14% of the pH and < 1% of salinity and temperature) were estimated using averages of the available data for the corresponding months based on the entire data set. Although there are trends in some of the data (increasing temperature, decreasing macronutrient and salinity: see Taylor et al. 2012), the monthly and annual variation is much larger than the trend over 15 years. A small number of observations (< 10 or 1%) of the environmental data were unrealistic and were removed from the data set.

Model specification

Our goal is to test whether there are significant effects of key environmental variables on phytoplankton abundance and community structure or vice versa.

Table I. Most frequently observed species at the Cariaco study site, with functional group classification, over the time series (November 1995–January 2011). Names without authorities are not in *algaebase.org* or *marinespecies.org*.

Diatom

Bacteriastrum delicatulum Cleve
Bacteriastrum sp. Shadbolt
Cerataulina pelagica (Cleve) Hendey
Chaetoceros affinis Lauder
Chaetoceros anastomosans Grunow
Chaetoceros compressus Lauder
Chaetoceros decipiens Cleve
Chaetoceros didymus Ehrenberg
Chaetoceros lorenzianus Grunow
Chaetoceros sp. Ehrenberg
Chaetoceros sp. 2
Cylindrotheca closterium (Ehrenberg) Reiman & Lewin
Dactyliosolen fragilissimus (Bergon) Hasle apud G.R. Hasle & Syvertsen
Eucampia zodiacus Ehrenberg
Guinardia delicatula (Cleve) Hasle
Guinardia flaccida (Castracane) Peragallo
Guinardia striata (Stolterfoth) Hasle
Haslea wawrikan (Husedt) Simonsen
Helicotheca tamesis Ricard
Hemiaulus hauckii Grunow in Van Heurck
Hemiaulus sinensis Greville
Lauderia annulata Cleve
Leptocylindrus danicus Cleve
Leptocylindrus mediterraneus (H. Peragallo) Hasle
Leptocylindrus minimus Grunow
Navicula sp. Bory de Saint-Vincent
Navicula yarrensis Grunow
Nitzschia fluminensis Grunow
Nitzschia longissima (Brébisson in Kützing) Ralfs in Pritchard
Proboscia alata (Brightwell) Sundström
Pseudo-nitzschia pseudodelicatissima (Hasle) Hasle
Pseudo-nitzschia pungens (Grunow ex P.T. Cleve) Hasle
Pseudo-nitzschia seriata (P.T. Cleve) H. Peragallo in H. & M. Peragallo
Pseudo-nitzschia sp. H. Peragallo in H. & M. Peragallo
Pseudo-nitzschia subfraudulenta (Hasle) Hasle
Rhizosolenia hebetate J. W. Bailey
Rhizosolenia imbricate Brightwell
Rhizosolenia setigera Brightwell
Rhizosolenia styliformis Brightwell

Skeletonema costatum Greville (Cleve)
Thalassionema delicatula
Thalassionema frauenfeldii (Grunow)
Thalassionema nitzschioides (Grunow) Mereschkowsky
Thalassiosira gravida P. T. Cleve
Thalassiosira rotula Meunier
Thalassiosira sp. Cleve
Thalassiosira subtilis (Ostenfeld) Gran

Dinoflagellate

Gonyaulax polygramma Stein
Gymnodinium mitratum Schiller
Gymnodinium sp. Stein
Gyrodinium fusus (Meunier) Akselman
Gyrodinium sp. Kofoid & Swezy
Heterocapsa triquetra (Ehrenberg) Stein
Neoceratium lineatum (Ehrenberg) F. Gomez, D. Moreira & P. Lopez-Garcia
Proocentrum gracile Schütt
Proocentrum micans Ehrenberg

Table I (Continued)

Scrippsiella sp. Balech ex A.R. Loeblich III
Scrippsiella trochoidea (Stein) Balech ex A.R. Loeblich III

Coccolithophore

Calcidiscus leptoporus (Murray & Blackman) Loeblich & Tappan
Calcidiscus sp. Kamptner
Calcipappus caudatus Gaarder & Ramsfjell
Calciosolenia murrayi Gran
Emiliania sp. Hay & Mohler, in Hay, Mohler, Roth, Schmidt & Boudreaux + *Gephyrocapsa* spp. Kamptner

Cyanobacteria

Synechococcus sp. Nägeli
Trichodesmium thiebautii Gomont

Ciliate

Mesodinium rubrum (Lohmann, 1908)

Silicoflagellate

Dictyocha fibula Ehrenberg

We describe the log abundance of each species as a linear function of scaled environmental factors. The estimated coefficients of each environmental variable can be interpreted as the expected change in the corresponding species' log abundance resulting from a change of one standard deviation in the value of the environmental variable when the rest of environmental variables are held at their mean values. While non-linear interactions are possible, a linear model usually provides a good approximation of a non-linear function around a given point if the argument's range is not too wide, by virtue of the first-order Taylor approximation. It is worth noting that on the original scale, our model assumes a non-linear relationship of the form $y = \alpha e^{\beta x}$, $\alpha > 0$, $\beta \in \mathfrak{R}$ between the abundance, y , of a given phytoplankton species and the value, x , of each environmental variable.

We have discarded temporal information by treating the phytoplankton abundances as a set of independent samples, conditionally on the included environmental variables. This is because phytoplankton abundances change over time scales of days to weeks and our samples were collected monthly, suggesting that there will be little temporal autocorrelation in the data. However, the validity of such an assumption is worth checking because temporal correlations may result from latent (unmeasured) environmental variables with slower time scales than phytoplankton. We examined this assumption (see Results) by verifying that the residual errors at different depths, after removing the effects of the involved environmental variables, are serially uncorrelated.

A traditional frequentist approach, such as a multiple linear regression, would impose no assumptions on the distribution of estimated effects, and the correlations among the environmental data would complicate the interpretation of the effects and their

significance. The hierarchical Bayesian approach adopted here allows us to estimate the joint distribution of the environmental effects, which involves their correlations. More specifically, letting $y_{s,d,k}$ denote the natural logarithm of the observed abundance of species s at depth $d = 1, 2, 3, 4, 5, 6, 7$ (corresponding, respectively, to 1, 7, 15, 25, 35, 55 and 75 m) in cruise k , we assume that:

$$y_{s,d,k} \sim N(\mu_{s,d,k}, \sigma_s^2) \quad (1)$$

and

$$\mu_{s,d,k} = \alpha_s + \sum_{j=1}^{\mathcal{J}} \beta_{s,j} x_{k,d}^j, \quad (2)$$

where α_s is a random intercept characterizing the log-abundance of species s when all predictors are held at their average values, and $\beta_{s,j}$ represents the effect of the j th environmental variable on the log abundance of species s . The raw environmental data are standardized to a mean of 0 and a variance of 1 using the data in Table II. If we let $\bar{\alpha}_s$ denote the posterior mean of α_s , then the expected log-abundance of species s with regard to variation in the j th environmental variable when all other variables are held at their mean values is $\bar{\alpha}_s + \beta_{s,j} x^j$. Noting that x^j is the standardized deviation of the observed value of the j th environmental variable from its mean value, it follows that a positive value of $\beta_{s,j}$ implies greater abundance of species s when the j th environmental variable is above its mean value.

Environmental data are often correlated. In a regression setup, strong correlations between predictors often lead to correlated regression coefficients, making it difficult or even impossible to tease apart their individual effects, as different parameter combinations can produce exactly the same fit. This parameter identifiability issue can be addressed by explicitly modelling and estimating the covariance structure of the regression coefficients. Here we assume that the vector $\boldsymbol{\beta}_s = (\beta_{s,1}, \dots, \beta_{s,\mathcal{J}})$ of environmental effects on the log-abundance of species s is a-priori multivariate normally distributed around the zero-vector $\mathbf{0}_{\mathcal{J}} = (0, \dots, 0)$, with a covariance matrix

denoted by Ω . The diagonal elements, $\Omega_{i,i}$, are the variances of the environmental effects, and the off-diagonal elements, $\Omega_{i,j}$ ($i \neq j$), are co-variances between the effects of the i th and j th environmental variables. Consequently, the correlation between the effects of the i th and j th environmental variables on the log-abundance of species s is given by

$$\rho_{i,j} = \Omega_{i,j} / \sqrt{\Omega_{i,i} \Omega_{j,j}} \quad (3)$$

The covariance matrix Ω is assigned a prior (see below), and estimated from the data alongside the other model parameters. Implicit in the definition of Ω is the assumption that the association between environmental effects on species abundance is identical across species, independently of their functional groups. A convenient prior for covariance matrices that guarantees the positive-definiteness of the posterior is the inverse-Wishart distribution (see e.g. Gelman et al. 2003). Here we assumed that $\Omega \sim \text{InvWish}_{\mathcal{J}}(I_{\mathcal{J} \times \mathcal{J}})$ where $I_{\mathcal{J} \times \mathcal{J}}$ denotes the $\mathcal{J} \times \mathcal{J}$ identity matrix, and $\Omega \sim \text{InvWish}_k(R)$, stands for the inverse Wishart distribution with scale matrix R and k degrees of freedom. We set the number of degrees of freedom to its maximum value \mathcal{J} (the rank of Ω) to convey a lack of prior information about Ω . We assigned $N(0, \sigma_\alpha^2)$ priors independently on the species-specific intercepts, α_s , and independent $Ga(a, b)$ priors on the species-specific variances σ_s^2 , where $Ga(a, b)$ denotes the Gamma distribution with mean a/b and variance a/b^2 . Finally, we placed $Ga(1, 1)$ priors on σ_α^2 and on the hyper-parameters a and b .

We used MCMC simulation (Gilks et al. 1996, p 1–19) through the Bayesian freeware OpenBUGS (Thomas et al. 2006) to sample from the joint posterior of the model parameters. We ran 100,000 iterations of three Markov chains, discarding the first 5000 samples from each Markov chain as burn-in, thinning the remainder to each 20th sample to reduce the correlation between consecutive MCMC samples. Although the chains proved to converge within the first 1000 iterations and were mixing very well, we decided to run the chains much longer, i.e. for 100,000 iterations which took roughly 24 h on a PC. With this large number of MCMC iterations, applying the relatively large thinning factor of 20 to reduce the correlation between consecutive MCMC draws still leaves us with a large posterior sample for inference.

We computed the correlations among the environmental variables and compared these with the estimated correlations between the effects $\beta_{s,j}$ to determine if the correlations between environmental effects on the log-abundance of individual species were following the correlation structure in the environmental data or were deviating from it. We performed a principal components analysis on the

Table II. Environmental variables observed in the Cariaco Ocean Time Series and selected for this analysis. All variables are standardized to mean 0 and variance 1 before being used in the model to facilitate comparison of the estimated effects.

Variable	Units	Mean	SD	Range
Irradiance	mol m ⁻² day ⁻¹	9.99	9.93	0.00–29.30
Temperature	°C	24.14	2.34	19.38–30.06
SiOH ₄	μmol l ⁻¹	2.23	3.43	0.00–63.94
PO ₄	μmol l ⁻¹	0.19	0.21	0.00–3.13
NO ₃	μmol l ⁻¹	2.27	3.59	0.00–29.80
Salinity	psu	36.78	0.17	35.84–37.07
pH	–	8.02	0.06	7.86–8.12

estimated effects $\beta_{s,j}$ to determine if the effects exhibited any structure at the species or functional group level using R (R Development Core Team 2011).

Species abundance patterns and community composition result from a complex interplay between demographic stochasticity, density-dependence regulation and environmental stochasticity (Mutshinda et al. 2009, 2011). Environmental stochasticity refers to the variability in species abundance patterns caused by fluctuations in environment conditions with regard to both abiotic factors, like climate and weather, and biotic factors (trophic and extra-trophic interactions). Population time series collected over many years are also prone to observation error. Teasing apart these different sources of variation is a difficult task that requires highly structured models with an explicit account of the diverse sources of variability. In our model, the variance parameters σ_s^2 describe the unexplained variability at the species level and the hyper-parameters a and b describe the unexplained variability at the community level. These hyper-parameters are estimated by borrowing strength across species through the data used to estimate the idiosyncratic variance parameters σ_s^2 for each species, thereby mitigating the estimation problem posed by the smaller sample sizes for some of the study species.

A direct evaluation of the overall amount of variation explained by the environmental predictors studied can be estimated by comparing the unexplained variances between the full model including the environmental predictors under consideration and the null model without any environmental covariates.

Results

Phytoplankton species richness varied from 4 to 100 species per cruise, with a mean of 40 over the CARIACO time series (167 cruises from November 1995 to March 2011). Over a third of the species (187 of 525) were only observed during one or two cruises in the entire study. To ensure sufficient data for each species, we restricted our analysis to the 67 species observed in 50 or more samples (Table I). This subsample includes an average of 85% of the cells counted on each cruise and thus describes most of the variation in phytoplankton cell abundance, but not most of the species richness. These species were found in an average of 139 samples each, but because there are 167 (cruises) \times 7 (depths) = 1169 samples, the abundance record for any one species is mostly comprised of zeros, due either to true absences or abundance below the detection threshold. We used an estimated abundance of half the detection limit (0.5 cells per 100 ml) for species not observed in a

sample. The mean and standard deviation of each environmental variable are summarized in Table II.

We identified significant relationships between environmental variables and abundance for 63 of the 67 individual taxa examined from 15 years of the Cariaco Ocean Time Series. The environmental linkages varied widely across species and environmental variables with many effects being significantly different from zero as indicated by the 95% credible intervals (Figure 2). While some environmental variables were found to have no effect on log abundance for some species, more than half of the environmental variables showed a significant relationship. There are broad trends in the sign and magnitude of environmental effects on log abundance across species and functional groups (Figures 1a,b). Temperature, irradiance, and pH coefficients are most likely to be significantly different from 0, and their effects have the largest magnitude. The sign of these effects was broadly consistent across species, but there was substantial variability across taxa, within and across functional groups. There were only four species with no significant effect for any environmental variable on log abundance (*Nitzschia fluminensis* and *Thalassionema fraueneldii*, both diatoms; *Gonyaulax polygramma*, a dinoflagellate; and *Dictyocha fibula*, a silicoflagellate). Our results indicate that changes in environmental conditions will change the abundance of most species and that there will be complex taxonomic structure changes due to the effects of changing environmental conditions.

The interpretation of the effect, $\beta_{s,j}$, of the j th environmental variable on the log-abundance of species s , is straightforward: it is the expected increase in log abundance of phytoplankton for a unit change in the standardized value of the environmental conditions when all other variables are held at their mean values. The reported effects are computed from the marginal posterior $p(\beta_{s,j}|Data)$, by integrating the joint posterior $p(\beta_s|Data)$ over all other parameters. The posterior estimate of the covariance matrix Ω reveals associations among the environmental effects which can be distinct from the correlations among the observed environmental variables themselves (Table III).

Temperature, pH and irradiance were strongly associated with phytoplankton abundance, with the most non-zero effects. The coefficients of temperature were broadly negative, particularly for the diatoms (Figure 1a), indicating lower cell concentrations at warmer temperatures. There are a few exceptions to this general trend. About one-third of the diatom species, mostly those with relatively low abundance, have no significant response to temperature, indicating a division within the diatoms in their traits for these variables. Dinoflagellates were mixed

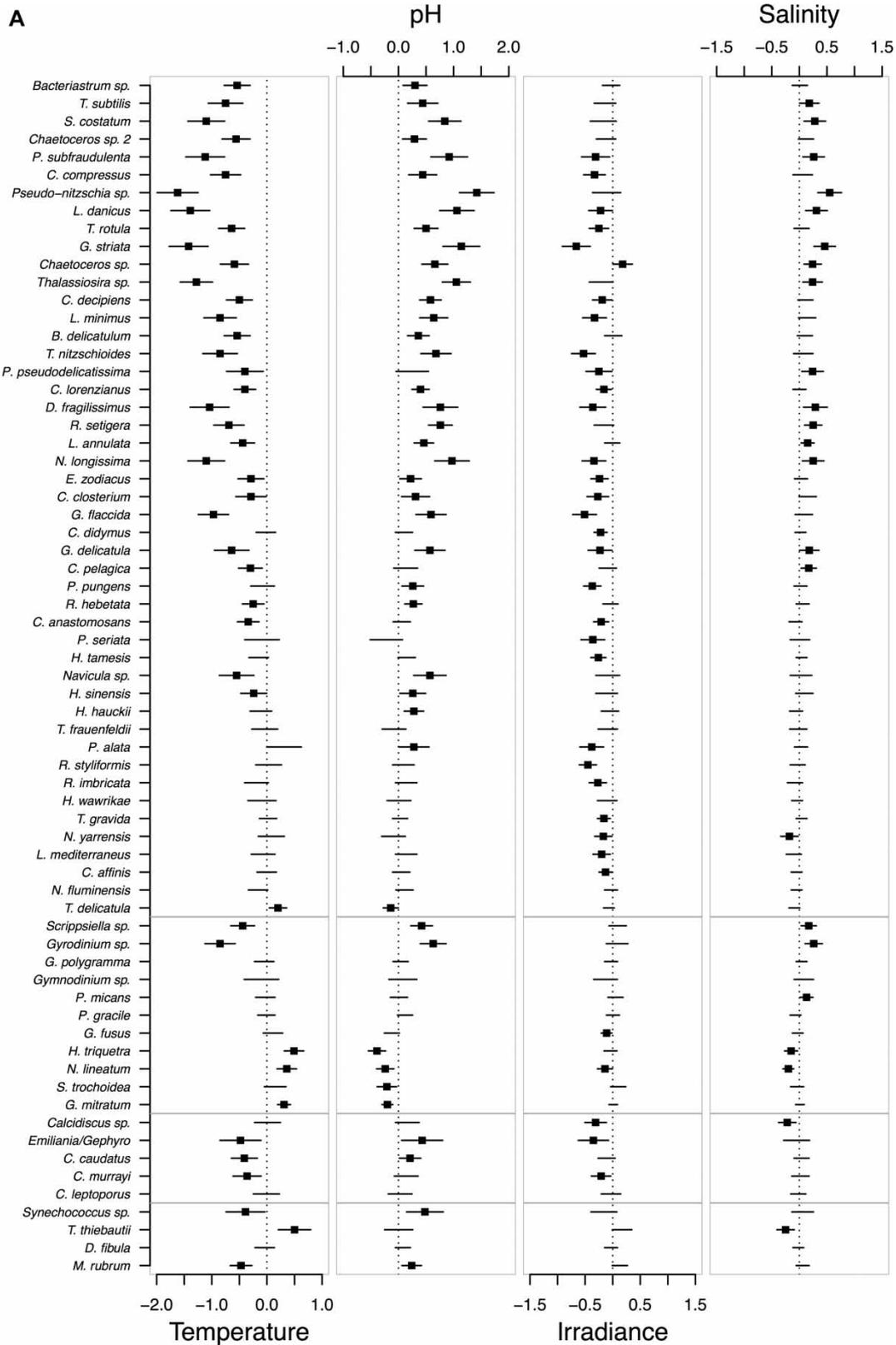


Figure 1. Posterior means (filled square if significantly different from zero, otherwise not shown) and 95% credible interval (horizontal line) of the effects for scaled (A) temperature, pH, irradiance, salinity, and (B) nitrate, phosphate and silicic acid concentration. A dashed vertical line indicates 0 effect. Horizontal lines divide diatoms, dinoflagellates, coccolithophores and other functional groups (Table I). Effect scales all have the same width (3 units). Species are sorted according to median abundance within each functional group. The last panel in (B) shows the median \log_{10} abundance (square), maximum \log_{10} abundance (diamond), and median ± 1 SD (horizontal line) for the samples each species was found in and not counting zeros.

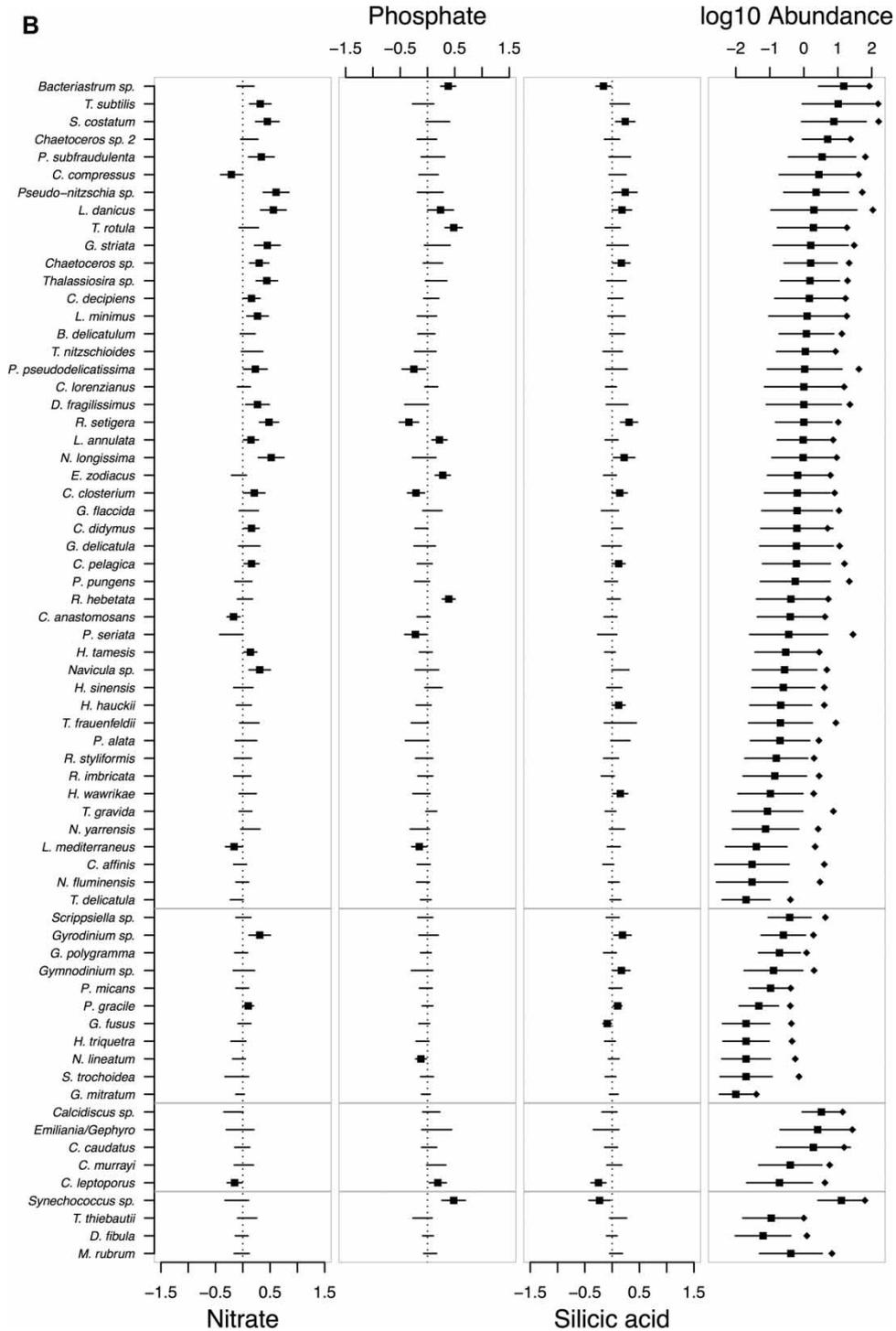


Figure 1. Continued

in their responses to temperature, indicating a range of strategies.

The coefficients associated with pH were broadly positive, with most exceptions among the dinoflagellates. On the other hand, the coefficients associated with irradiance were generally negative. However, these associations of pH and irradiance with phytoplankton abundance should be interpreted with

caution since phytoplankton abundance can affect irradiance and pH (see Discussion). In fact, as this is a statistical study of observational data, we did not make causal conclusions from these data alone. This limitation does not affect our ability to characterize how the environmental variables relate to the log abundance of individual phytoplankton species.

Table III. Correlations, r , between environmental data (top row) and correlations, ρ , between the posterior distributions of model effects estimated with the Bayesian model (bottom row, Equation 3); bold numbers indicate coefficients whose 95% credible intervals exclude zero.

	pH	Irradiance	Salinity	NO ₃ ⁻	PO ₄ ⁻	SiO ₄
Temperature	0.85	0.74	-0.51	-0.74	-0.70	-0.45
	-0.90	0.50	-0.59	-0.65	-0.17	-0.32
pH		0.73	-0.28	-0.86	-0.79	-0.60
		-0.49	0.59	0.68	0.20	0.33
Irradiance			-0.31	-0.67	-0.64	-0.47
			-0.24	-0.30	-0.01	-0.10
Salinity				0.13	0.09	-0.11
				0.51	0.09	0.27
NO ₃ ⁻					0.86	0.68
					0.00	0.27
PO ₄ ⁻						0.64
						-0.16

The point measurements of macronutrient concentration were generally less important as explanatory variables of phytoplankton abundance (Figure 1b) compared to temperature, pH and irradiance. Nitrate concentration was broadly positively associated with the abundance of diatoms, while its effects for other functional groups were mostly non-significant. Nitrate is not always limiting at the study site, but upwelling (estimated using temperature and salinity as proxy) stimulated community productivity due to nutrient additions (Müller-Karger et al. 2001, 2004). The effect of phosphate concentration was frequently non-significant, even among the diatoms. The coefficients of silicic acid concentration were broadly non-significant for the dinoflagellate, coccolithophorid, ciliate and silicoflagellate species, but positive for about half the diatoms under study. There are a few exceptions to the above trends: positive nitrate effects for *Gyrodinium* sp. and *Prorocentrum gracile* (dinoflagellates), and negative phosphate effects for *Neoceratium lineatum* (dinoflagellates), positive phosphate effects for *Calcidiscus leptoporus* (coccolithophorids), and a negative nitrate effect for *Calcidiscus leptoporus* and *Calcidiscus* sp. (coccolithophorids). The two cyanobacteria *Synechococcus* sp. and *Trichodesmium thiebautii* showed contrasting effects for most environmental variables, although not all differences were significant. *Trichodesmium* showed a significantly positive effect of temperature and a nearly significant effect of irradiance, while *Synechococcus* sp. had a large positive effect for phosphate concentration.

Salinity effects were significant for about half of the taxa. There was a small seasonal oscillation in salinity, varying from 36.6 to 37.0 psu for January–August and decreasing to 36 psu for the last four months of the year. The high abundance species and most of the diatoms had positive salinity effects (Figure 1a), consistent with their dominance during the productive months of January–August. The less-

abundant diatoms and dinoflagellates had negative salinity effects, indicating that they are more abundant towards the end of the year in lower productivity environments.

There was an overall positive relationship between the magnitude of the effects of environment and the median abundance of taxa (Figures 1a,b) for the most significant variables (temperature, pH, salinity, and nitrate concentration). Some of the species with the largest abundance (*Pseudo-nitzschia* sp. and *Synechococcus* sp.) had the largest magnitude effects, but this is not universal: the most abundant species in each functional group had smaller effects than intermediate abundance species, and the least abundant species still frequently have significant effects.

Scatterplots of pairs of environmental effects (Figure 2) showed taxonomic structure with diatoms and dinoflagellates (the two most diverse functional groups) appearing in largely non-overlapping clusters. We performed a principal component analysis of the environmental response profiles. The first two principal components accounted for 88% of variation in the effects of the environmental variables across species. The biplot of the first two principal components (Figure 3) illustrated the structure in the environmental effects within and among functional groups clearly distinguishing the responses of diatoms and dinoflagellates to environmental forcing. About half of the dinoflagellate species are outside the large cloud of diatoms and the other half are on the edge of the diatom clusters. Other taxonomic groups were represented by very few species and could not be separated out from the diatoms and dinoflagellates.

The posterior correlation matrix between environmental effects on species abundance estimated from Equation (3) and the correlations between the observed environmental variables (Table III) illustrates whether effects mirror the environment or contrast with them and reveal ecophysiological

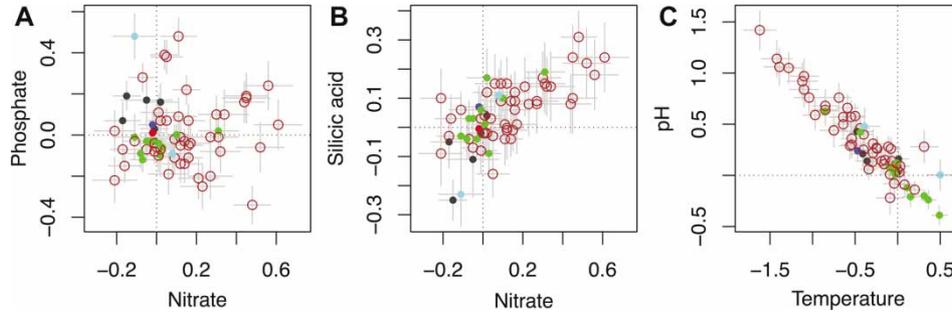


Figure 2. Scatterplots of estimated effects for select pairs of variables. (A) Nitrate and phosphate concentration, (B) nitrate and silicic acid concentration, (C) temperature and pH. Colours indicate the functional group of each species: diatoms (red), dinoflagellates (green), coccolithophores (grey), cyanobacteria (cyan), silicoflagellate (red), ciliate (blue). Diatoms are open circles, all other species are filled circles.

structure in the effects of the different environmental factors. For example, the correlation -0.51 between temperature and salinity in the environment is nearly identical to the correlation -0.59 between the effects of the same two variables, but the correlation 0.86 between phosphate and nitrate in the environment strongly contrasts with the essentially zero correlation in their effects. When the correlations are similar between the environmental variables and their effects, as is the case for temperature and salinity, we interpret this as meaning that the selective effect of each variable is fortuitously aligned with the environmental covariation, or that one variable has a relatively weak effect relative to the other, and that the estimated effect is due to the environmental correlation. On the other hand, the strongly contrasting correlations between nitrate and phosphate suggests a tradeoff in competitive ability for nitrate and phosphate leading to niche differentiation.

The species-specific variance parameters σ_s^2 are drawn from a gamma distribution with parameters a and b that are identical for all species. The expected

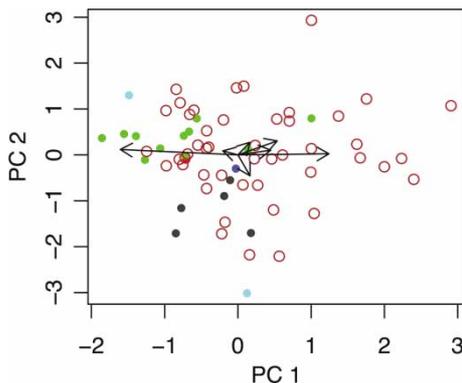


Figure 3. Principal component analysis of effects for all seven environmental variables. Colours and symbols indicate functional groups as in Figure 2. The arrows are projections of vectors in the direction of each environmental axis; clockwise from the top ('noon'): silicic acid, nitrate, salinity, pH, phosphate, temperature and irradiance.

value of σ_s^2 is the mean a/b of its $Ga(a,b)$ distribution. It follows that a/b provides a community-level measure of unexplained variability. The posterior means of the parameter vector (a, b) for the full and the null model were $(2.7, 0.4)$ and $(3.1, 0.28)$, respectively, implying that the included environmental variables account for 39% of the unexplained variance in the null model as $1 - (2.7/0.4)/(3.1/0.28) = 0.39$.

We examined the assumption of no serial correlation between the monthly species abundances conditionally on the included environmental variables, by computing the conditional residual errors $e_{s,k,d} = y_{s,k,d} - \mu_{s,k,d}$ for a random sample of 10 species at different depths, and checking their correlograms (autocorrelation plots) to see if they do not exhibit temporal dependence. The residuals were more or less serially uncorrelated as illustrated by the representative correlograms of four species namely, *Bacterium sp.*, *Chaetoceros compressus*, *Synechococcus sp.* and *Trichodesmium thiebautii* shown in Figure 4. The small autocorrelation observed in the residual correlograms of some species (e.g. *Synechococcus sp.*) indicates that there may be some memory, but the serial correlations are weak and unlikely to affect our results.

It may be interesting to ask how changing the value of an environmental variable would alter a species' growth rate to determine its future abundances. This requires more detailed data recorded at time intervals comparable to the generation time of the study species, which is much shorter than a month for phytoplankton. Our data were recorded monthly, at different depths and involve many missing observations. Moreover, the weak autocorrelation in the conditional residuals (Figure 4) implies that the abundance of a particular species at a specific depth is roughly independent of its abundance in the previous month after controlling for the environmental conditions. Similarly, the relatively long elapsed time between consecutive

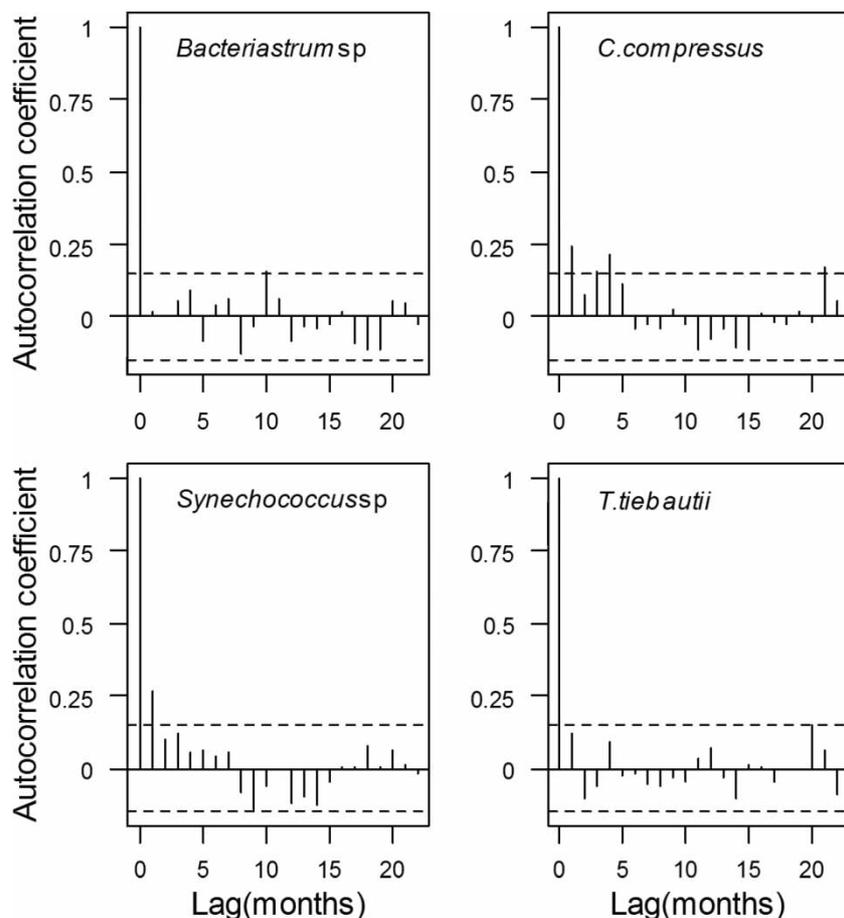


Figure 4. Correlograms of the model residuals for four selected species (*Bacteriastrum* sp., *Chaetoceros compressus*, *Synechococcus* sp. and *Trichodesmium thiebautii*) at 7 m depth. The vertical bars represent the estimated autocorrelation coefficients for observations that are k months (the lag) apart. The dotted horizontal lines are the 95% confidence bounds for a correlation of zero. These correlograms suggest that the residuals can reasonably be assumed to be serially uncorrelated, in line with our model assumption.

observations means that the data are not sufficient to describe the influence of intraspecific and interspecific interactions in the previous time step on the current abundance of a given species.

Discussion

Phytoplankton communities are diverse and exhibit changes in structure at a wide range of temporal and spatial scales (Martin 2003; Durham & Stocker 2011). Our results identified numerous differences in optimal environmental conditions across species, highlighting the role of habitat selection, niche differentiation and tradeoffs in the composition of a phytoplankton community. The environmental variables in our model can be categorized as either conditions (temperature, salinity) or resources (light, macronutrients, and DIC through the pH proxy). Changes in conditions could be expected to change community composition because individual species thrive at particular temperature or salinity based on ecophysiological differences (Eppley 1972; Brand

1984). Alternatively, changes in phytoplankton abundance may cause changes in environmental variables, particularly pH and irradiance. Changes in temperature and salinity may be related to advection of new water masses and new communities and prompt species succession which may not necessarily be due to environmental forcing. Regardless of the nature of the effect, our statistical results indicate a relationship between environmental variables and the abundance of individual species.

Nutrient availability can be assessed through the combination of the standing stock concentration of nutrient pools (nitrate, phosphate and silicic acid concentration) and the flux of those nutrients into the euphotic zone. Significant effects for nutrients can be interpreted as evidence species are resource limited in the case of positive effects, or inhibited in the case of negative effects, or they could be a signal of relative competitive ability for the nutrient across species. We do not have direct observations of nutrient flux at the CARIACO station, but temperature changes are a good proxy for nutrient supply

rate as the cold upwelled water has much higher nutrient concentrations than surface water. To interpret the complex signal resulting from both nutrient concentrations and temperature, it helps to consider two extreme situations: (1) upwelling of cold nutrient-rich waters and (2) a stratified warm layer depleted in nutrients. These two situations are expected to select different species of phytoplankton: elevated nutrient supply rates as indicated by decreasing temperature will support rapidly growing species with high nutrient demands while a stratified layer will select species with the best competitive ability for the limiting resources (Behrenfeld et al. 2006; Dutkiewicz et al. 2009). In practice, a particular community may illustrate elements of both situations.

Most species with significant temperature effects have increased abundance at lower temperatures, consistent with changes in turbulent mixing and the consequent changes in resource supply rate and short-term variability (Behrenfeld et al. 2006). Most of these opportunistic species are seasonally abundant diatoms, but there are seven species from the other functional groups with similar temperature effects. At the other extreme, positive temperature effects are found in some dinoflagellates and *Trichodesmium thiebautii* which are known to thrive in stratified waters. An analysis of environmental response functions for phytoplankton in the North Atlantic also showed that diatom species were more likely to be found in cooler waters compared to dinoflagellates (Irwin et al. 2012). At the CARIACO station, temperature has a strong seasonal cycle driven by changes in the winds with the lowest temperatures in March approximately 4°C cooler than the warmest temperatures in November (Müller-Karger et al. 2010), so the interpretation of temperature effects as due to seasonal succession driven by hydrographic changes and changes in resource availability seems reasonable. Macronutrient concentrations have a significant effect for about half the species studied, with the strongest effects for nitrate among the diatoms, but the effects are secondary to those of temperature. For most species, the nitrate effect is consistent with temperature effects (higher nutrient concentrations are associated with increased abundance), but there are a few species with negative nitrate concentration effects and negative temperature effects, or effects significant for only one of these variables, indicating there is distinct information regarding the availability of nitrate provided by each variable. As phosphate effects are generally smaller and much less likely to be significant, we conclude that few species are phosphate-limited or that the detection limit of phosphate is too high to resolve these effects; some

notable exceptions that have positive phosphate effects and non-significant nitrate effects are the diatoms *Bacteriastrum* sp., *Thalassiosira rotula*, *Eucampia zodiacus*, *Rhizosolenia hebetata* and the cyanobacterium *Synechococcus* sp.

In the environment, all three macronutrients are fairly strongly correlated, but the signal in the effects is more complex. Nitrate and phosphate effects are uncorrelated: a few species have effects with the same sign (3 species), a few have opposite sign (4 species), but most species with a significant effect for one of these macronutrients do not have a significant effect for the other (26 species). These results indicate there is a tradeoff between nitrate and phosphate affinity relative to demand, or that phytoplankton species are responding selectively to only one of nitrate and phosphate, respectively. Whatever the explanation, these correlations indicate that many species are limited by nitrate and fairly few are limited by phosphate. Silicic acid effects are positively correlated with nitrate effects among the diatoms and essentially uncorrelated for other taxa (Figure 3), although silicic acid effects are generally smaller and less likely to be significant than nitrate effects. This matches the expectation that diatoms, which require a specific amount of silicic acid to transit through their cell division cycle (Brzezinski 1992), and are more abundant at higher silicic acid concentrations, can achieve a tradeoff for nitrate relative to phosphate but not for nitrate relative to silicic acid. Species which do not require silicic acid (the dinoflagellates and cyanobacteria) are much less likely to have a significant silicic acid effect and the effects are uncorrelated with nitrate and phosphate concentration effects (Figure 3).

Irradiance is generally not limiting and may be slightly inhibiting the abundance of many species. The irradiance in the upper mixed layer (0–30 m) averages 18.6 mol photons m⁻² day⁻¹, corresponding to an average peak irradiance during the day of 215 μmol photons m⁻² s⁻¹. This indicates phytoplankton cells in the mixed layer can expect to see wide variations in that light during a single generation, including inhibitory irradiances (Kirk 2011, p 330–87). These irradiances are not likely to be subsaturating for growth for most of the species, and the fluctuations in the mixed layer are expected to impose significant photoacclimation and photoinhibition costs (Six et al. 2007; van de Poll et al. 2007; Alderkamp et al. 2010; Key et al. 2010). Thus it is not surprising that irradiance at this site has a negative effect for most species, is significant for half the species, and all but one of the significant effects are negative. Notable exceptions are *Chaetoceros* sp. (diatom), *Mesodinium rubrum* (ciliate) and *Trichodesmium thiebautii* sp. (cyanobacterium), which

have significant positive effects. *Trichodesmium thiebautii* is known to have a high energy demand because of the high cost of N_2 fixation (Andresen et al. 2010), although there is generally high irradiance at this tropical site so this physiological factor may be unimportant at the ecological scale. Increased phytoplankton abundance will lead to a decrease in irradiance with depth, so the negative effects of irradiance on abundance could be attributed to phytoplankton changing the light field. Note, however, that while temperature effects are generally largest for the most abundant species, there is little correlation between the magnitude of irradiance effects and individual species abundance (Figure 2), suggesting that the irradiance effects may be primarily physiological.

The simplest interpretation of pH effects is that increases in phytoplankton abundance are drawing down dissolved inorganic carbon (DIC) and increasing pH; the observed pH effects are a consequence of abundance, not a cause of them (Litt et al. 2010). Chemically, increased pH indicates lower pCO_2 , lower concentrations of DIC, HCO_3^- , and increased concentration of CO_3^{2-} , so the positive pH effects observed are opposite to what would be expected from a CO_2 fertilization perspective. The pH effect on abundance is not simply a result of its positive correlation with temperature since the correlation between pH and temperature effects is strongly negative across all taxa (Figure 2, Table III). A challenge to this interpretation arises for species with positive temperature effects and negative pH effects, which are more abundant at higher temperatures and lower pH (*Heterocapsa triquetra*, *Neoceratium lineatum*, *Gymnodinium mitratum* among the dinoflagellates and *Thalassionema delicatula* in the diatoms). These are among the least abundant of our 67 most frequently observed species; a positive effect of pH might be difficult to observe for these species which are a minor component of their communities and their pH effects may be a signal that these species are most abundant when other species are least abundant because of differences in their niches. Coccolithophores release DIC as a consequence of forming calcium carbonate liths, but the temperature–pH relationship does not seem to be weakened or reversed for these species, indicating a minor role for calcification on pH at this site, or lower pH waters may inhibit the growth of coccolithophores, leaving an apparent positive effect of pH.

The responses of the two cyanobacteria are unusual in that they are frequently contrasting both with each other and with the majority of the phytoplankton species (Figure 4, cyan symbols). This is expected because they have very different

niches: *Synechococcus* sp. is small, with a relatively high phosphate requirement while *Trichodesmium thiebautii* fixes N_2 gas as its source of nitrogen, and is more likely to be found in stratified, non-upwelling warm waters with little fixed nitrogen where it can grow with less competition. These traits are consistent with their estimated environmental effects. *Synechococcus* sp. has among the largest phosphate effect, and *Trichodesmium* sp. has one of the highest temperature and irradiance effects (although the irradiance effect is not quite significant). C : P and N : P ratios from laboratory studies are low for *Synechococcus* sp. (52 : 1 and 8.7 : 1, respectively) and high for *Trichodesmium* sp. (161 : 1 and 31 : 1, respectively) consistent with high P demand by *Synechococcus* sp. and high irradiance growth and N_2 fixation by *Trichodesmium* sp. (Quigg et al. 2010; see also the online supplementary material available on this article's webpage).

Phytoplankton communities exhibit a great deal of variation due to a combination of regulation of growth by bottom-up factors such as resource availability (Grover 1997), variation in mortality due to grazing, viral and parasitoid attack (Alpine & Cloern 1992; Mühling et al. 2005), and demographic drift among species (neutral variation). Bottom-up factors can be complex, depending on nutrient supply rates and ratios (Tilman 1982), light availability and photoinhibition (Six et al. 2007; Alderkamp et al. 2010), species-specific competitive ability for resources (Litchman 2007) and other factors influencing physiological responses such as temperature (Eppley 1972). Variation in mortality is much less frequently measured than the factors promoting growth due to the technical challenges in making measurements. None of our environmental variables are good proxies of grazing rates or whether the abundance of a particular population is determined by grazer control or resource supply at a particular time.

We have identified significant effects for one or more of our environmental variables in 63 of 67 species examined, demonstrating the important impact environmental variables have on the composition of the major species in the phytoplankton community observed at the CARIACO station. We restricted our model to bottom-up factors which accounted for 39% of the variation in log abundance. Much less of the total variance is explained by our model compared to models of total chlorophyll (Irwin & Finkel 2008) because of the tremendous challenge of predicting log abundance for 67 individual species. We have not ruled out the possibility of purely random fluctuations in the numerous infrequently observed species in this community; in fact, the trend of decreasing magni-

tudes of environmental effects with median abundance for each species (Figures 1a,b) suggests that many of the less-abundant species, most of which were omitted from our analysis, may be effectively random samples governed by stochastic variations. Despite these caveats, our results strongly indicate that environmental changes, whether short-term or a result of climate change, should be expected to have dramatic consequences on the relative abundance of the dominant phytoplankton species in a community.

Acknowledgements

CMM was supported by an Ace-Net post-doctoral fellowship and the Marjorie Young Bell faculty fund. AJI and ZVF were supported by NSERC Canada. The CARIACO Ocean Time Series Programme has been made possible through the support of the Venezuelan Fondo Nacional de Ciencia, Tecnología e Innovación (FONACIT) and the US National Science Foundation (NSF, grants OCE-0326268, OCE-0752139, and OCE-0963028 to FMK). We are grateful to the scientific, technical and administrative staff of the Fundación La Salle de Ciencias Naturales de Venezuela, in particular Ramón Varela and Yrene Astor, for their able support at sea and in leading the CARIACO programme. We are particularly indebted to the crew and captain of the *R/V Hno Gines* for their expert support at sea.

References

Alderkamp AC, de Baar HJW, Visser RJW, Arrigo KR. 2010. Can photoinhibition control phytoplankton abundance in deeply mixed water columns of the Southern Ocean? *Limnology and Oceanography* 55:1248–64.

Alpine AE, Cloern JE. 1992. Trophic interactions and direct physical effects control phytoplankton biomass and production in an estuary. *Limnology and Oceanography* 37:946–55.

Andresen E, Lohscheider J, Setlikova E, Adamska I, Simek M, Küpper H. 2010. Acclimation of *Trichodesmium erythraeum* ISM101 to high and low irradiance analysed on the physiological, biophysical and biochemical level. *New Phytologist* 185:173–88.

Behrenfeld MJ, O'Malley RT, Siegel DA, McClain CR, Sarmiento JL, Feldman GC, et al. 2006. Climate-driven trends in contemporary ocean productivity. *Nature* 444:752–55.

Berlinier LM. 1996. Hierarchical Bayesian time series models. In: Hanson K, Silver R, editors. (Maximum Entropy and Bayesian Methods). Dordrecht: Kluwer, p 15–22.

Black D, Abahazi M, Thunell R, Kaplan A, Tappa E, Peterson L. 2007. An eight-century tropical Atlantic SST record from the Cariaco Basin: Baseline variability, 20th century warming, and Atlantic hurricane frequency. *Paleoceanography* 22:PA4204. 14 pages.

Boyd PW, Strzepek R, Fu F, Hutchins DA. 2010. Environmental control of open-ocean phytoplankton groups: Now and in the future. *Limnology and Oceanography* 55:1353–76.

Brand LE. 1984. The salinity tolerance of forty-six marine phytoplankton isolates. *Estuarine, Coastal and Shelf Science* 18:543–56.

Brzezinski MA. 1992. Cell-cycle effects on the kinetics of silicic acid uptake and resource competition among diatoms. *Journal of Plankton Research* 14:1511–39.

Durham WM, Stocker R. 2012. Thin phytoplankton layers: Characteristics, mechanisms, and consequences. *Annual Review of Marine Science* 4:177–207.

Dutkiewicz S, Follows MJ, Bragg JG. 2009. Modeling the coupling of ocean ecology and biogeochemistry. *Global Biogeochemical Cycles* 23:GB4017: 1–15.

Eppley RW. 1972. Temperature and phytoplankton growth in the sea. *Fishery Bulletin* 70:1063–85.

Finkel ZV, Beardall J, Flynn KJ, Quigg A, Rees TAV, Raven JA. 2010. Phytoplankton in a changing world: Cell size and elemental stoichiometry. *Journal of Plankton Research* 32:119–37.

Gelman A, Carlin JB, Stern HS, Rubin DB. 2003. *Bayesian Data Analysis* 2nd Edn. London: Chapman & Hall, p 117–56.

Gilks WR, Richardson S, Spiegelhalter DJ. 1996. *Markov Chain Monte Carlo in Practice*. London: Chapman and Hall, p 1–19.

Grover, JP. 1997. *Resource competition*. Chapman and Hall, London, UK. 342 pages.

Hasle S. 1978. The inverted microscope method. In Sournia, editor. *Phytoplankton Manual*. SCOR-Unesco. 337 pages.

Hubbell, SP. 2001. *The Unified Neutral Theory of Biodiversity and Biogeography*, Princeton University Press. 448 pages.

Hubbell, SP. 2005. Neutral theory in ecology and the hypothesis of functional equivalence. *Functional Ecology* 19:166–172.

Irwin AJ, Finkel ZV. 2008. Mining a sea of data: Determining controls of ocean chlorophyll. *PLOS One* 3:e3836. 6 pages.

Irwin AJ, Nelles AM, Finkel ZV. 2012. Phytoplankton niches estimated from field data. *Limnology and Oceanography* 57:787–97.

Key T, McCarthy A, Campbell DA, Six C, Roy S, Finkel ZV. 2010. Cell size tradeoffs govern light exploitation strategies in marine phytoplankton. *Environmental Microbiology* 12: 95–104.

Kirk JTO. 2011. *Light and Photosynthesis in Aquatic Ecosystems* 3rd Edn. Cambridge: Cambridge University Press? p 330–87.

Litchman E. 2007. Resource competition and the ecological success of phytoplankton. In Falkowski PG, Knoll A, editors. *The Evolution of Eukaryotic Phytoplankton*. London and San Diego: Elsevier Academic Press, p 351–75.

Litt EJ, Hardman-Mountford NJ, Blackford JC, Mitchelson-Jacob G, Goodman A, Moore GF, et al. 2010. Biological control of pCO₂ at station L4 in the Western English Channel over 3 years. *Journal of Plankton Research* 32:621–29.

Lorenzoni L, Hu C, Varela R, Arias G, Guzmán L, Müller-Karger F. 2011. Bio-optical characteristics of Cariaco Basin (Caribbean Sea) waters. *Continental Shelf Research* 31:582–93.

Martin AP. 2003. Phytoplankton patchiness: The role of lateral stirring and mixing. *Progress In Oceanography* 57:125–74.

Mühling M, Fuller NJ, Millard A, Somerfield PJ, Marie D, Wilson WH, et al. 2005. Genetic diversity of marine *Synechococcus* and co-occurring cyanophage communities: Evidence for viral control of phytoplankton. *Environmental Microbiology* 7:499–508.

Müller-Karger FE, Varela R, Thunell R, Scranton M, Bohrer R, Taylor G, et al. 2001. Annual cycle of primary production in the Cariaco Basin: Response to upwelling and implications for vertical export. *Journal of Geophysical Research* 106:4527–42.

Müller-Karger FE, Varela R, Thunell R, Astor Y, Zhang H, Luerssen R, et al. 2004. Processes of coastal upwelling and

- carbon flux in the Cariaco Basin. *Deep Sea Research II* 51:927–43.
- Müller-Karger FE, Varela R, Thunell R, Luerssen R, Hu C, Walsh JJ. 2005. The importance of continental margins in the global carbon cycle. *Geophysical Research Letters* 32:L01602. 5 pages.
- Müller-Karger FE, Varela R, Thunell RC, Scranton MI, Taylor GT, Astor Y, et al. 2010. The CARIACO Oceanographic Time Series. In: Liu K-K, Atkinson L, Quinones R, Talaue-McManus L, editors. *Carbon and Nutrient Fluxes in Continental Margins: A Global Synthesis*. JGOFS Continental Margins Task Team (CMTT). Berlin: Springer-Verlag, p 454–463.
- Mutshinda CM, O'Hara RB, Woiwod IP. 2009. What drives community dynamics? *Proceedings of the Royal Society London B* 276:2923–29.
- Mutshinda CM, O'Hara RB, Woiwod IP. 2011. A multispecies perspective on ecological impacts of climatic forcing. *Journal of Animal Ecology* 80:101–07.
- Peterson LC, Haug GH. 2006. Variability in the mean latitude of the Atlantic Intertropical Convergence Zone as recorded by riverine input of sediments to the Cariaco Basin (Venezuela). *Palaeogeography Palaeoclimatology Palaeoecology* 234:97–113.
- Quigg AS, Irwin AJ, Finkel ZV. 2011. Evolutionary inheritance of elemental stoichiometry in phytoplankton. *Proceedings of the Royal Society London B* 278:526–34.
- R Development Core Team. 2011. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. Computer program.
- Richards FA. 1975. The Cariaco Basin (Trench). *Oceanography and Marine Biology Annual Review* 13:11–67.
- Six C, Finkel ZV, Irwin AJ, Campbell DA. 2007. Light variability illuminates niche-partitioning among marine picocyanobacteria. *PLOS One* 2(12):e1341. 6 pages.
- Taylor GT, Müller-Karger F, Thunell RC, Scranton MI, Astor Y, Varela R, Troccoli-Ghinaglia L. et al. 2012. Ecosystem responses in the southern Caribbean Sea to global climate change. *Nature Climate Change*. Abstract. *Ocean Sciences Meeting* 2012. <http://www.sgmeet.com/osm2012/viewabstract2.asp?AbstractID=9715> (accessed 24 November 2012).
- Tedesco K, Thunell R. 2003. High resolution tropical climate record for the last 6,000 years. *Geophysical Research Letters* 30:1–4.
- Tilman, D. 1982. *Resource competition and community structure*. Princeton University Press, Princeton, New Jersey, USA. 296 pages.
- Thomas A, O'Hara RB, Ligges U, Sturtz S. 2006. Making BUGS Open. *R News* 6:12–17.
- van de Poll WH, Visser RJW, Buma AGJ. 2007. Acclimation to a dynamic irradiance regime changes excessive irradiance sensitivity of *Emiliana huxleyi* and *Thalassiosira weissflogii*. *Limnology and Oceanography* 52:1430–38.

Editorial responsibility: David Montagnes