JOURNAL OF PLANKTON RESEARCH | VOLUME 33 | NUMBER 1 | PAGES 13–24 | 2011

Predicting marine phytoplankton community size structure from empirical relationships with remotely sensed variables

CAROLYN BARNES[1]*, XABIER IRIGOIEN[2], JOSÉ A. A. DE OLIVEIRA[1], DAVID MAXWELL[1] AND SIMON JENNINGS[1,3]

[1]CENTRE FOR ENVIRONMENT, FISHERIES AND AQUACULTURE SCIENCE, LOWESTOFT, SUFFOLK NR33 0HT, UK, [2]AZTI - TECNALIA/MARINE RESEARCH DIVISION, HERRERA KAIA PORTUALDEA Z/G 20110, PASAIA (GIPUZKOA), SPAIN AND [3]SCHOOL OF ENVIRONMENTAL SCIENCES, UNIVERSITY OF EAST ANGLIA, NORWICH NR7 4TJ, UK

*CORRESPONDING AUTHOR: carolyn.barnes@cefas.co.uk

Received February 9, 2010; accepted in principle June 18, 2010; accepted for publication June 25, 2010

Corresponding editor: William K.W. Li

The size composition of primary producers has a potential influence on the length of marine food chains and carbon sinking rates, thus on the proportion of primary production (PP) that is removed from the upper layers and available to higher trophic levels. While total rates of PP are widely reported, it is also necessary to account for the size composition of primary producers when developing food web models that predict consumer biomass and production. Empirical measurement of size composition over large space and time scales is not feasible, so one approach is to predict size composition from environmental variables that are measured and reported on relevant scales. Here, we describe relationships between the environment and the size composition of phytoplankton communities, using a collation of empirical measurements of size composition from sites that include polar, tropical and upwelling environments. The size composition of the phytoplankton communities can be predicted using two remotely sensed variables, chlorophyll-a concentration and sea surface temperature. Applying such relationships in combination allows prediction of the slope and location of phytoplankton size spectra and estimation of the percentage of different sized phytoplankton groups in communities.

KEYWORDS: primary production; community; phytoplankton; trophic level; transfer efficiency; size composition; picoplankton; nanoplankton; microplankton

INTRODUCTION

Global primary production (PP) by marine phytoplankton is around 5×10^{10} tonnes C year^{-1} (Carr *et al.*, 2006), but there are large regional variations in the proportion of this production being removed from the upper layer through sinking and therefore available to higher trophic levels. This is due to variations in absolute productivity among regions, with 50% of production estimated to come from 27% of ocean area (Longhurst *et al.*, 1995), and to regional differences in phytoplankton community structure. Regionally, the factors that affect the availability of phytoplankton to a given size class of consumers are their (i) spatial and temporal distribution, (ii) palatability, (iii) abundance and (iv) size composition [due to morphological constraints of consumers and because cell individual sinking rates are related to size (Smayda, 1971)].

doi:10.1093/plankt/fbq088, available online at www.plankt.oxfordjournals.org. Advance Access publication August 19, 2010

Marine pelagic food chains are strongly size based, with larger predators eating smaller prey. This size-based predation is predominantly responsible for the transfer of energy from phytoplankton to progressively larger animals and total production falls with body mass as trophic level rises (Sheldon *et al.*, 1972). Mean ratios between predator and prey body mass (predator–prey mass ratio, PPMR) are relatively constant in marine ecosystems (Barnes *et al.*, 2010). Consequently, when primary producers are smaller, there are, on average, more steps in a food chain to a predator of given size. Further, since mean annual trophic transfer efficiency at each step is also relatively constant (Barnes *et al.*, 2010), production by consumers of a given body size will be a smaller proportion of PP in regions where the primary producers are smaller. Such an effect has been shown in lakes (Sprules and Munawar, 1986), although it has not yet been reported in large-scale observations in marine systems (San Martin *et al.*, 2006b).

With knowledge of PPMR and the factors that influence trophic transfer efficiency, estimates of PP can be used to predict production at higher trophic levels (Dickie, 1976). For global scale analyses of the transfer of energy from phytoplankton to higher trophic levels, it is often convenient to use PP estimates based on satellite measurements of surface chlorophyll concentration provided by the Sea-viewing Wide Field-of-view Sensor (SeaWiFS) global time series (McClain *et al.*, 2004). Some progress has been made with estimating phytoplankton cell sizes by linking phytoplankton absorption to phytoplankton size classes using a single variable, the optical absorption by phytoplankton at 443 nm, which can be derived from the inversion of ocean colour data (Hirata *et al.*, 2008), but a complementary approach is to identify general relationships between remotely sensed environmental variables and the size composition of phytoplankton communities.

Here, we seek to identify relationships between the observed size composition of phytoplankton communities and remotely sensed environmental variables. Our aim is to use readily available remotely sensed environmental variables such as surface chlorophyll and sea surface temperature (SST) to enable estimations of phytoplankton size parameters for input to models. To provide a comprehensive description of the size and relative abundance of cells, we used four descriptors of community structure: (i) mean mass ($\log_{10}$), (ii) the variance of mass ($\log_{10}$), (iii) the slope of size spectra [relationship between the logarithm of total abundance by cell mass and the logarithm of cell mass, irrespective of species identity (Sheldon and Parsons, 1967)] and (iv) the range of cell masses that encompass a given proportion of total biomass or production. This information is necessary to predict phytoplankton production by size class in inputs to size-based models of production at higher trophic levels (Jennings *et al.*, 2008; Blanchard *et al.*, 2009).

METHOD

Abundance and species composition were determined for phytoplankton in 361 water samples collected at 12 sites: five transects from 48°N to 50°S in the Atlantic Ocean [hereafter AMT1-5, $n = 125$, taken at 7 m (second samples at each site, taken at the deep chlorophyll-a maximum were excluded from this analysis)], the Benguela upwelling ($n = 54$), mesocosms in the Bergen fjord ($n = 46$), the Irminger Sea ($n = 59$), Long Island Sound ($n = 7$), the North Sea ($n = 44$), the Norwegian Sea ($n = 19$) and the Oregon upwelling ($n = 7$). All included samples were taken in subsurface waters. See Irigoien *et al.* (Irigoien *et al.*, 2005) for details. Subsamples (100 mL) were settled [Utermőhl technique (Lund *et al.*, 1958)] and individuals counted at the species level with an inverted microscope. Heterotrophic species were excluded from the analysis. Picoplankton was measured using flow cytometry [see Irigoien *et al.* (2004) for details]. Biomass was calculated as the product of numerical abundance and cell mass. More details of sample positions, collection, processing and composition are provided by Irigoien *et al.* (Irigoien *et al.*, 2004, 2005).

To assess the proportion of phytoplankton biomass (B) that was attributable to cells in specified mass (M) ranges, we expressed cumulative B as a function of M by fitting:

$$B_{\mathrm{cum}}(M) = \frac{\left(\frac{a}{(b+1)}\right)(M^{b+1} - M_{\mathrm{min}}^{b+1})}{\left(\frac{a}{(b+1)}\right)(M_{\mathrm{max}}^{b+1} - M_{\mathrm{min}}^{b+1})} \quad (1)$$

Where b is size spectrum slope, log a is size spectrum intercept and M_{min} and M_{max} are the minimum and maximum M (see Supplementary material for derivation). Equation (1) was also fitted to cumulative production (P_{cum}) at M, where b and log a are the slope and intercept of the relationship between the logarithm of P at M and the logarithm of cell mass. The fitted relationships were used to predict the M ranges that contributed to 10, 50 and 90 of total B or P (Fig. 1). For example,

$$M_{\mathrm{B50}} = \left(\frac{50}{100}(M_{\mathrm{max}}^{b+1} - M_{\mathrm{min}}^{b+1}) + M_{\mathrm{min}}^{b+1}\right)^{\frac{1}{b+1}} \quad (2)$$

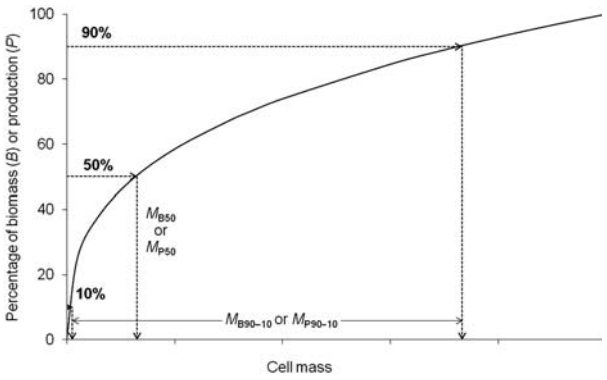

Fig. 1. A graphical summary of the definitions of cell mass ranges that account for various percentages of biomass and production. M_{B50} (M_{P50}) is the cell mass at which 50% of biomass (production) is reached. $M_{B90\text{-}10}$ ($M_{P90\text{-}10}$) is the range of cell masses that make up the mid 80% [i.e. $\log_{10}$(cell mass at 90%) $-\log_{10}$(cell mass at 10%) of biomass (production)].

Since P was not measured directly, we calculated the net production rate (R) of individuals from M, accounting for photosynthetic active radiation (PAR) following López-Urrutia *et al.* (López-Urrutia *et al.*, 2006):

$$\ln R = \ln(\mathcal{N}c) + \alpha \times \ln(M) - E \times \left(\frac{1}{kT}\right) + \ln\left(\frac{\mathrm{PAR}}{\mathrm{PAR} + K_\mathrm{m}}\right) \quad (3)$$

where $\mathcal{N}c$ is the normalization constant ($\ln(\mathcal{N}c) = -11.28$); α is the allometric exponent ($\alpha = 1.05$); E is the activation energy for photosynthetic reactions ($E = 0.29$) and PAR/(PAR + K_m) the Michaelis–Menten photosynthetic light response where PAR is the photosynthetically available radiation at the sample site and K_m is the half-saturation constant ($K_\mathrm{m} = 1.51$) that represents the amount of quanta at which half the maximum photosynthetic activity is reached. R is expressed as metabolic rate in mmol of O_2 day^{-1} and M is expressed in pg C, T is the absolute temperature and PAR is irradiance in mol photons (Einsteins) m^{-2} day^{-1} (2003 PAR values for each sample site were taken from http://oceandata.sci.gsfc.nasa.gov/SeaWiFS/Mapped/Annual/par/). Rate in mmol of O_2 day^{-1} was converted to rate in pg of C year^{-1} (molar mass of C = 12.01):

$$\text{Rate in pg C year}^{-1} = \frac{R}{12.01} \times 365 \times 10^{12} \quad (4)$$

Biomass of each species of phytoplankton in each sample was calculated by multiplying the abundance by species mean M. All data manipulation and analyses were performed using R (R-Development-Core-Team, 2007).

To provide a long-term description of the environment at the sampling sites, we estimated mean annual SST, surface chlorophyll-a concentration and PP at each site. SST data were derived from the Moderate-resolution Imaging Spectroradiometer (MODIS) aboard the NASA satellites. Monthly SST averages for 2003 were extracted through the Jet Propulsion Laboratory physical oceanography DAAC web portal (http://poet.jpl.nasa.gov/) and averaged to give a mean annual value at a scale of 36 km^2. Estimates of chlorophyll-a were taken from Sea-viewing Wide Field-of-view Sensor (SeaWiFS) data (http://orca.science.oregonstate.edu/1080.by.2160.monthly.hdf.chl.seawifs.php) by calculating an annual mean value for each sampling location from data extracted for each month in 2003. Where no value was available for a given month at a particular location then the value from the nearest available location for that month was used. PP was computed from a wavelength- and depth-resolved model (Mélin, 2003), building on the approach of Longhurst *et al.* (Longhurst *et al.*, 1995). The main biological input to the models was the surface concentration of chlorophyll-a pigment provided by SeaWiFS (see above). All changes from the implementation of Longhurst *et al.* (Longhurst *et al.*, 1995) are detailed in Mélin (Mélin, 2003). Outputs were calculated on a 36 km grid. Annual PP was obtained by averaging positive values (mg C m^{-2} day^{-1}) over the number of available months, except at high latitudes where it was normalized to 12 (as ocean colour has no good coverage of wintertime high latitudes, owing to the presence of cloud cover and sea ice). The mean annual values were assigned to each station location. When station locations had to be matched to point locations rather than onto a grid, we attempted the match in the following order; (i) number of degrees to one decimal place for both latitude and longitude, (ii) number of degrees to one decimal place for latitude, rounded whole number of degrees for longitude, (iii) rounded whole number of degrees for latitude and number of degrees with one decimal place for longitude, (iv) rounded whole number of degrees for both latitude and longitude. All were successfully matched (split between the four priorities 0, 8, 5 and 87%, respectively).

For some comparisons, cell size had to be converted between equivalent spherical diameter (ESD) and M. For example, a phytoplankton cell with ESD 2 μm has carbon content of approximately 0.8 pg (−0.08 on the $\log_{10}$ pg scale used in our Figures) using the conversion

$$\text{pg C year}^{-1} = 0.216 \times \text{volume}^{0.939}\,\mu\text{m}^{-3} \quad (5)$$

reported for taxonomically diverse protist plankton (Menden-Deuer and Lessard, 2000).

Relationships were explored between the remotely sensed environmental variables PP, SST and chlorophyll-a and (i) cell mass that accounted for 50% of total phytoplankton biomass (i.e. 50% on cumulative biomass curve) (hereafter M_{B50}), (ii) cell mass range that included 80% of total biomass (range of cell masses from 10 to 90% of biomass on cumulative biomass curve) ($M_{B90\text{-}10}$), (iii) cell mass that accounted for 50% of total phytoplankton production (i.e. 50% on cumulative production curve) (M_{P50}), (iv) cell mass range that included 80% of total production ($M_{P90\text{-}10}$), (v) size spectra slope (b) and (vi) size spectra intercept (a). Prior to performing ANOVA, where necessary the data were log-transformed to achieve normality and equality of variance.

To identify the remotely sensed environmental variables that best predicted the properties of the phytoplankton communities, we randomly divided the data into equal-sized "training" and "predicting" data sets (Tian *et al.*, 2007). We fitted linear models to the cases in the training set and then evaluated how well these models predicted the properties of the phytoplankton communities in the predicting set. The training set models were of the form:

$$\Upsilon(t) = f\,(X(t);\theta) + \text{error} \tag{6}$$

where the $\Upsilon(t)$ variable was related to a function of a subset of the environmental variables $X(t)$ and the parameters θ were estimated by least squares. The models were assumed to be normally distributed with mean zero and constant variance. The effect of each variable as a predictor was tested individually; then each was tested as the second variable along with the variable previously identified as the best predictor.

The performance of the training set models was evaluated using the prediction data set and we described performance with the following summary statistic (Tian *et al.*, 2007):

$$D = |\Upsilon(p) - \mathcal{Z}(p);\hat{\theta}| \tag{7}$$

where $\Upsilon(p)$ are the Υ values in the prediction data set and $\mathcal{Z}(p)$ contains the predicted values of $\Upsilon(p)$ based on the training data set model with its estimated parameter vector $\hat{\theta}$. This statistic can be interpreted as prediction error.

We report the mean values of D from 1000 calculations based on different random choices of the training and prediction data sets. This approach ensures that our results are not affected by the selection of training and prediction data sets. We refer to the mean D as $\bar{D}$ and use it to evaluate our models (rather than a summary of the fit of the model) because we want to use our model for prediction (models with the best fit often include more variables than those that give the best prediction). We then investigate the explanatory power of the best prediction model by fitting the model to the whole data set.

Applying these relationships in combination allows prediction of the slope, intercept and location of phytoplankton size spectra with respect to M. The location of the cell mass range is calculated so that the integrated biomass to either side of the mid-point is equal. Thus the value for mass at 10% (M_{B10}) is calculated as:

$$M_{B10} = M_{B50}\left[\frac{10^{(b+1)\log_{10}(M_{B90-10})}+1}{2}\right]^{\frac{-1}{b+1}} \tag{8}$$

where M_{B50} is mid-point mass, b is slope and $\log_{10}(M_{B90\text{-}10})$ is $(\log_{10}(M_{B90})-\log_{10}(M_{B10}))$, the values predicted by the models. The derivation is shown in the Supplementary Material. M_{B90} is then calculated as:

$$M_{B90} = 10^{(\log_{10}(M_{B10})+\log_{10}(M_{B90-10})} \tag{9}$$

Equation (1) did not always adequately fit the very variable tails of the cumulative distributions. So, to calculate 100% of cumulative B, the values of M_{B0} and M_{B100} were each calculated to incorporate the additional 20% of integrated biomass.

$$M_{B0} = \left(-\frac{5}{4}M_{B50}^{b+1} + \frac{5}{4}M_{B10}^{b+1} + M_{B50}^{b+1}\right)^{\frac{1}{b+1}} \tag{10}$$

$$M_{B100} = \left(\frac{5}{4}M_{B90}^{b+1} - \frac{5}{4}M_{B50}^{b+1} + M_{B50}^{b+1}\right)^{\frac{1}{b+1}} \tag{11}$$

The percentage contribution of a particular cell mass to total community biomass or production can then be calculated for a given SST and chlorophyll-a. The derivations of the equations are shown in Supplementary Material.

RESULTS

Mean cell mass ranged from 243 pg of C (AMT3) to 4740 pg of C (Bergen fjord) and variance was lowest in AMT1 and highest in the Bergen fjord (Table I). M_{B50} ranged from 0.6 pg C (AMT3) to 1314 pg C (Long Island Sound) and M_{B90-10} ranged from 43 pg C (Long Island Sound) to 13 438 pg C (AMT4) (Table I). M_{P50} ranged from 1.1 pg C (AMT3) to 1476 pg C (Long Island Sound) and M_{P90-10} ranged from 33 pg C (Long Island Sound) to 18 157 pg C (AMT4) (Table I).

Across all sites the mean size spectra slope was −1.18, ranging from −1.61 in the Norwegian Sea to

Table I: Phytoplankton cell mass (pg C) statistics from 12 sites that include polar, tropical and upwelling environments

	%	AMT1	AMT2	AMT3	AMT4	AMT5	Benguela	Bergen fjord	Irminger	Long Island Sound	North Sea	Norwegian Sea	Oregon upwelling
Mean cell mass		253	279	243	316	246	912	4740	381	742	788	267	702
1 standard deviation		670	1150	817	941	954	4844	99418	1083	2133	3379	1052	1728
1 standard error		5.5	9.9	6.6	7.3	7.5	28	635	6.4	33	25	14	24
Smallest cell mass		0.1	0.1	0.01	0.01	0.01	0.1	1.0	0.2	1.0	1.0	1.0	0.2
Largest cell mass		8206	27 123	10 686	15 221	27 123	49 930	2 799 208	14 709	27 123	143 195	7000	14 709
Cell mass that accounts for % of total biomass	10	0.55	0.18	0.018	0.018	0.022	1.2	3.7	0.41	122	2.7	1.3	0.74
	25	4.7	0.70	0.052	0.055	0.12	16	18	1.2	454	14	2.3	4.5
	50	40	8.8	0.58	0.82	2.6	160	88	11	1314	131	11	70
	75	161	79	15	23	36	904	612	136	2878	805	72	781
	80	209	123	32	47	61	1292	1221	237	3441	1139	108	1231
	90	392	346	188	237	204	2950	7537	793	5291	2352	255	3064
Cell mass range that accounts for the mid-80% of total biomass		709	1934	10 471	13 439	9303	2383	2022	1922	43	861	197	4145
Cell mass that accounts for % of total production	10	0.62	0.20	0.020	0.020	0.031	3.3	7.8	0.45	170	3.4	1.3	0.9
	25	5.3	0.95	0.07	0.08	0.30	32.8	33.7	1.6	561	21	2.4	7.1
	50	43	12	1.1	1.7	5.7	250	159	16	1476	185	12	112
	75	174	103	29	43	57	1207	2120	190	3111	1000	79	1029
	80	228	158	59	85	90	1670	4191	318	3702	1373	118	1548
	90	454	442	291	365	286	3537	18 409	963	5584	2658	272	3507
Cell mass range that accounts for the mid-80% of total production		732	2207	14 500	18 157	9159	1063	2351	2121	33	776	207	3734

−0.64 in Long Island Sound. Mean intercept ($\log_{10}$ C pg) was 9.5 ranging from 8.89 in AMT5 to 10.57 in the Bergen fjord, all $P < 0.001$ (Table II).

There was a positive linear relationship between M_{B50} ($\log_{10}$) and both, primary production ($\log_{10}$) and chlorophyll-a ($\log_{10}$) and an inverse linear relationship between M_{B50} ($\log_{10}$) and sea surface temperature. There was a positive linear relationship between M_{B90-10} ($\log_{10}$) and both primary production ($\log_{10}$) and sea surface temperature and an inverse linear relationship between M_{B90-10} ($\log_{10}$) and chlorophyll-a ($\log_{10}$) but the p-value for the slope was not significant (Fig. 2 and Supplementary material Table S1). The production relationships were similar to those for biomass, except that the M_{P90-10} ($\log_{10}$) and primary production ($\log_{10}$) relationship was not significant at the 95% confidence level ($P = 0.93$) and both slope and intercept were significant for $M_{P90\text{-}10}$ ($\log_{10}$) against chlorophyll-a ($\log_{10}$) (Fig. 3 and Supplementary material Table S1).

There were significant positive linear relationships between size spectra slope and the variables primary production ($\log_{10}$) and chlorophyll-a ($\log_{10}$) and but no significant relationship between slope and sea surface temperature (Fig. 4 and Supplementary material Table S1). Linear relationships between size spectra intercept ($\log_{10}$) and both primary production ($\log_{10}$) and chlorophyll-a ($\log_{10}$) were also significant and positive and there was a significant inverse relationship with sea surface temperature (Fig. 4 and Supplementary material Table S1).

Table II: Properties and statistical analysis of normalized biomass size spectra from 12 sites that include polar, tropical and upwelling environments

	AMT1	AMT2	AMT3	AMT4	AMT5	Benguela	Bergen fjord	Irminger	Long Island Sound	North Sea	Norwegian Sea	Oregon upwelling
No. samples	25	25	25	26	24	54	46	59	7	44	19	7
Mean slope	−1.28	−1.24	−1.18	−1.23	−1.18	−1.16	−1.27	−1.18	−0.64	−1.16	−1.61	−1.03
Mean p-value	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
Mean intercept ($\log_{10}$ pg C)	9.36	9.28	9.06	9.12	8.89	9.89	10.57	9.61	9.07	9.69	9.59	9.91
Slope coefficient of variation	−0.180	−0.097	−0.068	−0.098	−0.093	−0.147	−0.126	−0.110	−0.594	−0.216	−0.168	−0.146
Intercept coefficient of variation	0.035	0.028	0.029	0.026	0.029	0.043	0.044	0.053	0.072	0.073	0.043	0.047

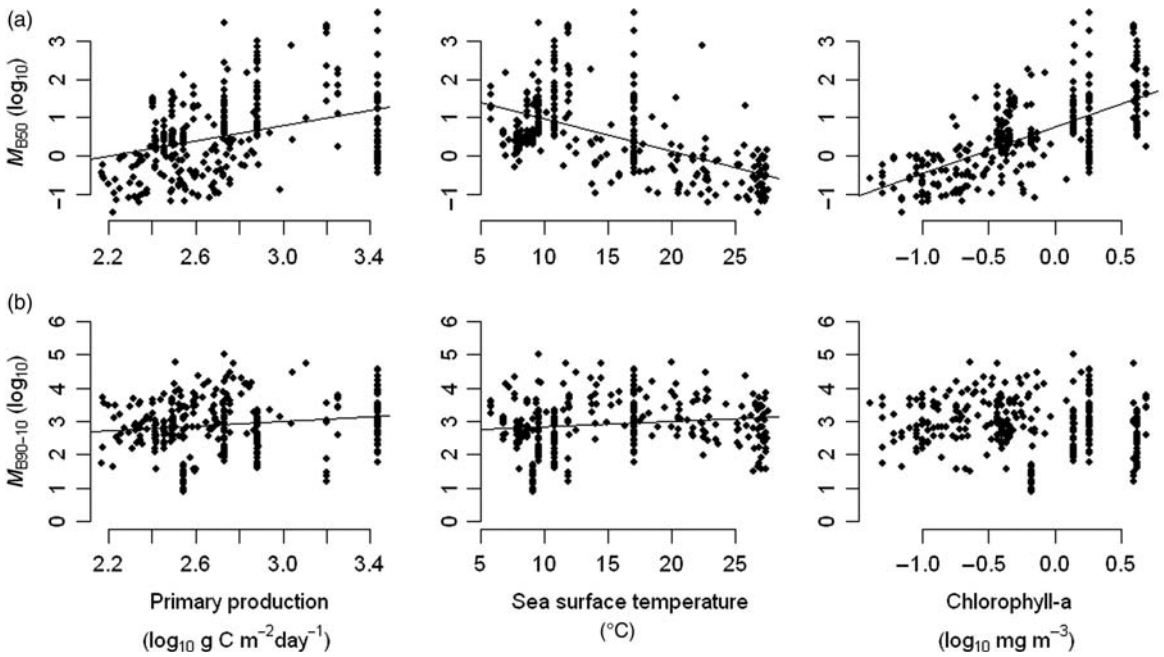

Fig. 2. Relationships between (**a**) M_{B50}, the mass of phytoplankton cells that account for 50% of total biomass and (**b**) $M_{B90\text{-}10}$, the range of phytoplankton cell masses that account for the mid 80% of total biomass and remotely sensed estimates of primary production, sea surface temperature and chlorophyll-a at 12 sites.

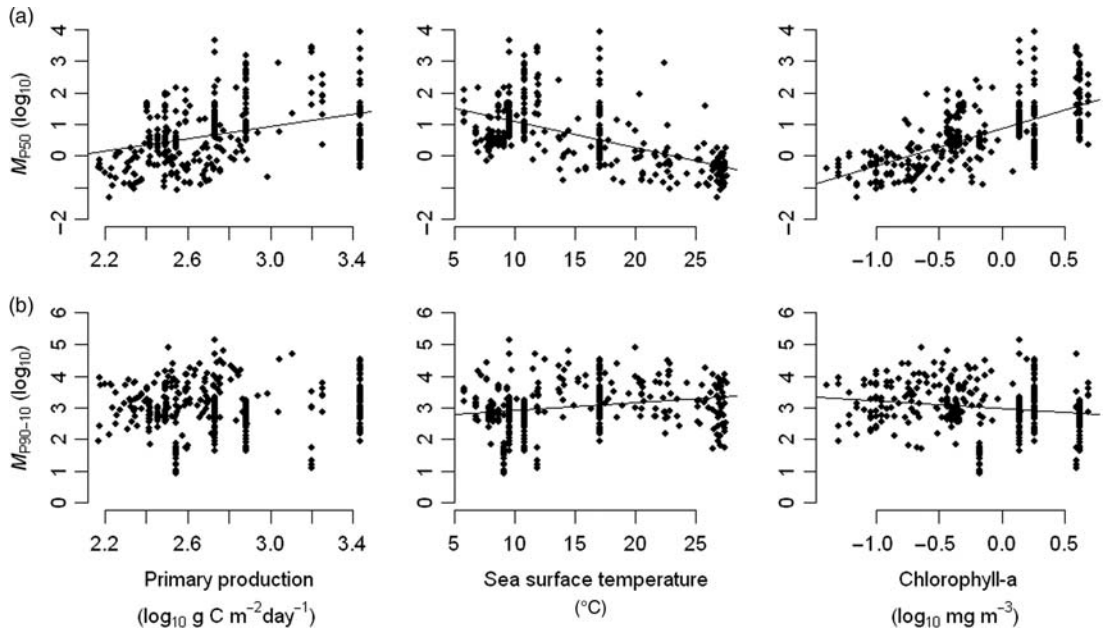

Fig. 3. Relationships between (**a**) M_{P50}, the mass of phytoplankton cells that account for 50% of total production and (**b**) $M_{P90\text{-}10}$, the range of phytoplankton cell mass that account for the mid 80% of total production and remotely sensed estimates of primary production, sea surface temperature and chlorophyll-a at 12 sites.

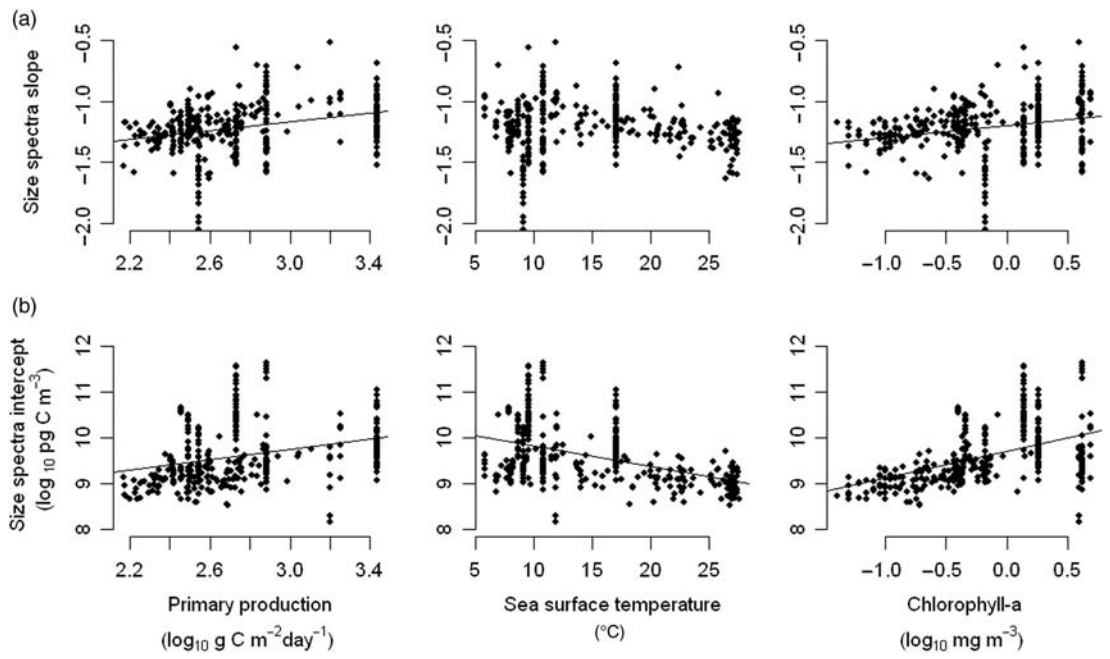

Fig. 4. Relationships between (**a**) size spectra slopes and (**b**) size spectra intercepts and remotely sensed estimates of primary production, sea surface temperature and chlorophyll-a at 12 sites.

Prediction results

Prediction models are useful only for application on large space and time scales when the explanatory variables are readily available, so we investigated only the prediction using primary production, sea surface temperature and chlorophyll estimates as these can easily be derived from remote sensing.

Chlorophyll-a ($\log_{10}$) and SST together successfully predicted M_{B50} and M_{P50}. The ranges M_{B90-10} and M_{P90-10} were less predictable, although SST could be

used to predict both ranges better than just taking the mean value of all samples. Similarly, the slope of the size spectrum was slightly more predictable using chlorophyll-a ($\log_{10}$) than just taking the mean value of all samples. The size spectrum intercept ($\log_{10}$) could be predicted well using the single variable chlorophyll-a ($\log_{10}$). The prediction equations are given in Table III and the statistical results are given in Supplementary Table Supplementary material Table S2.

Applying these relationships in combination allows prediction of the slope, height and location of the phytoplankton size spectra over relevant size ranges (Fig. 5) and the make-up of communities by size class (see Supplementary Material). For example, where SST is 15°C and chlorophyll-a concentration is 1.0 mg m^{-3} there is total biomass of 31.8 $\log_{10}$ pg C, of which 13% is picoplankton (ESD $< 2\mu$m), 72% is nanoplankton (ESD > 2 but $< 20\,\mu$m) and 15% is microplankton (ESD $> 20\,\mu$m).

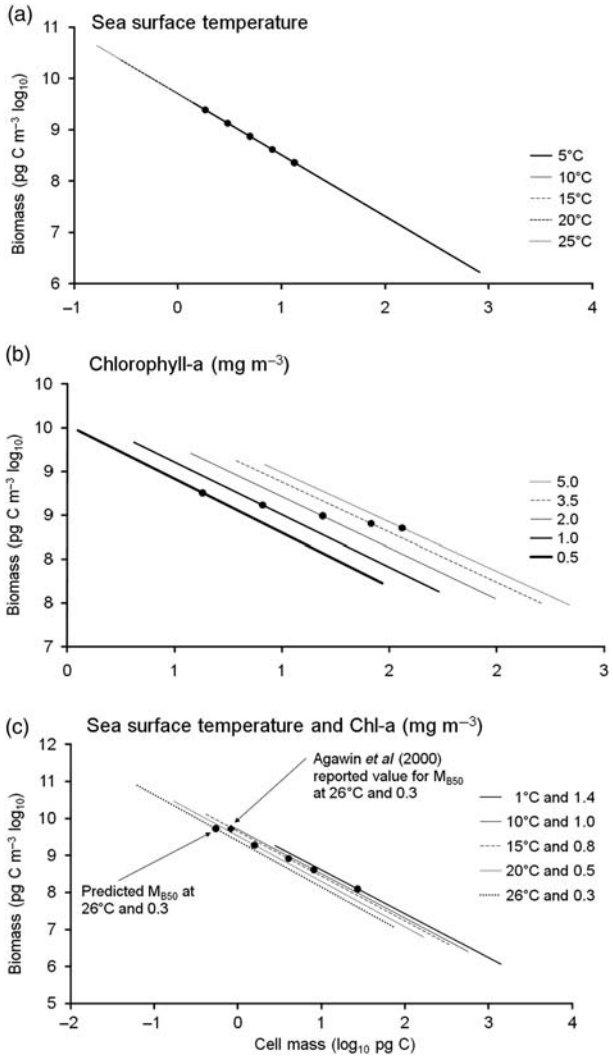

Fig. 5. Size spectra predictions for (**a**) a range of sea surface temperatures at an intermediate chlorophyll-a level of 1 mg m^{-3}; (**b**) a range of chlorophyll-a levels at 10°C and (**c**) a range of sea surface temperatures and a range of chlorophyll-a levels. Circle (filled) show position of M_{B50}, the mass of phytoplankton cells that account for 50% of total biomass. See text for explanation for mid-point prediction comparison with Agawin *et al.* (2000).

DISCUSSION

The analyses reveal that relationships between the size composition of phytoplankton communities and the environment can be used to predict mean cell mass, the slopes of size spectra and the range of cell masses that encompass a given proportion of total biomass or production. Since predictions are based on temperature and chlorophyll estimates that are available at high resolution and over large spatial scales, for example from remote sensing, they can be used to predict community size structure at the same scales. Such predictions provide necessary inputs to models that seek to link PP to production at higher trophic levels.

Chlorophyll concentration varies due to a variety of physical and chemical factors. Irradiance over the mixed layer depth, surface nitrate, SST and latitude and longitude together can predict 83% of the variation in log chlorophyll in the North Atlantic (Irwin and Finkel, 2008). We are not here investigating the mechanisms leading to the correlation between phytoplankton size composition and chlorophyll; instead, we accept the practical value of the relationship pending research that identifies mechanistic relationships. Tight correlations

Table III: Equations for predicting phytoplankton cell sizes from remotely sensed variables

	Predicted Value	SST (°C) multiplier	$\log_{10}$ [chlorophyll-a (mg m^{-3})] multiplier	Constant
M_{B50}	$\log_{10}$ [cell size (pg C) for 50% of biomass]	−0.043	0.929	1.340
M_{B90-10}	$\log_{10}$ [cell size range (pg C) for mid 80% of biomass]	0.015	0	2.689
M_{P50}	$\log_{10}$ [cell size (pg C) for 50% of production]	−0.040	0.916	1.432
M_{P90-10}	$\log_{10}$ [cell size range (pg C) for mid 80% of production]	0.025	0	2.655
Slope	Slope of the size spectrum	0	0.099	−1.196
Intercept	$\log_{10}$ (intercept of the size spectrum)	0	0.585	9.704

between temperature, chlorophyll-a and nitrate concentrations are known suggesting that the environmental factors are highly intertwined and strongly regulate the phytoplankton average cell size (Chen and Liu, 2010).

Picophytoplankton made up 50% of the biomass at chlorophyll-a concentrations less than 0.21 mg m^{-3} and at temperatures over 22.3°C. At this temperature, 50% of production was from cells up to 2.3 μm. These relationships are broadly consistent with the findings of Agawin *et al.* (Agawin *et al.*, 2000) who reported that picophytoplankton dominated (≥50%) where chlorophyll-a concentration was lower than 0.3 mg m^{-3} and they reported picophytoplankton dominating in waters over 26°C. Using our prediction models, which use both temperature and chlorophyll-a concentration, we estimate that at 26°C and 0.3 mg m^{-3} chlorophyll concentration, phytoplankton cells up to 1.7 μm diameter (0.55 pg C) make up 50% of the biomass (Fig. 5c) and up to 2.0 μm diameter (0.82 pg C) make up 50% of the production.

Slopes of phytoplankton size spectra are predicted to be steeper when total biomass is low but unaffected by temperature. If the size-spectrum slope is −1, then the trend is for all size classes to have equal biomass; if the slope is greater than −1 then biomass tends to increase with increasing cell mass; if the abundance–mass slope is less than −1 then biomass decreases with increasing cell mass. Modellers have often assumed slopes of either −0.75 or −1.0 when describing the size structure of the phytoplankton community but the model: "b = 0.099 $\log_{10}$(chlorophyll-a(mg m^{-3})) −1.196" suggests a slope of −1.2 at intermediate levels of chlorophyll-a. There are relatively few empirical estimates of size-spectrum slope for phytoplankton communities with which to compare our estimates, as most estimates are for communities that include zooplankton. However, those available provide qualitative support for our results. For example, Marañón *et al.* (Marañón *et al.*, 2007) reported that unproductive ecosystems were characterized by more negative slopes (−1.3 to −1.1) than productive ones (−0.8 and −0.6), but the slopes were rather shallower than reported in our study in which 11 of the 12 sites had a slope steeper (more negative) than −1. San Martin *et al.* (San Martin *et al.*, 2006a) also found that the slopes of biomass size spectra for the picoplankton and microplankton size ranges were positively related to biomass (but the pattern disappeared with the addition of mesozooplankton). They did not report relationships with temperature.

Size spectrum slopes as steep as −1.39 have been reported for pure phytoplankton (Huete-Ortega *et al.*, 2010). These values, like ours, are much steeper than the value of −0.75 that is theoretically predicted to result from the allocation of energy among competing individuals (Belgrano *et al.*, 2002). Our values are also steeper than the slope of −0.78 (Li, 2002) and slopes ranging from −0.74 to −1.06 (Cermeño and Figueiras, 2008) reported for phytoplankton. Differences in slope could be attributed to sampling artefacts or ecological processes. Smaller plankton are not effectively sampled by some gears, are not so well preserved after sampling and can be under-sampled during microscopic identification so we expect that any bias in the data will lead to the under representation of small cells and lower rather than higher estimates of slope (Harris *et al.*, 2000). Of necessity, deriving a size spectrum using mean values rather than individual mass will introduce bias. Biases introduced by working with mean mass in studies of relationships between abundance and body mass are greatest when working with species that have indeterminate growth and when the range of body sizes considered is narrow (Jennings *et al.*, 2007). These biases will be minimal for phytoplankton that have limited and determinate growth and a range in mean mass that spans over five orders of magnitude. Since sampling and analytical artefacts are unlikely to explain why our slopes are relatively negative, there may be an influence of ecological processes. For example, consumer density increases with temperature, leading to increased grazing pressure (O'Connor *et al.*, 2009), thus the slopes may increase as a result of greater predation rates on larger phytoplankton or uniform predation rates that have a greater impact on larger phytoplankton owing to their slower turnover times.

Our analyses do not determine cause and effect and are primarily intended to predict the size composition of phytoplankton communities from readily available large scale and high-resolution remote sensing data to support parameterization of food web models. However, direct linkages between temperature and the size composition of phytoplankton communities have been proposed (López-Urrutia, 2008). He suggests that large cells may dominate in colder systems owing to the different temperature dependence of heterotrophic and autotrophic rates. In colder areas, heterotrophs may not grow fast enough to control autotrophs, while in warmer areas large phytoplankton cells are likely to have relatively longer division times than smaller ones and may not be able to attain high levels of abundance given the high abundance of heterotrophic predators. We also expect community structure to be influenced by nutrient supply and the colder seas (as represented in the database) often correspond to areas with more nutrients (e.g. upwellings and coastal areas) than the warmer ones (e.g. oligotrophic subtropical gyres).

Other methods are being developed to predict the size composition of phytoplankton communities at high spatial resolution over large spatial scales. For example, Uitz *et al.* (Uitz *et al.*, 2006, 2008) have determined size spectra of phytoplankton communities from near-surface chlorophyll-a concentration using accessory pigments as markers for pico-, nano- and micro-plankton to infer the column-integrated phytoplankton biomass, its vertical distribution, and ultimately the community composition by quantifying on a global scale the phytoplankton biomass associated with each of the three algal assemblages. However, pigment analysis, based on high-performance liquid chromatography, although a widely used method for studying phytoplankton community composition does not provide any information on size structure in each group (Chen and Liu, 2010).

Our approach can be used to allocate phytoplankton biomass and PP to cell mass classes and thus improve the description of the primary producer community in size-based models that have been used to link PP and production at higher trophic levels. This is useful because estimates of phytoplankton production from NPZD (Nutrient Phytoplankton Zooplankton Detritus) models and remote sensing rarely provide information on the size composition of the phytoplankton community alone. At an SST of 5°C, our predictions suggest that the phytoplankton cell size at 50% of biomass ranges from 13 to 19 pg C, at 15°C it ranges from 5.0–7.2 pg C and at 25°C it ranges from 1.8 to 2.7 pg C for intermediate chlorophyll-a concentrations from 1 to 1.5 mg m^{-3}. At an SST of 5°C, the phytoplankton cell size at 50% of production ranges from 17 to 25 pg C, at 15°C from 6.9 to 10 pg C and at 25°C from 2.7 to 4.0 pg C over the same range of chlorophyll-a concentrations. The range of cell sizes that make up the majority of the community (the mid-80% of biomass/production) is unaffected by production but wider at warmer temperatures. One consequence of the dominance of smaller primary producers in less productive, warmer waters is that food chains will be longer. The ratio between the sizes of consumers and their prey (reported as the PPMR) does not depend on production or temperature (Barnes *et al.*, 2010), so the mean trophic level of a given size class of consumers would be expected to be higher in low productivity areas such as ocean gyres. Since smaller primary producers are linked to lower PP, but PPMR and transfer efficiency may be unrelated to the environment, production at higher trophic levels is disproportionately low when primary productivity is low.

In addition to predicting fluxes of energy to higher trophic levels and the biomass of consumer communities based on measurements of PP and temperature, our predictive relationships may also be valuable for predicting how the composition of phytoplankton communities may change in relation to environmental change. For example, Morán *et al.* (Morán *et al.*, 2009) have reported consistent relationships among temperature, cell size and picophytoplankton abundance and speculate that the size of cells in phytoplankton assemblages will gradually decrease as temperatures rise. Li *et al.* (Li *et al.*, 2009) concur that a reduction in community average cell size because of an increase in the abundance of individuals belonging to small-sized species may be a common response to increasing sea temperatures. There is evidence that reduced body size is the third universal ecological response to global warming besides the shift of species ranges toward higher altitudes and latitudes and the seasonal shifts in life-cycle events (Daufresne *et al.*, 2009). Such influences need to be considered in models that seek to predict how future changes in PP and temperature will affect production at higher trophic levels.

CONCLUSIONS

The size composition of primary producers is an essential input to enable estimations of food chain length, consumer biomass and production in any given location. We have described relationships between the environment and the size composition of phytoplankton communities using environmental variables that are easily estimated from ocean colour satellite measurements. Estimates of consumer biomass, production and trophic level depend on the length of food chains that support this biomass, a consequence of predator–prey size relationships and the size composition of primary producers. As mean predator–prey size ratios in marine ecosystems do not depend on temperature or PP, the size composition of the phytoplankton community has an overriding influence on food chain length which, in turn, can be used to further explore fish production, fisheries catch potential and the bioaccumulation of contaminants.

SUPPLEMENTARY MATERIAL

Supplementary data can be found online at http://plankt.oxfordjournals.org.

ACKNOWLEDGEMENTS

We would like to thank Roger Harris, Kevin Flynn, Ángel López-Urrutia, Doug Glazier and Rodney Forster for their thought-provoking comments and suggestions on how to improve the manuscript. We also

thank Frédéric Mélin for providing remote sensing data and Dan Reuman and Jon Barry for assistance with the statistical analyses. We thank two anonymous reviewers for their detailed comments that also helped us to improve the manuscript.

FUNDING

This work was supported by the United Kingdom Department for Environment, Food and Rural Affairs (M1001). The collection of the AMT data was supported by the UK Natural Environment Research Council through the Atlantic Meridional Transect consortium (NER/O/S/2001/00680). This is contribution number 192 of the AMT programme.

REFERENCES

Agawin, N. S. R., Duarte, C. M. and Agusti, S. (2000) Nutrient and temperature control of the contribution of picoplankton to phytoplankton biomass and production. *Limnol. Oceanogr.*, **45**, 591–600.

Barnes, C., Maxwell, D., Reuman, D. C. *et al.* (2010) Global patterns in predator-prey size relationships reveal size-dependency of trophic transfer efficiency. *Ecology*, **91**, 222–232.

Belgrano, A., Allen, A. P., Enquist, B. J. *et al.* (2002) Allometric scaling of maximum population density: a common rule for marine phytoplankton and terrestrial plants. *Ecol. Lett.*, **5**, 611–613.

Blanchard, J. L., Jennings, S., Law, R. *et al.* (2009) How does abundance scale with body size in coupled size structured food webs? *J. Anim. Ecol.*, **78**, 270–280.

Carr, M. E., Friedrichs, M. A. M., Schmeltz, M. *et al.* (2006) A comparison of global estimates of marine primary production from ocean color. *Deep-Sea Res. II Top. Stud. Oceanogr.*, **53**, 741–770.

Cermeño, P. and Figueiras, F. G. (2008) Species richness and cell-size distribution: Size structure of phytoplankton communities. *Mar. Ecol. Prog. Ser.*, **357**, 79–85.

Chen, B. and Liu, H. (2010) Relationships between phytoplankton growth and cell size in surface oceans: interactive effects of temperature, nutrients and grazing. *Limnol. Oceanogr.*, **55**, 965–972.

Daufresne, M., Lengfellner, K. and Sommer, U. (2009) Global warming benefits the small in aquatic ecosystems. *Proc. Natl Acad. Sci. USA*, **106**, 12788–12793.

Dickie, L. M. (1976) Predation, yield, and ecological efficiency in aquatic food chains. *J. Fish. Res. Board Can.*, **33**, 313–316.

Harris, R., Wiebe, P., Lenz, J. *et al.* (2000) *ICES Zooplankton Methodology Manual*. Academic Press, 684 pp.

Hirata, T., Aiken, J., Hardman-Mountford, N. J. *et al.* (2008) An absorption model to determine phytoplankton size classes from satellite ocean colour. *Remote Sens. Environ.*, **112**, 3153–3159.

Huete-Ortega, M., Marañón, E., Varela, M. *et al.* (2010) General patterns in the size scaling of phytoplankton abundance in coastal waters during a 10-year time series. *J. Plankton Res.*, **32**, 1–14.

Irigoien, X., Huisman, J. and Harris, R. P. (2004) Global biodiversity patterns of marine phytoplankton and zooplankton. *Nature*, **429**, 863–867.

Irigoien, X., Flynn, K. J. and Harris, R. P. (2005) Phytoplankton blooms: a 'loophole' in microzooplankton grazing impact? *J. Plankton Res.*, **27**, 313–321.

Irwin, A. J. and Finkel, Z. V. (2008) Mining a sea of data: Deducing the environmental controls of ocean chlorophyll. *PLoS ONE*, **3**.

Jennings, S., De Oliveira, J. A. A. and Warr, K. J. (2007) Measurement of body size and abundance in tests of macroecological and food web theory. *J. Anim. Ecol.*, **76**, 72–82.

Jennings, S., Mélin, F., Blanchard, J. L. *et al.* (2008) Global-scale predictions of community and ecosystem properties from simple ecological theory. *Proc. R. Soc. B*, **275**, 1375–1383.

Li, W. K. W. (2002) Macroecological patterns of phytoplankton in the northwestern North Atlantic Ocean. *Nature*, **419**, 154–157.

Li, W. K. W., McLaughlin, F. A., Lovejoy, C. *et al.* (2009) Smallest algae thrive as the Arctic Ocean freshens. *Science*, **326**, 539.

Longhurst, A., Sathyendranath, S., Platt, T. *et al.* (1995) An estimate of global primary production in the ocean from satellite radiometer data. *J. Plankton Res.*, **17**, 1245–1271.

López-Urrutia, Á. (2008) The metabolic theory of ecology and algal bloom formation. *Limnol. Oceanogr.*, **53**, 2046–2047.

López-Urrutia, Á., San Martin, E., Harris, R. P. *et al.* (2006) Scaling the metabolic balance of the oceans. *Proc. Natl Acad. Sci. USA*, **103**, 8739–8744.

Lund, J. W. G., Kipling, C. and Le Cren, E. D. (1958) The inverted microscope method of estimating algal numbers and the statistical basis of estimations by counting. *Hydrobiologia*, **11**, 143–170.

Marañón, E., Cermeño, P., Rodríguez, J. *et al.* (2007) Scaling of phytoplankton photosynthesis and cell size in the ocean. *Limnol. Oceanogr.*, **52**, 2190–2198.

McClain, C. R., Feldman, G. C. and Hooker, S. B. (2004) An overview of the SeaWiFS project and strategies for producing a climate research quality global ocean bio-optical time series. *Deep Sea Res. (II Top. Stud. Oceanogr.)*, **51**, 5–42.

Mélin, F. (2003) Potential of remote sensing for the analysis of the optical properties of the ocean-atmosphere system and application to estimates of phytoplankton photosynthesis. PhD Thesis. Université Paul Sabatier, Toulouse, France, 514 pp.

Menden-Deuer, S. and Lessard, E. J. (2000) Carbon to volume relationships for dinoflagellates, diatoms, and other protist plankton. *Limnol. Oceanogr.*, **45**, 569–579.

Morán, X. A. G., López-Urrutia, Á., Calvo-Díaz, a. *et al.* (2009) Increasing importance of small phytoplankton in a warmer ocean. *Global Change Biol.* (in press).

O'Connor, M. I., Piehler, M. F., Leech, D. M. *et al.* (2009) Warming and resource availability shift food web structure and metabolism. *PLoS Biol.*, **7**, e1000178.

R-Development-Core-Team. (2007) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

San Martin, E., Harris, R. P. and Irigoien, X. (2006a) Latitudinal variation in plankton size spectra in the Atlantic Ocean. *Deep-Sea Res. II Top. Stud. Oceanogr.*, **53**, 1560–1572.

San Martin, E., Irigoien, X., Harris, R. P. *et al.* (2006b) Variation in the transfer of energy in marine plankton along a productivity gradient in the Atlantic Ocean. *Limnol. Oceanogr.*, **51**, 2084–2091.

Sheldon, R. W. and Parsons, T. R. (1967) A continuous size spectrum for particulate matter in the sea. *J. Fish. Res. Board Can.*, **24**, 909–915.

Sheldon, R. W., Prakash, A. and Sutcliffe, W. H. (1972) The size distribution of particles in the Ocean. *Limnol. Oceanogr.*, **17**, 327–340.

Smayda, T. J. (1971) Normal and accelerated sinking of phytoplankton in the sea. *Mar. Geol.*, **11**, 105–122.

Sprules, W. G. and Munawar, M. (1986) Plankton size spectra in relation to ecosystem productivity, size, and perturbation. *Can. J. Fish. Aquat. Sci.*, **43**, 1789–1794.

Tian, L. U., Cai, T., Goetghebeur, E. *et al.* (2007) Model evaluation based on the sampling distribution of estimated absolute prediction error. *Biometrika*, **94**, 297–311.

Uitz, J., Claustre, H., Morel, A. *et al.* (2006) Vertical distribution of phytoplankton communities in open ocean: an assessment based on surface chlorophyll. *J. Geophys. Res.*, **111**.

Uitz, J., Huot, Y., Bruyant, F. *et al.* (2008) Relating phytoplankton photophysiological properties to community structure on large scales. *Limnol. Oceanogr.*, **53**, 614–630.

Vidondo, B., Prairie, Y. T., Blanco, J. M. *et al.* (1997) Some aspects of the analysis of size spectra in aquatic ecology. *Limnol. Oceanogr.*, **42**, 184–192.