

Combined constraints on global ocean primary production using observations and models

Erik T. Buitenhuis,¹ Taketo Hashioka,^{1,2,3} and Corinne Le Quéré¹

Received 27 September 2012; revised 30 May 2013; accepted 28 July 2013; published 30 August 2013.

[1] Primary production is at the base of the marine food web and plays a central role for global biogeochemical cycles. Yet global ocean primary production is known to only a factor of ~2, with previous estimates ranging from 38 to 65 Pg C yr⁻¹ and no formal uncertainty analysis. Here, we present an improved global ocean biogeochemistry model that includes a mechanistic representation of photosynthesis and a new observational database of net primary production (NPP) in the ocean. We combine the model and observations to constrain particulate NPP in the ocean with statistical metrics. The PlankTOM5.3 model includes a new photosynthesis formulation with a dynamic representation of iron-light colimitation, which leads to a considerable improvement of the interannual variability of surface chlorophyll. The database includes a consistent set of 50,050 measurements of ¹⁴C primary production. The model best reproduces observations when global NPP is 58 ± 7 Pg C yr⁻¹, with a most probable value of 56 Pg C yr⁻¹. The most probable value is robust to the model used. The uncertainty represents 95% confidence intervals. It considers all random errors in the model and observations, but not potential biases in the observations. We show that tropical regions (23°S–23°N) contribute half of the global NPP, while NPPs in the Northern and Southern Hemispheres are approximately equal in spite of the larger ocean area in the South.

Citation: Buitenhuis, E. T., T. Hashioka, and C. L. Quéré (2013), Combined constraints on global ocean primary production using observations and models, *Global Biogeochem. Cycles*, 27, 847–858, doi:10.1002/gbc.20074.

1. Introduction

[2] Primary production forms the basis of the food web and as such controls how much material and energy is available for the biosphere as a whole. It also forms the entry point of inorganic nutrients and carbon into the food web, and therefore, primary production is the ecosystem flux that is most directly influenced by anthropogenic changes in global biogeochemical cycles. Primary production is of similar magnitude in the land and marine biospheres, or ~100 Pg C yr⁻¹ for gross production and ~50 Pg C yr⁻¹ for net production (net primary production (NPP) = gross production minus autotrophic respiration [Prentice *et al.*, 2001]). This is where the similarity stops. Heterotrophic activity in the land biosphere is dominated by detritivory by bacteria, while in the

marine biosphere, it is dominated by herbivory by zooplankton. This leads to very different standing stocks of autotrophs: ~500 Pg C in land plants [Prentice *et al.*, 2001] and ~1 Pg C in marine autotrophic protists [Le Quéré *et al.*, 2005], with an average turnover time for autotrophs of ~10 years on land and ~1 week in the ocean.

[3] Primary production in the ocean has mostly been measured by incubating samples with ¹⁴C and measuring the difference in uptake between light and dark incubated bottles [Steemann Nielsen, 1952]. Interpretation of what ¹⁴C incubations measure varies, but there is evidence supporting the conclusion that daytime ¹⁴C uptake (which is the case for most in situ measurements) measures NPP [Marra, 2009]. Because there is no consensus on this question, we will assume that ¹⁴C uptake measures particulate net primary production (NPP), though this might introduce some unquantified bias into our results.

[4] NPP is the flux that is the most relevant to global ocean biogeochemical cycles and, therefore, the focus of this study. Initial estimates of global NPP were made by extrapolating in situ measurements in space and time (Figure 1). Large gaps in the data coverage led to a considerable spread in the estimated NPP from in situ measurements. More recently, global NPP has been estimated using remote sensing of surface chlorophyll [Antoine *et al.*, 1996; Behrenfeld and Falkowski, 1997; Mélin, 2003] or surface carbon [Behrenfeld *et al.*, 2005]. These satellite-based estimates all rely on assumptions about the vertical distribution of NPP and the C:Chl (or C:backscattering coefficient) ratio. The improved

Additional supporting information may be found in the online version of this article.

¹Tyndall Centre for Climate Change Research, School of Environmental Sciences, University of East Anglia, Norwich, UK.

²Graduate School of Environmental Earth Science, Hokkaido University, Sapporo, Japan.

³Core Research for Evolutional Science and Technology, Japan Science and Technology Agency, Tokyo, Japan.

Corresponding author: E. T. Buitenhuis, Tyndall Centre for Climate Change Research, School of Environmental Sciences, University of East Anglia, Norwich Research Park, Norwich NR4 7TJ, UK. (<http://tinyurl.com/contacterik>)

©2013. American Geophysical Union. All Rights Reserved.
0886-6236/13/10.1002/gbc.20074

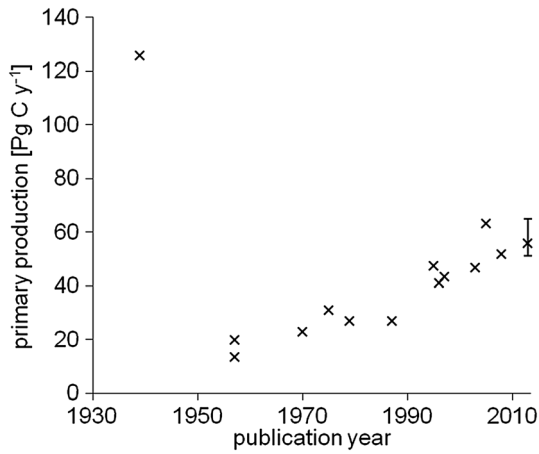


Figure 1. Incomplete record of published estimates of global particulate net primary production. Estimates before 1990 are extrapolations of in situ observations, estimates between 1995 and 2008 are satellite algorithms [Riley, 1939; Steeman Nielsen and Jensen, 1957; Gessner, 1957; Koblenz-Mishke, 1970; Platt and Subba Rao, 1975; Eppley and Peterson, 1979; Berger et al., 1987; Longhurst et al., 1995; Antoine et al., 1996; Behrenfeld and Falkowski, 1997; Mélin, 2003; Behrenfeld et al., 2005; Westberry et al., 2008], and the 2013 estimate is our ocean biogeochemical model (this study: depth-resolved PlankTOM5.3).

spatial and temporal coverage provided by remote sensing data has narrowed down the range of estimated values to 46–65 Pg C yr⁻¹, but a large spread remains, mainly because of the uncertainty in the vertical distribution of NPP. Furthermore, no formal error analysis has been done previously, no central (most likely) estimate exists, and the satellite-based estimates have not ruled out the lower NPP values determined from in situ measurements alone.

[5] Here, we combine model results and observations to constrain global NPP and its uncertainty. We first improve the vertical representation of NPP in the PlankTOM5 global biogeochemical model of Buitenhuis et al. [2010] by implementing and validating the dynamic iron-light photosynthesis formulation of Buitenhuis and Geider [2010]. This mechanistic formulation reproduces the effects of iron and light availability on the C:Chl:Fe stoichiometry of phytoplankton and the dynamical influence of this stoichiometry on photosynthetic performance. It is thus well suited to study the spatiotemporal (including depth) distribution of phytoplankton carbon biomass and their activity in a changing light-Fe environment. We use the model and a range of perturbation experiments to calculate global NPP when the model best fits the observations and estimate its uncertainty. Finally, we verify the results using the NSI-MEM (Nitrogen, Silicon and Iron Regulated Marine Ecosystem Model) ocean biogeochemistry model [Shigemitsu et al., 2012].

2. Model Description

2.1. PlankTOM5 Biogeochemical Model

[6] PlankTOM5.3 is a global ocean biogeochemical model representing five plankton functional types (PFTs): three phytoplankton PFTs (diatoms, coccolithophores, and mixed

phytoplankton) and two zooplankton PFTs (microzooplankton and mesozooplankton). The model has 25 state variables, including 10 phytoplankton components (three PFTs with variable Fe:C, Chl:C, and Si:C ratios and fixed macronutrient:C ratios), 2 zooplankton components, 6 dissolved/nutrient components (dissolved inorganic carbon, alkalinity, oxygen, nitrate (NO₃), silicate (SiO₃), and Fe), and 7 dead particulate/detritus components (small and big particulate organic carbon and iron, dissolved organic matter, and detrital CaCO₃ and SiO₂). PlankTOM5.3 was developed from PlankTOM5.2 of Buitenhuis et al. [2010] and includes (1) the dynamical photosynthesis model of Buitenhuis and Geider [2010], (2) the biophysical feedback through heat absorption by chlorophyll for each phytoplankton plankton functional type (pPFT) of Manizza [2006; see also Manizza et al., 2008], and (3) a new parameterization of the ballasting effect based on the observations of drag coefficient as a function of the Reynolds number from Ploug et al. [2008]. Here, we present only these new equations. Documentation of the other compartments can be found at http://lgmwebweb.env.uea.ac.uk/green_ocean/.

2.1.1. Dynamic Photosynthesis Model

[7] The change in phytoplankton concentration is calculated as

$$\frac{\delta P_i^e}{\delta t} = (1 - d)P_i^e P_i^C - w_{Pi}P_i^C P_i^e - \sum_j g_{Zj}^{Pi} Z_j P_i^e \quad (1)$$

in which P_i^e is the concentration of element e in phytoplankton PFTs i , in which the elements e are C, Fe, or Chl (N, P, and O occur in fixed ratios to C). d is the fraction of primary production that is exuded as dissolved organic carbon (DOC), P_i^C is the assimilation rate of phytoplankton (equations (2), (4), and (5)), w_{Pi} is the total loss rate (equation (6)), and g_{Zj}^{Pi} is the grazing rate of zooplankton Z_j on phytoplankton P_i [Buitenhuis et al., 2006, 2010]. The loss rates are the same for all elements.

[8] The assimilation rates are based on the iron-light colimitation model of Buitenhuis and Geider [2010] as follows: for carbon,

$$P_{Pi}^C P_{C_{max}} \times \left(1 - \exp\left(\frac{-\alpha^{Chl} \theta^C I}{P_{C_{max}}}\right)\right) \quad (2)$$

where $P_{C_{max}}$ is the maximum photosynthesis rate (equation (3)), α^{Chl} is the constant initial slope of the PI (photosynthesis-irradiance) curve, θ^C is the variable chlorophyll-to-carbon ratio, and I is the light intensity (blue plus red) in mol photons m⁻² s⁻¹. Light intensity is converted from W m⁻² to mol photons m⁻² s⁻¹ using a conversion factor of 3.99 for blue light and 5.24 for red light.

[9] PlankTOM5.3 has fixed C:N ratios and variable C:Fe ratios. Fixed ratio models typically use saturation kinetics to model dependence of growth rate on external nutrient concentrations, while variable ratio quota models use a linear dependence of growth rate on internal nutrient quota. Droop [1974] showed that the law of the minimum can be used to calculate the effect of multiple potentially limiting nutrients. Here we assume that limitation by a saturation model and a quota model can be combined in a

single minimum function to represent the maximum photosynthesis rate:

$$P_{\max}^C = P_{\text{ref},0^\circ}^C \times Q_{10}^{\wedge T/10} \cdot \min \left(\frac{Q - Q_{\min}}{Q_{\text{opt}} - Q_{\min}}, \frac{\text{NO}_3}{K_{1/2,\text{NO}_3} + \text{NO}_3}, \frac{\text{SiO}_3}{K_{1/2,\text{SiO}_3} + \text{SiO}_3}, 1 \right) \quad (3)$$

where P_{ref}^C is the maximum assimilation rate at 0°C , Q_{10} is the temperature dependence of growth, T is temperature, Q is the internal phytoplankton Fe:C quota, NO_3 and SiO_3 are the seawater nutrient concentrations, and $K_{1/2}$ are the half saturation concentrations for growth.

[10] The assimilation rate for iron is

$$P_{\text{Pi}}^{\text{Fe}} = \left(\rho_{\max}^{\text{hi}} - (\rho_{\max}^{\text{hi}} - \rho_{\max}^{\text{lo}}) \times \left(\frac{Q - Q_{\min}}{Q_{\max} - Q_{\min}} \right) \right) \cdot Q_{10}^{\wedge T/10} \times \frac{\text{Fe}'}{K_{1/2,\text{Fe}} + \text{Fe}'} \quad (4)$$

where ρ_{\max}^{hi} is the maximum iron uptake rate at iron limitation, $\rho_{\max}^{\text{lo}} (= \mu_{\max} Q_{\max})$ is the steady state iron uptake rate at saturating iron concentrations, Fe' is the dissolved iron concentration, and $K_{1/2,\text{Fe}}$ is the half saturation concentration for iron uptake.

[11] The assimilation rate for chlorophyll is

$$P_{\text{Pi}}^{\text{Chl}} = \left(P_{\max}^C \times \left(1 - \exp \left(\frac{-\alpha^{\text{Chl}} \theta^{\text{C}} I}{P_{\max}^C} \right) \right) \right)^2 \times \frac{\theta_{\max}^{\text{C}}}{-\alpha^{\text{Chl}} \theta^{\text{C}} I} \quad (5)$$

where θ_{\max}^{C} is the maximum chlorophyll-to-carbon ratio.

[12] We simplified the phytoplankton loss rates to a single term that represents respiration, aggregation, and mortality because insufficient data were available to separately represent these processes. Compared to the previous model version [Buitenhuis *et al.*, 2010], the loss term dependent on nutrient limitation was removed, but the quadratic loss term was retained, as follows:

$$w_{\text{Pi}} = \left(L + L_{\text{dia}} \times \left(1 - \min \left(\frac{\text{Fe}'}{K_{1/2,\text{Fe}}}, \frac{\text{NO}_3}{K_{1/2,\text{NO}_3}}, \frac{\text{SiO}_3}{K_{1/2,\text{SiO}_3}}, 1 \right) \right) \right) \quad (6)$$

where L is the loss rate and L_{dia} is the additional loss term for nutrient-limited diatoms.

2.1.2. Biophysical Feedback

[13] The change in temperature through heat absorption by chlorophyll for each phytoplankton plankton functional type (pPFT) was calculated according to Manizza [2006; see also Manizza *et al.*, 2008], as

$$\frac{\delta T}{\delta t} = \frac{\sum (E_z(\lambda) - E_{z+\Delta z}(\lambda))}{\rho C_p} \quad (7)$$

and

$$E_{z+\Delta z}(\lambda) = E_z(\lambda) e^{-\sum (k_w(\lambda) + \sum \chi(\lambda, \text{PFT}) \text{Chl}_{\text{PFT}}) \Delta z} \quad (8)$$

[14] in which T is the temperature, Σ are the sums over the wavelength bands λ (blue, red, and infrared) or pPFTs, E is the light intensity in W m^{-2} , z is depth at the top of each model box, Δz is the depth of each model box, ρ is the density of seawater, C_p is the specific heat of seawater, $k_w(\lambda)$ is the extinction coefficient of pure water, and $\chi(\lambda, \text{PFT})$ is the PFT-specific extinction coefficient of chlorophyll.

2.1.3. Ballasting of Sinking Particulate Organic Carbon by CaCO_3 and SiO_2

[15] We recalculated the ballasting parameters of the model, which increases the sinking speed of big particles as their content of detrital CaCO_3 and SiO_2 increases. All detrital CaCO_3 and SiO_2 end up in big particles. Small organic carbon particles sink at a fixed rate of 3 m d^{-1} . First, the dependence of the drag coefficient (C_d) on the Reynolds number (Re) was fit to the data from Ploug *et al.* [2008, Figure 2], using their measurements on fecal pellets and the data they reproduce from Taghon *et al.* [1984]. This gave $C_d = 51 \text{ Re}^{-0.56}$. Then, the drag equations [Buitenhuis *et al.*, 2001] were solved offline by iteration. This gave pairs of particle density and sinking speed. Lastly, a simple function between particle density and sinking speed was derived for use in the model. With the new data, the function changed from concave to convex (i.e., from changing faster at high densities to changing faster at low densities). We therefore changed the function to

$$v_{\text{sink}} = a \times (\rho - \rho_{\text{sw}})^{\wedge b} \quad (9)$$

in which v_{sink} is the sinking speed of big particles; ρ is the density of the particle, which is calculated from the composition in organic matter, CaCO_3 , and SiO_2 ; and ρ_{sw} is the density of seawater. The densities were calculated from Ploug *et al.* [2008, Table 3], assuming that the organic matter in the three types of fecal pellets had the same density, giving $\rho_{\text{organic}} = 1.08 \text{ kg L}^{-1}$, $\rho_{\text{CaCO}_3} = 1.34 \text{ kg L}^{-1}$, $\rho_{\text{SiO}_2} = 1.2 \text{ kg L}^{-1}$, $a = 0.0303$, and $b = 0.6923$.

2.2. Nucleus for European Modelling of the Ocean (NEMO) Physical Model and Model Forcing

[16] The PlankTOM5.3 biogeochemical model was run online in the ocean general circulation model (OGCM) NEMOv2.3 [Madec, 2008]. We use the ORCA configuration with a horizontal resolution of 2° longitude and on average 1.1° latitude, and a vertical resolution of 10 m in the top 100 m, increasing to 500 m at 5 km depth. The model has a free surface height [Roullet and Madec, 2000]. It is coupled to a dynamic-thermodynamic sea ice model [Timmermann *et al.*, 2005]. The vertical mixing is calculated at all depths using a turbulent kinetic energy model [Gaspar *et al.*, 1990]. Subgrid eddy-induced mixing is parameterized according to Gent and McWilliams [1990].

[17] The standard simulation was run from 1920 to 2009. The model was spun up from 1920 to 1947 using constant atmospheric forcing of daily wind and precipitation from the National Centers for Environmental Prediction reanalysis [Kalnay *et al.*, 1996] repeated every year using values for year 1980, followed by varying atmospheric forcing corresponding to each year from 1948 to 2009. Perturbation experiments were initialized with the output of the standard simulation and were run from 2004 to 2009. Results, unless otherwise stated, are analyzed for year 2009.

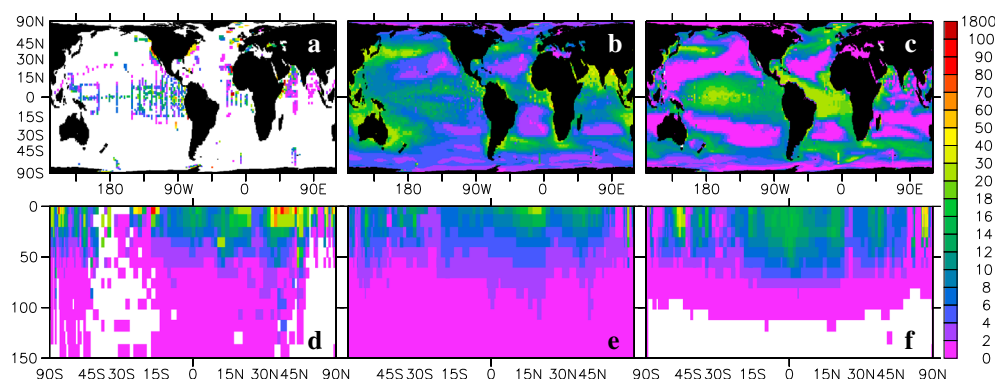


Figure 2. (a) ^{14}C measured surface (0–10 m) primary production ($\mu\text{g C}\cdot(\text{L}\cdot\text{day})^{-1}$). The white area indicates no observations. (b) Surface primary production from PlankTOM5.3 best simulation. Model results are for the same months where there are observations and annual averages everywhere else. (c) Surface primary production NSI-MEM best simulation. (d) ^{14}C measured zonally averaged primary production ($\mu\text{g C}\cdot(\text{L}\cdot\text{day})^{-1}$). (e) Zonally averaged primary production from PlankTOM5.3 best simulation. Model results are for the same longitudes and months where there are observations and zonal annual averages everywhere else. (f) Zonally averaged primary production NSI-MEM best simulation. The white area indicates negative NPP.

3. Materials and Methods

3.1. Database and Treatment of Observations

[18] We synthesized observations of ^{14}C NPP in the ocean (Table S1 in the supporting information, Figures 2a and 2d). The average relative error of the measurements is 19%, decreasing to 12% for values over $10 \mu\text{g C L}^{-1} \text{d}^{-1}$. The error was calculated from the replicated, nonzero California Cooperative Oceanic Fisheries Investigations data only ($n = 8094$).

[19] The NPP database contains 50,050 data points. The data are available at http://lgmacweb.env.uea.ac.uk/green_ocean/biogeochemdata.html. The data were binned on the model grid (2° in longitude, 1.1° average in latitude, 31 vertical levels, and a monthly climatological year), leaving 22,017 grid points. From a simple average of the observations alone, about half of NPP takes place in the top 20 m and 90% in the top 100 m (Figure 2d). The database contains more data in the low latitudes (Figure 2a), with 69% of the gridded data points in the tropics (23°S – 23°N), which make up 42% of the surface area of the ocean. There are no observations in the south of 25°S in winter. We also computed the vertically integrated NPP containing 7509 data points. The vertical integration greatly reduces the number of data, leaving 3132 data points on the model grid, or almost 7 times less than the depth-resolved database.

[20] The ^{14}C technique measures only particulate NPP. We therefore compared the observations to the particulate NPP in the model. In addition, the model phytoplankton produce 5% of primary production as DOC (equation (1), Table S2), based on the nutrient sufficient data compiled by Nagata [2000]. Validating this dissolved NPP falls outside the scope of this paper.

[21] In addition to NPP, the model was evaluated using the World Ocean Atlas 2005 data set of in situ chlorophyll (Figure 3d), which contains 104,689 data points on the model grid. We used World Ocean Atlas (WOA) chlorophyll (Chl) because it is depth resolved and therefore helps to evaluate our depth-resolved model. Sea-viewing Wide Field-of-view Sensor (SeaWiFS) chlorophyll is used

for graphical comparison only, because its horizontal coverage is more complete. In fact, coverage is so good that between $\sim 50^\circ\text{S}$ and $\sim 50^\circ\text{N}$, we also calculate interannual variability of surface Chl. Mixed layer depth (Figure 3d) was calculated from World Ocean Atlas 2005 temperature and salinity, using a density criterion of 0.03 kg L^{-1} [de Boyer Montégut et al., 2004].

3.2. Parameterization of the PlankTOM5.3 Biogeochemical Model

[22] The maximum assimilation rate was calculated as the maximum growth rate divided by the fraction of primary production that is particulate (i.e., not DOC):

$$P_{\text{ref},0^\circ}^{\text{C}} = \mu_{\text{max},0^\circ} / (1 - d) \quad (10)$$

[23] The maximum growth rate parameters ($\mu_{\text{max},0^\circ}$, Q_{10}) were fit to laboratory data (Table S2, Figure 4). For diatoms and mixed phytoplankton, we used the data from the Liverpool Phytoplankton Database [Bissinger et al., 2008]. This database did not include any measurements on coccolithophores. Therefore, for coccolithophores, we used the data of Buitenhuis et al. [2008].

[24] The light limitation parameters (α^{Chl} , $\theta_{\text{max}}^{\text{C}}$) were taken as the average for the three pPFTs from Geider et al. [1997]. The iron limitation parameters ($\rho_{\text{max}}^{\text{hi}}/\rho_{\text{max}}^{\text{lo}}$, $K_{1/2}$, Fe , Q_{min} , Q_{opt} , Q_{max}) for diatoms and coccolithophores were taken from the data for *Thalassiosira oceanica* and *Emiliania huxleyi* in Buitenhuis and Geider [2010]. The iron limitation parameters for mixed phytoplankton were based on the *Pelagomonas calceolata* data in Sunda and Huntsman [1995].

3.3. Tuning of the Standard Simulation of PlankTOM5.3

[25] The above selection of parameters without prior model adjustments led to a strong overestimate in coccolithophores, accounting for $\sim 60\%$ of the phytoplankton biomass, an unrealistically high figure [Le Quéré et al., 2005]. As a consequence,

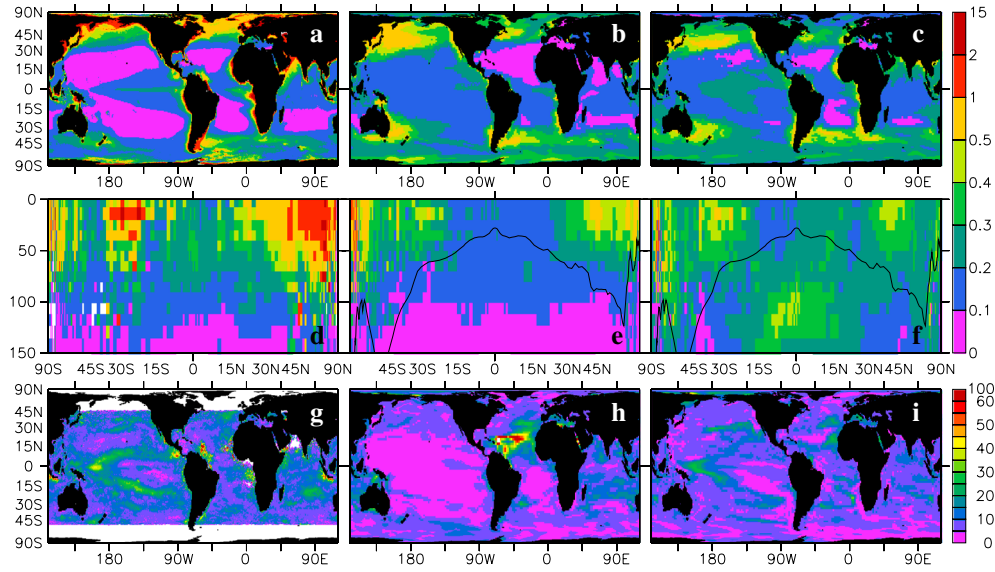


Figure 3. (a) Surface chlorophyll *a* ($\mu\text{g L}^{-1}$, SeaWiFS). (b) Surface chlorophyll *a* (PlankTOM5.2). (c) Surface chlorophyll *a* (PlankTOM5.3 standard simulation). (d) Chlorophyll *a* ($\mu\text{g L}^{-1}$, WOA 2005) and mixed layer depth (black line), [de Boyer Montégut et al., 2004]. (e) Chlorophyll *a* and mixed layer depth (PlankTOM5.2). Model results are for the same longitudes and months as the observations and zonal annual averages everywhere else. (f) Chlorophyll *a* and mixed layer depth (PlankTOM5.3). (g) Relative standard deviation (%) of surface chlorophyll *a* (SeaWiFS). (h) Relative standard deviation of surface chlorophyll *a* (PlankTOM5.2). (i) Relative standard deviation of surface chlorophyll *a* (PlankTOM5.3).

CaCO_3 export at 3.4 Pg C yr^{-1} was also unrealistically high [Lee, 2001]. We adjusted coccolithophore model parameters to obtain a realistic ecosystem distribution. We decreased the competitiveness of coccolithophores by increasing the relative preference of mesozooplankton for coccolithophores to 2.5 times that of mixed phytoplankton, doubling their $K_{1/2, \text{Fe}}$ to 2.6 nmol L^{-1} relative to the observed value for *E. huxleyi* [Buitenhuis and Geider, 2010], and increasing their $K_{1/2, \text{NO}_3}$ to $0.4 \mu\text{mol L}^{-1}$. Rather than change one parameter a lot, we chose to change these three parameters in an effort to keep them within the (poorly constrained) range of observed/realistic values. This gave a realistic coccolithophore biomass of 25% of the phytoplankton and a more realistic CaCO_3 export of 1.6 Pg C yr^{-1} . As in previous versions of the model, we tuned the particle degradation rate to get a realistic air-sea CO_2 flux of 2.1 Pg C yr^{-1} in the 1990s, within 5% of Denman et al. [2007].

3.4. Model Evaluation and Statistics

[26] As in Buitenhuis et al. [2010], we evaluate the model using the following cost function:

$$\text{cost function} = 10^{\text{average}(|10 \log(\text{model}/\text{observation})|)} \quad (11)$$

[27] This cost function gives the same penalty when the model is half the observed value or when the model is twice the observed value. It is also a relative measure of error and thus gives the same penalty for small and large values. We also use the square root of the average residual sum of squares:

$$\text{RSS}^{0.5} = \left(\left(\sum (\text{model} - \text{observation})^2 \right) / n \right)^{0.5} \quad (12)$$

[28] This formulation is dominated by errors in the large values. It has the same units as the observations. For both these evaluations, the observations were binned onto the model grid.

[29] We calculated the 95% confidence intervals of NPP from the ratio of two residual sum of squares (RSS) values, using the following formula [Abramowitz and Stegun, 1972]:

$$\frac{\text{RSS}}{\text{RSS}_{\min}} = 1.645 \times \frac{n}{n-2} \times \sqrt{\left(\frac{2(2n-2)}{n(n-4)} \right) + \frac{n}{n-2}} \quad (13)$$

where RSS_{\min} is the value for the model simulation that best fit the observations, RSS are the values for the model simulations that are inside or just outside the confidence

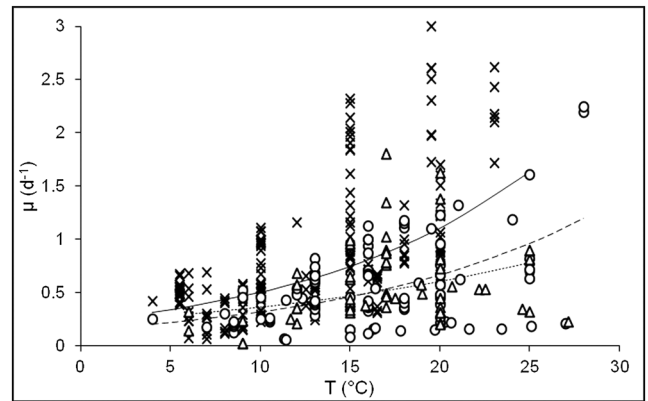


Figure 4. Phytoplankton maximum growth rates (d^{-1}) as a function of temperature ($^{\circ}\text{C}$). Crosses, solid line: diatoms [Bissinger et al., 2008]. Triangles, short dashed line: coccolithophores [Buitenhuis et al., 2008]. Circles, long dashed line: mixed phytoplankton [Bissinger et al., 2008].

interval, 1.645 is the F distribution value for $p = 0.05$, and n is the degrees of freedom. An F distribution is appropriate for the ratio of two χ^2 distributions, such as squared residuals [Berry and Lindgren, 1990]. We approximated n with the number of observations (binned by month on the model grid) because most model parameters were constrained by additional observations and were therefore not free parameters. In addition, the number of parameters in the biogeochemical model (~ 100) is small relative to the number of observations (20,394 for NPP and 104,689 for Chl). This procedure for calculating confidence intervals assumes that the model simulations are independent and that the average residual of the simulation with RSS_{\min} is 0. It is clear that the model simulations are not independent, and the average residual of the simulation with RSS_{\min} is -13% relative to the observations. Nevertheless, the model results that best match the observations and their confidence intervals should be robustly estimated given that we used very large sample sizes [Donaldson, 1968]. It should be noted that the confidence intervals only account for the uncertainty arising from the model-observation mismatch, including the measurement error in the observations, but not from any biases in the observations, the most likely of which are spatial bias and the possible underestimation of NPP by up to 13% noted in the introduction. To analyze the potential for spatial bias in the observations to introduce a bias in the best estimate of global NPP, we repeat the analysis by regions.

3.5. NSI-MEM Biogeochemical Model

[30] To check the robustness of our statistical methodology and results, we also use the NSI-MEM model [Shigemitsu *et al.*, 2012]. NSI-MEM is based on the NEMURO (North Pacific Ecosystem Model for Understanding Regional Oceanography) model [Kishi *et al.*, 2007] and includes an iron cycle [Shigemitsu *et al.*, 2012]. The model was developed in a 1-D setting, but here we run it in the OGCM COCO (CCSR Ocean Component Model) [Aita *et al.*, 2007]. Thus, both the biogeochemical and physical models are different from PlankTOM5-NEMO. The parameters used by Shigemitsu *et al.* [2012] were tuned for application in the western North Pacific. Here, we tuned the parameters for application in the global ocean as given in Table S3.

3.6. Perturbation Experiments

[31] The statistical evaluation requires a number of model simulations that span the expected ranges of global NPP. These ranges were obtained by performing perturbation experiments of PFT turnover rates.

[32] In PlankTOM5.3, PFT turnover rates were increased to match the upper 99% of observed rates, following the concept introduced by Eppley [1972] that the phytoplankton with the highest growth rates will outgrow the others and thus are most representative of the population growth rate. Combining these high growth rates with observed resource efficiencies would create a model organism that could not exist in nature [cf. Litchman *et al.*, 2007]. Therefore, in the standard simulation, we have used growth rates that were fit to all data. We used the 99% approach in the perturbation experiments, not only for phytoplankton, but for zooplankton rates as well. We increased the turnover rate until 99% of the data (or at least two points for coccolithophore growth rate) were under the curve at a constant Q_{10} . Note that by keeping

the Q_{10} constant, it becomes easier to interpret the results than if we had increased both the turnover rate and the Q_{10} . The following turnover rates were increased: phytoplankton growth rate, zooplankton grazing rates, zooplankton respiration and mortality, and organic particulate detritus degradation rates. The latter was insufficiently constrained by observations, so we increased it by factors 1.5, 2, and 3. This resulted in a wide range of NPP rates for the statistical evaluation. For the regional analysis of NPP, a wider range of perturbations was needed. Parameter values are given in Table S4.

[33] In NSI-MEM, perturbation experiments used the same general approach as for PlankTOM5.3, increasing phytoplankton maximum nutrient uptake rate and particulate organic matter degradation rates. Because of the different model structure, we do not use the same observational data to calculate the parameters in the perturbation experiments but use simple scaling factors (Table S5).

[34] About 3 times more perturbation experiments were performed with both models than are reported here. The experiments cover changes in zooplankton and remineralization parameters and all the phytoplankton parameters in equations (1)–(5). We report only the simulations that showed the lowest RSS at each global NPP. This ensures that the reported RSS_{\min} is the best constrained by the observational database and that the reported 95% confidence interval covered was as good a sample of the total parameter space as we could achieve and thus at its widest.

4. Results

4.1. Evaluation of New Model Components

[35] We evaluated the PlankTOM5.3 model against 4-D fields (latitude, longitude, depth, and month) of NPP, chlorophyll concentration, phytoplankton growth rate, microzooplankton-caused phytoplankton mortality rate, microzooplankton concentration, 2-D fields (latitude, longitude) of mesozooplankton concentration, export at 100 m, and the global rate of mesozooplankton grazing on phytoplankton. Data sources for observations are given in Table 1. Compared to the PlankTOM5.2 model [Buitenhuis *et al.*, 2010], PlankTOM5.3 shows improvements in all cost functions relative to the observational databases (Table 2) and also improved average rates and concentrations for NPP, average chlorophyll and microzooplankton concentration, microzooplankton-caused phytoplankton mortality rate, and global mesozooplankton grazing rates, the same average phytoplankton growth rate, but deteriorated average mesozooplankton concentration and export (Table 1).

[36] The chlorophyll concentration in PlankTOM5.3 has clearly improved relative to PlankTOM5.2 (Tables 1 and 2; Figures 3a–3c). We performed sensitivity simulations without the biophysical feedback or the new ballasting formulation (see sections 2.1.2 and 2.1.3). These simulations showed virtually the same improvements in chlorophyll concentration, NPP, and other cost functions (Table 2), showing that the improvements are due to the new photosynthesis model. The model now includes a clear deep chlorophyll maximum in the tropics with no deep productivity maximum (Figure 2), as observed. However, the deep chlorophyll maximum is too deep, probably due to the fact that the physical model produces an upper mixed layer that is too deep (Figures 3d and 3f).

Table 1. Average Concentrations and Globally Integrated Rates in Perturbation Tests with PlankTOM5.3

Parameters		PP ^{a,b} μg C·(L·d) ⁻¹	Chl ^{b,c} μg·L ⁻¹	Phygro ^{b,d} d ⁻¹	Phymor ^{b,d} d ⁻¹	Microz ^{b,e} μg C·L ⁻¹	Mesoz ^{b,f} μg C·L ⁻¹	Mesphy ^g Pg C·yr ⁻¹	Export ^h Pg C·yr ⁻¹
Observation		7.0	0.22	0.63	0.43	2.8	7.1	5.5	9.6
PlankTOM5.2	<i>Buitenhuis et al.</i> [2010]	4.8	0.12	0.32	0.10	1.7	2.1	14.7	9.1
Standard	Table S2	5.3	0.17	0.32	0.16	2.5	1.2	10.4	8.4
No Chl heating		5.3	0.17	0.32	0.17	2.4	1.2	10.4	8.4
Old ballast		5.7	0.17	0.31	0.16	2.4	1.4	14.3	7.2
Perturbation	99% phyto growth ⁱ	5.5	0.18	0.35	0.18	3.4	1.6	13.4	11.23
	99% micro growth ^j	2.4	0.10	0.36	0.25	2.3	0.4	3.7	3.5
	intermediate micro growth ^j	3.7	0.12	0.36	0.24	2.7	0.7	6.0	5.3
	ⁱ +99% zoo growth ^k	0.9	0.01	0.68	0.00	0.1	0.5	6.9	2.6
	^k +99% zoo loss	2.2	0.05	0.57	0.00	0.1	0.6	17.2	5.2
	ⁱ +1.5×	6.1	0.19	0.38	0.14	3.6	1.8	15.8	10.4
	part degr ^m ⁱ +2×	6.7	0.20	0.41	0.16	4.0	1.9	18	9.8
	part degr ^m ⁱ +3×	7.9	0.22	0.45	0.19	4.3	2.1	22.1	8.7
	part degr ^m ⁱ +4×	9.1	0.23	0.49	0.24	4.7	2.3	25.1	7.8
	part degr ^m ⁱ +5×	10.1	0.24	0.52	0.26	4.9	2.5	28.2	7.0
	part degr ^m								

^aParticulate net primary production.

^bThe model was sampled where evaluation data were available.

^cChlorophyll, World Ocean Atlas 2005.

^dPhytoplankton growth rate and microzooplankton-caused phytoplankton mortality rate from dilution experiments [Buitenhuis et al., 2010].

^eMicrozooplankton biomass [Buitenhuis et al., 2010].

^fMesozooplankton biomass [Buitenhuis et al., 2006].

^gMesozooplankton grazing on phytoplankton [Calbet, 2001].

^hExport at 100 m [Schlitzer, 2004].

^{i,j}All other parameters as in standard simulation.

^{k,m}Include the changes made in ⁱ.

^lIncludes the changes made in ⁱ and ^k.

[37] There is a marked improvement in the interannual variability of chlorophyll in PlankTOM5.3 compared to PlankTOM5.2 (Figures 3g–3i). Comparison of the relative interannual variability in Chl:C, phytoplankton C, and NPP shows that the former is low, while the latter two closely resemble the interannual variability in Chl in the respective model versions (data not shown). This suggests that the improvement in the model is not primarily due to a better representation of phytoplankton chlorophyll content but that the signal originates in a better representation of NPP and propagates into phytoplankton carbon. This interpretation is consistent with the fact that PlankTOM5.2 already included variable Chl:C, while PlankTOM5.3 has added the effect of the Chl:C ratio on NPP. It thus suggests that the new formulation, which is more directly based on observations for its equations and parameters, better captures the sensitivity of (carbon) NPP to variability in the atmospheric forcing. This will be an important improvement when exploring the feedbacks between climate change and global biogeochemical cycles in future studies.

[38] As noted above, the cost function of export at 100 m has improved relative to the inversion results of Schlitzer [2004]. Since ocean biogeochemistry, in particular air-sea CO₂ flux, is very sensitive to export, this result warrants further analysis. We compared the global average nutrient profile of PlankTOM5.2, which has slightly higher export, and the standard simulation of PlankTOM5.3 to the World Ocean Atlas observations (Figure 5). Both model runs were initialized with these observations, and the standard simulation has remained closer to the observations,

even though PlankTOM5.3 was run longer (90 years) than PlankTOM5.2 (60 years). As a consequence, the cost function has decreased from 1.5 to 1.3. This would only be an indication that the model has improved if the ocean had been in steady state over that time, which we know it was not [Le Quéré et al., 2007]. However, the interannual variability of chlorophyll has also improved considerably in the standard simulation (Figures 3g–3i), which is sensitive to the nutrient concentration gradient across the permanent thermocline. This gradient is sharper and thus closer to the observations in the standard simulation (Figure 5c). We therefore conclude that the lower export constitutes a real improvement in the model.

[39] The supporting information presents additional graphs comparing observations and model for microzooplankton concentration, microzooplankton-caused phytoplankton mortality, phytoplankton growth rate, and mesozooplankton and phosphate concentration and export (Figure S1).

4.2. Global and Regional Estimates of NPP

[40] The previous model version, PlankTOM5.2, already was the ocean biogeochemical model with the best fit to the observations of NPP at Bermuda Atlantic Timeseries Station (BATS) and Hawaii Ocean Timeseries (HOT) out of 12 models tested [Saba et al., 2010]. Here, we make a further slight improvement in the cost function (Table 2) and the depth-resolved and vertically integrated RSS^{0.5} (Figures 6a and 6b) of PlankTOM5.3 relative to the observational database of NPP (equations (11) and (12)).

Table 2. Cost Functions in Perturbation Tests with PlankTOM5.3 (Equation (11))

	Parameter Values	PP	Chl	Phygro	Phymor	Microz	Mesoz	Export
PlankTOM5.2	<i>Buitenhuis et al.</i> [2010]	3.5	5.1	2.5	41	2.8	6.7	2.1
Standard	Table S2	3.3	4.3	2.3	15	2.6	6.4	2.0
No Chl heating		3.3	4.3	2.3	15	2.6	6.4	2.0
Old ballast		3.3	4.3	2.3	15	2.6	6.1	2.0
Perturbation	99% phyto growth ⁱ	3.2	4.3	2.4	24	2.7	5.7	2.0
	99% micro growth ⁱ	3.4	5.3	2.1	3.0	2.2	16.1	3.6
	intermediate micro growth ⁱ	3.2	4.6	2.1	4.5	2.4	9.7	2.5
	ⁱ +99% zoo growth ^j	7.1	13.5	1.9	3134	3.6	8.0	4.3
	^j +99% zoo loss ^k	4.3	6.3	2.1	2975	3.6	9.0	2.8
	ⁱ +1.5×	3.3	4.3	2.3	22	2.7	5.2	2.0
	part degr ^l							
	ⁱ +2×	3.4	4.3	2.2	16	2.6	4.9	2.0
	part degr ^l							
	ⁱ +3×	3.6	4.4	2.2	12	2.6	4.5	2.0
	part degr ^l							
	ⁱ +4×	3.6	4.4	2.1	8.6	2.5	4.3	2.0
	part degr ^l							
	ⁱ +5×	3.8	4.5	2.0	8.5	2.5	4.1	2.0
	part degr ^l							

^{i,j,k,l}As in Table 1.

[41] The perturbation experiments with both models cover a range of NPP from ~ 10 to ~ 85 Pg C yr⁻¹ (parameter values in Tables S4 and S5). The simulations have clear increases of the RSS^{0.5} at the high and low ends which are used to estimate the 95% confidence intervals. The global modeled ocean NPP most consistent with the depth-resolved ¹⁴C observations is 56 Pg C yr⁻¹ with a 95% confidence interval of 51–65 Pg C yr⁻¹. Perturbation experiments with the NSI-MEM model suggest the same global NPP of 56 Pg C yr⁻¹, but with a wider 95% confidence interval of 42–68 Pg C yr⁻¹ (Figure 6a).

[42] The global modeled ocean NPP most consistent with the vertically integrated ¹⁴C observations is 51 Pg C yr⁻¹ (95% confidence interval 37–63 Pg C yr⁻¹) for PlankTOM5.3 and 44 Pg C yr⁻¹ (95% confidence interval 23–53 Pg C yr⁻¹) for NSI-MEM (Figure 6b). The vertically integrated estimate for PlankTOM5.3 agrees well with the depth-resolved estimate. However, the 95% confidence interval is almost twice as wide because of the smaller number of data points. Because NSI-MEM overestimates NPP in the tropics (Figure 2c) where there are about twice as many observations as in the rest of the ocean

(Figure 2a), the lower estimate of vertically integrated NPP in NSI-MEM can be discounted as biased.

[43] We also used the perturbation experiments to calculate the best estimates of NPP in 10 regions: 9 open ocean regions in the Pacific, Atlantic, and Indian Ocean basins divided into tropics (23°S–23°N) and the extratropics in both hemispheres, and the coastal (<200 m depth) ocean (Table 3). We also estimated the 95% confidence intervals for each region with the PlankTOM5.3 model. The confidence intervals from the NSI-MEM model for some of the regions were not sufficiently constrained by the perturbation experiments. In both models, the largest contributions to global NPP are made by the Pacific Ocean and the tropical region, both about 50%, and roughly a quarter by the Atlantic and Indian Oceans and also by the southern and northern extratropics, while the coastal ocean contributes 8% in PlankTOM5.3 and 9% in NSI-MEM. Relative to the region areas, both models agree that the largest contribution is made by the coastal ocean, 77% larger than expected from its area in PlankTOM5.3 and 98% larger in NSI-MEM, followed by the North Pacific, 54% and 40% larger than expected from

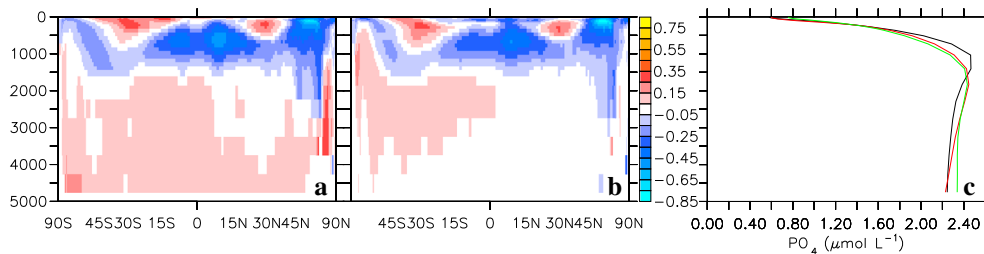


Figure 5. (a) Zonal average difference between PlankTOM5.2 macronutrient concentration in 2007 and World Ocean Atlas 2005 PO₄ concentration (μmol L⁻¹, depth in meters). (b) Zonal average difference between PlankTOM5.3 macronutrient concentration in 2007 and World Ocean Atlas 2005 PO₄ concentration. (c) Area weighted average nutrient concentration. Black line: World Ocean Atlas 2005 PO₄ concentration. Also, model macronutrient concentration initial conditions. Green line: PlankTOM5.2 macronutrient concentration in 2007. Red line: PlankTOM5.3 standard simulation macronutrient concentration in 2007.

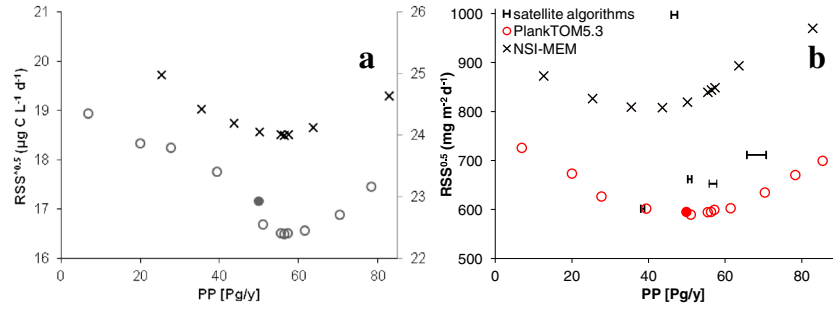


Figure 6. $\text{RSS}^{0.5}$ for different models. Filled symbols: standard simulation. Open symbols: perturbation simulations. (a) 3-D primary production. Left y axis: PlankTOM. Right y axis: NSI-MEM. (b) 2-D integrated primary production. Symbols without error bars: biogeochemical model simulations. Error bars without symbols: ranges from satellite-based algorithms (see text for details). These are, from low to high productivity, *Antoine et al.* [1996], *Behrenfeld and Falkowski* [1997], *Mélin* [2003], *Westberry et al.* [2008], and *Behrenfeld et al.* [2005]. Circles: PlankTOM5.3. Crosses: NSI-MEM.

its area, and the equatorial Indian Ocean, 26% and 49% larger than expected from its area, and also that the smallest relative contribution is made by the South Atlantic, 37% and 67% smaller than expected from its area. The South Indian and South Atlantic are least constrained by the observations, 87% and 84% fewer observations than expected from their areas. The models disagree about the absolute amount of NPP in the South Pacific and South Indian, with lower NPP in NSI-MEM in the South Pacific and higher in the South Indian, and the NPP in the equatorial Pacific is also slightly lower than the PlankTOM5.3 95% confidence interval.

[44] We compare our results to four satellite-based algorithms of NPP. Although satellite algorithms are sometimes treated as if they were observational data by ocean biogeochemical modelers, these algorithms are in fact themselves models with considerable uncertainties associated with their equations, parameters, and input data choices [*Friedrichs et al.*, 2009; *Saba et al.*, 2010]. Satellites do not cover the whole ocean, mostly because of cloud cover. The results with PlankTOM5.3 suggest that for the SeaWiFS satellite, NPP at grid cells where chlorophyll data are not available from the SeaWiFS satellite is about half of the average NPP over the SeaWiFS-covered part of the ocean. Therefore, the range of global NPP for each satellite algorithm was estimated between a low estimate assuming zero NPP when no chlorophyll data were available and a

high estimate assuming an average NPP (Figure 6b). Our calculated global NPP rates for the satellite algorithms are slightly different from the original publications because we calculated the results on our model grid. For *Antoine et al.* [1996] we calculate $38.0\text{--}39.0 \text{ Pg C yr}^{-1}$, for *Behrenfeld and Falkowski* [1997] $46.0\text{--}47.6 \text{ Pg C yr}^{-1}$, for *Mélin* [2003] $50.3\text{--}51.3 \text{ Pg C yr}^{-1}$, and for *Behrenfeld et al.* [2005] $65.7\text{--}70.7 \text{ Pg C yr}^{-1}$. For the *Behrenfeld and Falkowski* [1997] algorithm, our high estimate matches the corrected estimate from this algorithm in *Field et al.* [1998]. The $\text{RSS}^{0.5}$ for the *Behrenfeld and Falkowski* [1997] algorithm is comparatively high. This result is dominated by overestimations of more than $5000 \text{ mg C m}^{-2} \text{ d}^{-1}$ in only 23 coastal sites (<1% of the observations). The *Antoine et al.* [1996] algorithm falls below both of the depth-resolved 95% confidence intervals, and the *Behrenfeld and Falkowski* [1997] algorithm falls below the PlankTOM5.3 depth-resolved confidence interval. The *Behrenfeld et al.* [2005] algorithm falls just above the PlankTOM5.3 depth-resolved confidence interval, only partially overlaps with the NSI-MEM depth-resolved confidence interval, and is well above both vertically integrated confidence intervals. The *Mélin* [2003] and *Westberry et al.* [2008] algorithms are closest to our best estimates of depth-resolved global NPP and fall within the PlankTOM5.3 depth-resolved confidence interval. Our results confirm, on

Table 3. Regional Analysis of Primary Production (Pg C/yr)

Open Ocean PlankTOM5.3	Pacific	Atlantic	Indian	Latitude Total	Latitude Total NSI-MEM
North	8.7 (5.4–9.2) n = 1,867	3.8 (3.3–6.0) n = 1,306	0.10 (0 ^a –0.12) n = 22	12.6 (9.3–14.9)	11.9 (6.1–NC ^b)
Tropics (23°S–23°N)	11.9 (9.9–15.6) n = 10,946	5.2 (2.1–7.6) n = 1,168	7.5 (7.0–8.3) n = 1,896	24.7 (20.9–29.2)	24.2 (15.9–31.4)
South	10.3 (9.0–11.1) n = 1,885	3.1 (0 ^a –4.0) n = 293	2.5 (1.6–3.4) n = 248	15.9 (12.4–17.4)	9.2 (5.6–11.7)
Coast				4.5 (3.1–4.9) n = 763	4.3 (0.4–NC ^b)
Total	30.9 (27.7–31.4)	12.2 (7.7–15.5)	10.1 (9.1–11.3)	57.7 (51.4–63.0)	49.6 (38.1–NC ^b)

^aThe perturbation simulations did not constrain the lower 95% confidence intervals, but a priori knowledge says they must be ≥ 0 .

^bNC: not constrained by the simulations.

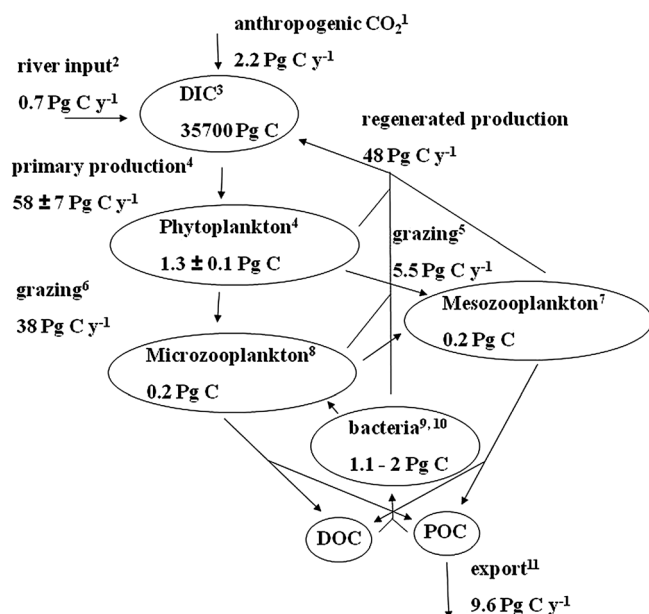


Figure 7. Fluxes and standing stocks in the global marine ecosystem from observations and model extrapolated or inverted observations. ¹Denman *et al.* [2007], ²da Cunha *et al.* [2007], ³GLODAP gridded data with Arctic bottle data as in Buitenhuis *et al.* [2010], ⁴this study, ⁵Calbet [2001], ⁶PP¹ × microzooplankton-caused phytoplankton mortality/phytoplankton growth (Table 1), ⁷Buitenhuis *et al.* [2006], ⁸Buitenhuis *et al.* [2010], ⁹Buitenhuis *et al.* [2012], ¹⁰Whitman *et al.* [1998], and ¹¹Schlitzer [2004].

a global scale, the analysis of Saba *et al.* [2010] at BATS and HOT that “the difference in overall skill between the best BOGCM and the best ocean color model at each site was not significant” (Figure 6b).

[45] Our analysis suggests that the Behrenfeld and Falkowski [1997] model underestimates global NPP, whereas Milutinovic and Bertino [2011] suggest it has a positive bias. This discrepancy occurs despite the fact that Milutinovic and Bertino [2011] exclude the coastal ocean, which we find to be overestimated and the biggest contributor to RSS^{0.5} in the Behrenfeld and Falkowski [1997] model. The reason for this discrepancy is due to methodology: Milutinovic and Bertino [2011] assume that the Behrenfeld and Falkowski [1997] model is flawless and that the uncertainty of the output stems from the uncertainty in the input data, whereas we directly assign a cost function to the output using observed local ¹⁴C primary production. We make no a priori assumption that our standard simulation is the best model (it is not) but use the observations to estimate both the most likely global NPP and its confidence interval.

5. Discussion

[46] By combining state-of-the-art models of ocean biogeochemistry and physics, we are able to analyze a model that is depth resolved and has a physically realistic, mechanistic representation of photosynthesis, light extinction, temperature distribution, and nutrient transport. The inclusion of a new photosynthesis model in PlankTOM5.3 has improved the representation of NPP and all the other cost functions of

biogeochemical concentrations and rates (Table 2). The interannual variability of the detrended NPP in the standard simulation is 0.4 Pg C yr⁻¹, or a peak-to-peak variability of 1.7 Pg C yr⁻¹. Thus, interannual variability is small relative to the 95% confidence interval of the year 2009 model results.

[47] The two main ways to narrow the confidence interval are to improve the model (decrease RSS_{min}, equation (13)) and to increase the number of observations. Analysis of the spatial distribution of residuals between NSI-MEM and the observations suggests that the 95% confidence interval is larger than in PlankTOM5.3 because in the simulation that fits best to the observations (RSS_{min}, equation (13)), there is a trade-off between some regions being overestimated and some regions being underestimated. Because of this, perturbation experiments lead to improvements in some regions, resulting in a slow increase in the RSS away from RSS_{min} and thus a large confidence interval. To some extent, this trade-off is probably an inevitable product both of using point observations with measurement errors to evaluate relatively large regions in space and time, and of inadequacies in the models. However, the results are consistent with the expectation that a better model would suffer less from such a trade-off between localized improvements and deteriorations, and would therefore have a faster increase in RSS away from RSS_{min} and therefore have a smaller 95% interval, as is found for PlankTOM5.3. In addition to model improvements, the other way in which we have constrained the confidence interval is by effectively increasing the number of observations. The depth-resolved confidence intervals, which are based on 7 times more data, are smaller than the vertically integrated confidence intervals for both models. For the depth-resolved evaluation of PlankTOM5.3, which has the lowest RSS_{min}, the model is now able to provide a confidence interval that is smaller than the range of satellite algorithms published since 1996.

[48] The regional analysis with PlankTOM5.3 is well within the 95% confidence interval of the global analysis with both models. This shows that despite a spatial bias in the number of observations, the model is mechanistically realistic enough that it can produce reliable results in regions where NPP is less well constrained by the observations. Though the regional analysis with NSI-MEM gives a lower global total, the upper 95% confidence interval of this estimate is unconstrained and thus does not contradict these results. More observations in the Southern Ocean could improve the constraints on the estimated NPP and resolve the disagreement between the regional analysis of the two models in the South Pacific and South Indian.

[49] Our analysis shows that implementing the dynamic iron-light colimitation model of Buitenhuis and Geider [2010] leads to a clear improvement of the ability of the PlankTOM5.3 model to reproduce the observations, in particular the interannual variability of chlorophyll. However, the parameterization of the model is based on few data and could probably be improved further by additional constraints on phytoplankton physiological parameters. Improvements in the mixed layer depth in the OGCM might also result in a narrower confidence interval. This scope for improvement can be seen in the depth of the deep chlorophyll maximum, which is too deep in PlankTOM5.3 (Figure 3). The too deep chlorophyll maximum also leads to an overestimation of NPP below 60 m (Figure 2). NPP below 60 m contributes ~20% to

the global total, but only 2% to the RSS, because NPP variability decreases with NPP (data not shown). Therefore, this overestimation has only a small impact on the confidence interval of global NPP.

[50] Figure 7 summarizes the current understanding of the global marine ecosystem carbon fluxes and standing stocks. As has been shown before [Calbet, 2001; Calbet and Landry, 2004], ~80% of NPP is grazed by zooplankton. We are not aware of estimates on how the remaining 20% gets partitioned between phytoplankton mortality, viral loss, and direct export of (aggregating) phytoplankton. After passing through the epipelagic food web, about one sixth of NPP is exported below 100 m, which implies that five sixth of NPP is regenerated production.

[51] In conclusion, our results with vertically integrated NPP support the wide range of previously reported global estimates, but it provides a quantitative measure of error and a most likely estimate. The depth-resolved evaluations of model NPP against observations contain more information than the vertically integrated evaluation and therefore give a better constrained estimate. Both models agree that our best estimate of global ocean particulate NPP from the depth-resolved evaluations is 56 Pg C yr^{-1} , while the 95% confidence interval from the better constrained PlankTOM5.3 model is $51\text{--}65$ or $58 \pm 7 \text{ Pg C yr}^{-1}$ (Figure 6a).

[52] **Acknowledgments.** We thank the many scientists and institutions that made, compiled, reanalysed, and made publicly available all the measurements that were used to initialize, force, and evaluate the model, in particular the Ocean Primary Production Working Group (OPPWG), Jan Bissinger, and Richard Geider. We thank Clare Enright for programming support, Gareth Janacek for help with the statistics, the Laboratoire d'Océanographie et du Climat for making the code of the NEMO model available, and the Research Computing Service at the University of East Anglia for use of the Linux cluster. We thank two anonymous reviewers and Toby Westberry for their helpful reviews. We thank the Laboratoire d'Océanologie de Villefranche-sur-mer and NERC (contracts NE/C516079/1 and NE/G006725/1) for financial support.

References

- Abramowitz, M., and I. A. Stegun (1972), *Handbook of Mathematical Functions*, National Bureau of Standards Applied Mathematics Series 55, 1046 pp., National Bureau of Standards, Washington, D.C.
- Aita, M. N., Y. Yamanaka, and M. J. Kishi (2007), Interdecadal variation of the lower trophic ecosystem in the North Pacific between 1948 and 2002, in a 3-D implementation of the NEMO model, *Ecol. Modeling.*, 202(1–2), 81–94, doi:10.1016/j.ecolmodel.2006.07.045.
- Antoine, D., J.-M. Andre, and A. Morel (1996), Oceanic primary production 2. Estimation at global scale from satellite (coastal zone color scanner) chlorophyll, *Global Biogeochem. Cycles*, 10, 57–69.
- Behrenfeld, M. J., and P. G. Falkowski (1997), Photosynthetic rates derived from satellite-based chlorophyll concentration, *Limnol. Oceanogr.*, 42, 1–20.
- Behrenfeld, M. J., E. Boss, D. A. Siegel, and D. M. Shea (2005), Carbon-based ocean productivity and phytoplankton physiology from space, *Global Biogeochem. Cycles*, 19, GB1006, doi:10.1029/2004GB002299.
- Berger, W. H., K. Fischer, C. Lai, and G. Wu (1987), Oceanic primary productivity and organic carbon flux, Part 1, Overview and maps of primary production and export, *Scripps Inst. Oceanogr.*, 87-30, 1–67.
- Berry, D. A., and B. W. Lindgren (1990), *Statistics: Theory and methods*. Thomson Brooks/Cole.
- Bissinger, J. E., D. J. S. Montagnes, J. Sharples, and D. Atkinson (2008), Predicting marine phytoplankton maximum growth rates from temperature: Improving on the Eppley curve using quantile regression, *Limnol. Oceanogr.*, 53, 487–493.
- de Boyer Montégut, C., G. Madec, A. S. Fischer, A. Lazar, and D. Iudicone (2004), Mixed layer depth over the global ocean: An examination of profile data and a profile-based climatology, *J. Geophys. Res.*, 109, C12003, doi:10.1029/2004JC002378.
- Buitenhuis, E. T., and R. J. Geider (2010), A model of phytoplankton acclimation to iron-light colimitation, *Limnol. Oceanogr.*, 55, 714–724.
- Buitenhuis, E. T., P. van der Wal, and H. J. W. de Baar (2001), Blooms of *Emiliania huxleyi* are sinks of atmospheric carbon dioxide: A field and mesocosm study derived simulation, *Global Biogeochem. Cycles*, 15, 577–587.
- Buitenhuis, E. T., C. Le Quéré, O. Aumont, G. Beaugrand, A. Bunker, A. Hirst, T. Ikeda, T. O'Brien, S. Piontkovski, and D. Straile (2006), Biogeochemical fluxes through mesozooplankton, *Global Biogeochem. Cycles*, 20, GB2003, doi:10.1029/2005GB002511.
- Buitenhuis, E. T., T. Pangere, D. Franklin, C. Le Quéré, and G. Malin (2008), Growth rates of six coccolithophorid strains as a function of temperature, *Limnol. Oceanogr.*, 53(3), 1181–1185.
- Buitenhuis, E., R. Rivkin, S. Sailley, and C. Le Quéré (2010), Biogeochemical fluxes through microzooplankton, *Global Biogeochem. Cycles*, 24, GB4015, doi:10.1029/2009GB003601.
- Buitenhuis, E. T., W. K. W. Li, M. W. Lomas, D. M. Karl, M. R. Landry, and S. Jacquet (2012), Picoheterotroph (Bacteria and Archaea) biomass distribution in the global ocean, *Earth Syst. Sci. Data*, 4, 101–106, doi:10.5194/essd-4-101-2012.
- Calbet, A. (2001), Mesozooplankton grazing effect on primary production: A global comparative analysis in marine ecosystems, *Limnol. Oceanogr.*, 46, 1824–1830.
- Calbet, A., and M. R. Landry (2004), Phytoplankton growth, microzooplankton grazing, and carbon cycling in marine systems, *Limnol. Oceanogr.*, 49, 51–57.
- da Cunha, L. C., C. Le Quéré, E. T. Buitenhuis, X. Giraud, and W. Ludwig (2007), Potential impact of changes in river nutrient supply on global ocean biogeochemistry, *Global Biogeochem. Cycles*, 21, GB4007, doi:10.1029/2006GB002718.
- Denman, K. L., et al. (2007), Couplings between changes in the climate system and biogeochemistry, in *Climate Change 2007: The Physical Science Basis*. Contribution of working group I to the fourth assessment report of the intergovernmental panel on climate change, edited by S. Solomon et al., pp. 499–587, Cambridge University Press, Cambridge, U.K.
- Donaldson, T. S. (1968), Robustness of the F-test to errors of both kinds and the correlation between the numerator and denominator of the F-ratio, *J. Am. Statist. Ass.*, 60, 660–676.
- Droop, M. R. (1974), The nutrient status of algal cells in continuous culture, *J. Mar. Biol. Ass. UK*, 54, 825–855.
- Eppley, R. W. (1972), Temperature and phytoplankton growth in the sea, *Fish. Bull.*, 70, 1063–1085.
- Eppley, R. W., and B. J. Peterson (1979), Particulate organic-matter flux and planktonic new production in the deep ocean, *Nature*, 282, 677–680, doi:10.1038/282677a0.
- Field, C. B., M. J. Behrenfeld, J. T. Randerson, and P. Falkowski (1998), Primary production of the biosphere: Integrating terrestrial and oceanic components, *Science*, 281, 237–240.
- Friedrichs, M. A. M., et al. (2009), Assessing the uncertainties of model estimates of primary productivity in the tropical Pacific Ocean, *J. Mar. Sys.*, 76, 113–133.
- Gaspar, P., Y. Gregoris, and J. M. Lefèvre (1990), A simple eddy kinetic energy model for simulations of the oceanic vertical mixing: Tests at station Papa and long-term upper ocean study site, *J. Geophys. Res.*, 95(16), 179–193.
- Geider, R. J., H. L. MacIntyre, and T. M. Kana (1997), Dynamic model of phytoplankton growth and acclimation: Responses of the balanced growth rate and the chlorophyll a: Carbon ratio to light, nutrient-limitation and temperature, *Mar. Ecol. Progr. Ser.*, 148, 187–200.
- Gent, P. R., and J. C. McWilliams (1990), Isopycnal mixing in the ocean circulation models, *J. Phys. Oceanogr.*, 20, 150–155.
- Gessner, F. (1957), *Hydrobotanik*, II. *Stoffhaushalt*, 638 pp., Deutscher Verlag der Wissenschaften, Berlin.
- Kalnay, E., et al. (1996), The NCEP/NCAR 40-year reanalysis project, *Bull. Am. Meteorol. Soc.*, 77(3), 437–471.
- Koblentz-Mishke, O. J. L., V. V. Volcovinsky, and J. G. Kabanova (1970), Plankton primary production of the world ocean, in *Scientific Exploration of the South Pacific*, edited by W. S. Wooster, pp. 83–193, National Academy of Sciences, Washington, D.C.
- Le Quéré, C., et al. (2005), Ecosystem dynamics based on plankton functional types for global ocean biogeochemistry models, *Glob. Chang. Biol.*, 11, 2016–2040, doi:10.1111/j.1365-2486.2005.001004.x.
- Le Quéré, C., et al. (2007), Saturation of the Southern Ocean CO₂ sink due to recent climate change, *Science*, 316, 1735–1738, doi:10.1126/science.1136188.
- Lee, K. (2001), Global net community production estimated from the annual cycle of surface water total dissolved inorganic carbon, *Limnol. Oceanogr.*, 46, 1287–1297.
- Litchman, E., C. A. Klausmeier, O. M. Schofield, and P. G. Falkowski (2007), The role of functional traits and trade-offs in structuring phytoplankton communities: Scaling from cellular to ecosystem level, *Ecol. Lett.*, 10, 1170–1181.
- Longhurst, A., S. Sathyendranath, T. Platt, and C. Caverhill (1995), An estimate of global primary production in the ocean from satellite radiometer data, *J. Plankton Res.*, 17, 1245–1271.

- Madec, G. (2008), NEMO ocean engine. Note du Pole de modélisation, Institut Pierre-Simon Laplace (IPSL), France, No 27, http://www.nemo-ocean.eu/content/download/21612/97924/file/NEMO_book_3_4.pdf.
- Manizza, M. (2006), Modelling phytoplankton-light feedback and its biogeochemical implications, PhD thesis, University of East Anglia. http://lmgmacweb.env.uea.ac.uk/green_ocean/publications/Manizza_PhD_thesis_2006.pdf.
- Manizza, M., C. Le Quéré, A. J. Watson, and E. T. Buitenhuis (2008), Ocean biogeochemical response to phytoplankton-light feedback in a global model, *J. Geophys. Res.*, **113**, C10010, doi:10.1029/2007JC004478.
- Marra, J. (2009), Net and gross productivity: Weighting in with ^{14}C , *Aquat. Microb. Ecol.*, **56**, 123–131.
- Mélin, F. (2003), Potentiel de la télédétection pour l'analyse des propriétés optiques du système océan-atmosphère et application à l'estimation de la photosynthèse phytoplanktonique, PhD thesis Université Toulouse III.
- Milutinovic, S. and L. Bertino (2011), Assessment and propagation of uncertainties in input terms through an ocean-color-based model of primary production, *Rem. Sens. Env.*, **115**, 1906–1917, doi:10.1016/j.rse.2011.03.013.
- Nagata, T. (2000), Production mechanisms of dissolved organic matter, in *Microbial Ecology of the Ocean*, edited by D. L. Kirchmann, Wiley-Liss, New York, USA.
- Platt, T. and J. J. V. Subba Rao (1975), Primary production of marine microphytes, in *Photosynthesis and Productivity in Different Environments*, edited by J. P. Cooper, pp. 249–280, Cambridge University Press, Cambridge, U.K.
- Ploug, H., M. H. Iversen, M. Koski, and E. T. Buitenhuis (2008), Production, oxygen respiration rates, and sinking velocity of copepod fecal pellets: Direct measurements of ballasting by opal and calcite, *Limnol. Oceanogr.*, **53**, 469–476.
- Prentice, I. C., G. D. Farquhar, M. J. R. Fasham, M. L. Goulden, M. Heimann, F. J. Jaramillo, H. S. Khashgi, C. Le Quere, R. J. Scholes, and D. W. R. Wallace (2001), The carbon cycle and atmospheric carbon dioxide, in *Climate Change 2001: The Scientific Basis*. Contribution of working group I to the third assessment report of the intergovernmental panel on climate change, edited by J. T. Houghton et al., pp. 185–237, Cambridge University Press, Cambridge, U.K.
- Riley, G.A. (1939), Plankton studies. II. The western North Atlantic, May–June, 1939, *Mar. Res.*, **2**, 145–162.
- Roullet, G., and G. Madec (2000), Salt conservation, free surface, and varying levels: A new formulation for ocean general circulation models, *J. Geophys. Res.*, **105**(C10), 23,927–23,942.
- Saba, V. S., et al. (2010) The challenges of modeling marine primary productivity through multidecadal climate shifts: A case study at BATS and HOT, *Global Biogeochem. Cycles*, **24**, GB3020, doi:10.1029/2009GB003655.
- Schlitzer, R. (2004), Export production in the equatorial and north Pacific derived from dissolved oxygen, nutrient and carbon data, *J. Oceanogr.*, **60**, 53–62.
- Shigemitsu, M., T. Okunishi, J. Nishioka, H. Sumata, T. Hashioka, M. N. Aita, S. L. Smith, N. Yoshie, N. Okada, and Y. Yamanaka (2012), Development of a one-dimensional ecosystem model including the iron cycle applied to the Oyashio region, western subarctic Pacific, *J. Geophys. Res.*, **117**, C06021, doi:10.1029/2011JC007689.
- Steemann Nielsen, E. (1952), The use of radioactive carbon (C^{14}) for measuring organic production in the sea, *J. Cons. Int. Explor. Mer.*, **18**, 117–140.
- Steemann Nielsen, E., and E. A. Jensen (1957), The autotrophic production of organic matter in the oceans, *Calathea Rep.*, **1**, 49–124.
- Sunda, W. G., and S. A. Huntsman (1995), Iron uptake and growth limitation in oceanic and coastal phytoplankton, *Mar. Chem.*, **50**, 189–206.
- Taghon, G. L., A. R. M. Nowell, and P. A. Jumars (1984), Transport and breakdown of fecal pellets: Biological and sedimentological consequences, *Limnol. Oceanogr.*, **29**, 64–72.
- Timmermann, R., H. Goosse, G. Madec, T. Fichefet, C. Etche, and V. Duliere (2005), On the representation of high latitude processes in the ORCA-LIM global coupled sea ice-ocean model, *Ocean Model.*, **8**, 175–201.
- Westberry, T., M. J. Behrenfeld, D. A. Siegel, and E. Boss (2008), Carbon based primary productivity modeling with vertically resolved photoacclimation, *Global Biogeochem. Cycles*, **22**, GB2024, doi:10.1029/2007GB003078.
- Whitman, W. B., D. C. Coleman, and W. J. Wiebe (1998), Prokaryotes: The unseen majority, *Proc. Nat. Acad. Sci. U.S.A.*, **95**, 6578–6583.