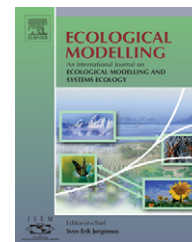


available at www.sciencedirect.comjournal homepage: www.elsevier.com/locate/ecolmodel

Short communication

Quantitative comparison of photoacclimation models for marine phytoplankton

S. Lan Smith*, Yasuhiro Yamanaka

Ecosystem Change Research Program, Frontier Research Center for Global Change, Japan Agency for Marine-Earth Science and Technology, 3173-25 Showa-machi, Kanazawa-ku, Yokohama 236-0001, Japan

ARTICLE INFO

Article history:

Received 14 December 2005

Received in revised form

8 September 2006

Accepted 20 September 2006

Keywords:

Model

Phytoplankton

Data assimilation

Photoacclimation

ABSTRACT

Photoacclimation models for marine phytoplankton describe the changes in their composition (typically C, N and chlorophyll) and growth in response to changing light and nutrient environment. We compared two such models: that of Geider et al. (Geider, R.J., MacIntyre, H.L., Kana, T.M., 1998. A dynamic regulatory model of phytoplankton acclimation to light, nutrients, and temperature. *Limnol. Oceanogr.* 43, 679–694), hereafter the G model, and that of Pahlow (Pahlow, M., 2005. Linking chlorophyll-nutrient dynamics to the redfield N:C ratio with a model of optimal phytoplankton growth. *Mar. Ecol. Prog. Ser.* 287, 33–43), hereafter the P model. Using the Monte Carlo Markov Chain method, we fitted both models to the data set from an incubation experiment by Flynn et al. (Flynn, K.J., Davidson, K., Leftley, J.W., 1994. Carbon–nitrogen relations at whole-cell and free amino-acid levels, during batch growth of *isochrysis galbana* (prymnesiophyceae) under conditions of alternating light and dark. *Mar. Biol.* 118, 229–237). Data consisted of measured concentrations for particulate organic N, particulate organic C, chlorophyll and ammonium. The authors of the G model began their simulation from day 5 for this experiment, claiming that their model could not reproduce the initial lag phase (slow growth during the first few days of the experiment). The author of the P model claimed that its ability to reproduce this initial lag phase (starting from the beginning of the experiment) was a significant improvement over previous models.

Our fitting revealed that the G model can reproduce this initial lag phase as well as the P model, and that both models simulate the data set well. In the best-fits of both models, chlorophyll synthesis during the initial lag phase was limited by the rate of carbon assimilation, even though the G model was designed to have chlorophyll synthesis limited only by the rate of N assimilation. The requirement of organic carbon for energy to assimilate nitrogen results in this indirect limitation of chlorophyll synthesis by C assimilation in the G model, whereas chlorophyll synthesis is explicitly limited by the rate of C assimilation in the P model. This suggests that chlorophyll synthesis is in fact limited by the rate of C assimilation during the initial lag phase. As in the hand-tuned simulations previously published, the P model simulated the initial decrease in Chl:N ratio (as observed) whereas the G model simulated an initial increase in this ratio (contrary to the observations). We also discuss the relative merits of the two models for applications at large scales. Although data

* Corresponding author. Tel.: +81 45 778 5581; fax: 81 45 778 5706.

E-mail addresses: lanimal@jamstec.go.jp (S.L. Smith), galapen@ees.hokudai.ac.jp (Y. Yamanaka).

0304-3800/\$ – see front matter © 2004 Elsevier B.V. All rights reserved.

doi:10.1016/j.ecolmodel.2006.09.016

assimilation is not perfectly objective, because it requires certain choices such as weights for various data types and which data to include when fitting, our results show the advantage of mathematically rigorous fitting as opposed to hand-tuning of models. Our best-fits were significantly better than the hand-tuned fits originally published, especially for the G model, and this yielded insight into the mechanism responsible for the initial lag in phytoplankton growth.

© 2004 Elsevier B.V. All rights reserved.

1. Introduction

Phytoplankton change their composition in response to changes in light and nutrient availability. Such changes are important to marine ecosystems because phytoplankton are the primary producers of organic matter and energy. These variations in composition are also key to interpreting observations of phytoplankton chlorophyll (Chl) in terms of N or C biomass.

To describe such changes, Geider et al. (1998) developed a model (hereafter the G model) of phytoplankton acclimation to light and nutrients, a Photoacclimation model. It describes how the C, N and Chl content of phytoplankton change in response to changes in light and nutrient availability. Recently Pahlow (2005) developed a different photoacclimation model (hereafter the P model) based on the principal of optimizing the daily mean net growth of phytoplankton.

Both studies simulated an experiment by Flynn et al. (1994), in which *Isochrysis galbana* was incubated under alternating light–dark conditions. Both obtained generally good simulations of the experiment, without using any mathematically rigorous fitting algorithm. Pahlow (2005) began the simulation from the start of the experiment, whereas Geider et al. (1998) began from the 5th day because, they reported, their model could not simulate the initial lag phase (a period of slow growth during the first few days of the incubation). Pahlow (2005) claimed that a key advantage of his formulation was that it allowed realistic simulation of this initial lag.

However, in both studies the “hand-tuned” fits of model to data were subjective. Geider et al. (1998)’s decision to begin their simulation from day 5 of the incubation was apparently based only on such “hand-tuning”. Without applying a mathematically rigorous algorithm, one cannot say that the G model cannot simulate the initial lag or that the P model can better simulate it. We sought to quantitatively compare these two models’ abilities to simulate this experiment, using data assimilation.

Flynn et al. (2001) compared various photoacclimation models, by fitting them to data from an experiment in which plankton were incubated under nutrient-replete conditions with step changes in irradiance (Anning et al., 2000). However, that experiment did not include exponential phase growth of phytoplankton followed by depletion of nutrients, as did that of Flynn et al. (1994). Such growth with nutrient depletion approximates an oceanic bloom, and the mechanisms underlying the initial lag phase are likely important to the development of oceanic blooms. Accurate simulations of such blooms are a major goal of large scale marine ecosystem models. Part of the motivation for this study was to evaluate which

of the models considered would be better suited for large scale simulations.

2. Experiment

Flynn et al. (1994) incubated *Isochrysis galbana* under alternating light–dark conditions (an idealized daily cycle) for 26 days. Ammonium was the only form of inorganic nitrogen supplied, and the incubations were conducted in batch mode so that it was depleted as phytoplankton grew. Concentrations of particulate organic carbon, particulate organic nitrogen, chlorophyll and ammonium were measured at 12 h intervals (Fig. 1). They observed an initial lag phase (slow growth) lasting about 5 days, followed by exponential growth, then a stationary phase (with no net C fixation). The initially N-starved cells first took up ammonium, raising their N:C ratio. They then synthesized chlorophyll, which allowed them to grow faster, terminating the lag phase. After nutrient depletion, the phytoplankton continued to assimilate carbon in excess of nitrogen, lowering their N:C ratio until it leveled off around day 20.

There was an imbalance in the total observed nitrogen concentration ($\text{NH}_4 + \text{N}$) of as much as 15% in the experiment. These errors in the data would adversely affect the data assimilation, because the model can not accurately simulate such an imbalance. We therefore excluded observations of organic nitrogen for any time at which this nitrogen imbalance was greater than twice the standard deviation of the nitrogen imbalance for all remaining observations. Thus the observed N was excluded for time = 2.5, 3, 3.5, 4, 4.5 and 6 days.

3. Models

Both models simulate the concentrations of C, N and Chl in phytoplankton, and account for changes in cellular composition in response to light and nutrient environment. Both are quota models, meaning that they are formulated in terms of the intracellular quota of N, expressed as the ratio of N:C.

3.1. G model

Geider et al. (1998) developed the first dynamic model of phytoplankton acclimation to variations in light, inorganic nitrogen availability and temperature. In addition to those environmental variables, rates of assimilation of C and N and of Chl synthesis depend on intracellular N:C and Chl:C ratios. The model was based on a mechanistic understanding of these processes and was formulated to consistently simulate general observations about variations in phytoplankton composition and growth. Chl synthesis depends directly on N assimilation, based on the requirement of N for protein synthesis. Exclud-

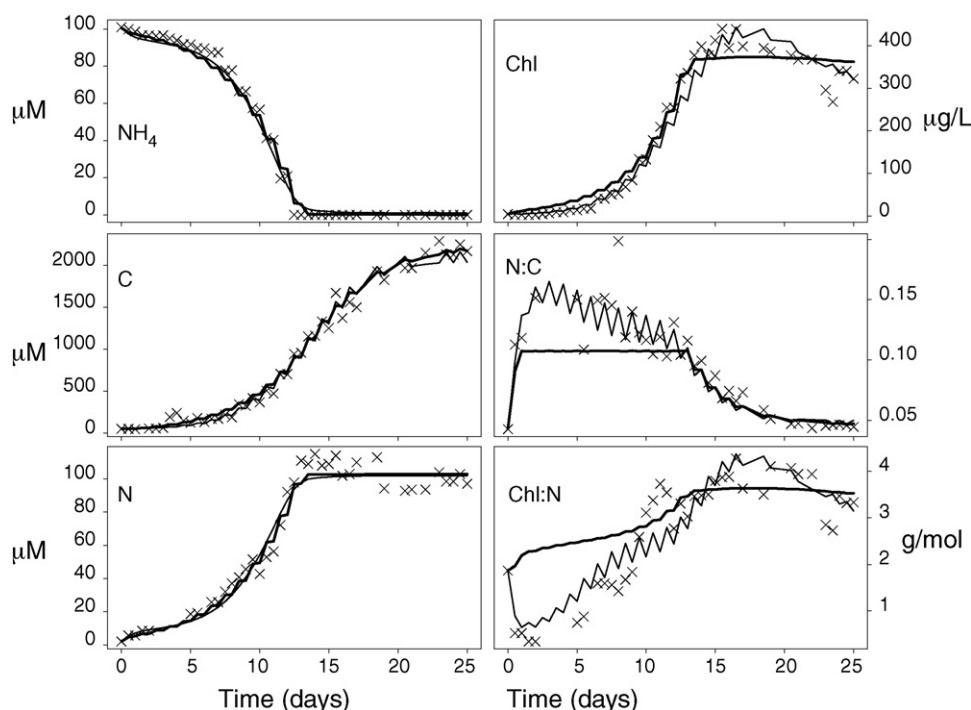


Fig. 1 – Best-fit simulations (thick lines, G model; thin lines, P model) and observations (X's) for NH_4 , phytoplankton C (C), phytoplankton N (N) and chlorophyll (Chl), molar N:C ratio and Chl:N ratio.

ing the activation energy (temperature dependence) which we do not consider here, the G model has 10 parameters (Table 1) or 9, if not counting the N uptake rate as a free parameter, because its value is fixed in terms of two others.

3.2. P model

Pahlow (2005) developed a model based on the concept that phytoplankton adjust their composition in order to maximize

their daily average net growth rate, by optimally allocating resources among competing requirements for nutrient uptake, light-harvesting and growth. While it also applies mechanistic formulations for these processes, it further solves for the optimal distribution of N and Chl, in terms of the intracellular N:C and Chl:C ratios. Chl synthesis depends directly on C assimilation, based on the requirement of energy for this process, and indirectly on N assimilation through the cellular N:C ratio. The P model has eight parameters (Table 1).

Table 1 – Parameters in the models

Symbol	Value	Units	Description
G model			
α^{Chl}	0.460×10^{-5}	$\text{g C } (\mu\text{mol phot g Chl})^{-1} \text{ m}^{-2}$	Chl-specific light adsorption coefficient
p^{ref}	3.00	day^{-1}	Maximum growth rate at T_{ref}
Q_{min}	0.0340	(molar ratio, N:C)	Minimum cell quota
Q_{max}	0.143	(molar ratio, N:C)	Maximum cell quota
$V_{\text{C}}^{\text{N,ref}}$	$Q_{\text{max}} p^{\text{ref}}$	day^{-1}	Maximum possible DIN uptake rate
R^0	0.0250	day^{-1}	Respiration rate at T_{ref}
ζ_{N}	1.80	mol C/mol N	Cost of assimilating NH_4
$\varphi^{\text{N,max}}$	0.300	g Chl/g N	Maximum ratio of Chl-to-N
K_{NH_4}	0.100	$\mu\text{mol N l}^{-1}$	Half-saturation constant, NH_4 uptake
n	0.0500		Shape parameter for N uptake
P model			
α^{Chl}	0.957×10^{-5}	$\text{g C } (\mu\text{mol phot g Chl})^{-1} \text{ m}^{-2}$	Chl-specific light adsorption coefficient
μ_{d}	4.45	day^{-1}	Growth rate
R^0	0.180	day^{-1}	Respiration rate at T_{ref}
Q_0	0.0400	(molar ratio, N:C)	Reference cell quota
r_{d}	0.500		Ratio of dark-light rates
ζ_{N}	1.80	mol C/mol N	Cost of assimilating NH_4
ξ^{Chl}	10.7	g Chl/g C	Reference ratio of Chl-to-C
A_0	0.0100	$\text{m}^3 (\text{mmol N day})^{-1}$	Potential affinity for NH_4 uptake

4. Solution method

Discrete approximations in time to the governing equations were solved with a fourth order Runge-Kutta method. Small timesteps are required to simulate this experiment because ammonium is depleted to very low concentrations. Timesteps were adjusted by trial and error to be small enough that doubling the timestep did not significantly change the calculations. This resulted in a timestep of 4 min for the G model and 1 h for the P model. The difference results from the formulation of the G model in terms of instantaneous gross rates (i.e., separately for uptake and respiration, both of which vary in time), versus that of the P model in terms of daily mean net rates (i.e., only the sum of uptake and respiration, averaged over the daily cycle). The G model depleted ammonium to lower concentrations than the P model, requiring a smaller timestep.

5. Assimilation method

To fit the models to the observations, we applied the Monte Carlo Markov Chain (MCMC) (Harmon and Challenor, 1997; Hargreaves and Annan, 2002; Smith et al., 2006). Briefly, this method conducts a random walk through parameter space, such that it seeks to sample the complete joint probability distribution (pdf) of the parameter space given the observations and any prior information about the parameter values. It is similar to simulated annealing, as both methods are Metropolis-Hastings algorithms (Metropolis et al., 1953). The key difference is that because MCMC seeks to sample the whole pdf, it provides more information (in the form of the ensemble of values of parameters and simulated quantities), which can be useful in data assimilation studies. MCMC therefore requires more calculations than simulated annealing, which converges more efficiently to the same optimal solution, but does not provide this additional information. Ideally, any optimization method should suffice, provided that it has a solid mathematical basis.

We define our cost function (which the method aims to minimize) and describe our selection of which parameters to

vary in the assimilations. We fit both models to all data for concentrations, except for certain organic N data that were excluded because of the observed mass imbalance (Section 2). To calculate the cost function we used the common approach of using the reciprocal of the estimated standard deviation of measurement, σ_i , as a weight for each data type i . The total cost is:

$$\text{cost} = \frac{1}{W} \sum_n \sum_i \left(\frac{C_{i,n}^{\text{sim}} - C_{i,n}^{\text{obs}}}{\sigma_i} \right)^2 \quad (1)$$

in which C_i^{sim} is the simulated value corresponding to observed value C_i^{obs} , and n sums over the observation times. We estimated the standard deviations as $0.050 \mu\text{M}$ for ammonium (NH_4), $0.10 \mu\text{M}$ for organic nitrogen (N), $1.0 \mu\text{M}$ for organic carbon (C), and $0.40 \mu\text{g L}^{-1}$ for chlorophyll (Chl). W is a "widening factor" (Annan et al., 2005), applied in order to achieve a realistic spread in parameter values (to account for model inadequacy). We selected its value, $W = 500$, by trial-and-error. This affects neither the position of the optimum in parameter space nor the relative ranges of the different parameters, but it allows a wider spread of results than would be achieved under the perfect model assumption.

We used a series of paired assimilations to choose which subset of parameters to vary in the assimilations as did Smith et al. (2006). For each model, one of the paired assimilations started from the initial guess values, and the other started from 60% of those values. The initial guess parameter values for each model were the values from each original study (except that some of the published values were in error for the P model; Markus Pahlow, personal communication). If the final answer is independent of the initial guesses (at least for the values applied), each parameter should converge to the same value in both assimilations. If the two paired assimilations converge to different parameter values, this means that either 1 at least one of the assimilations has converged to a local minimum, or 2 the data is insufficient to constrain all parameter values. If the paired assimilations did not approach convergence within a few million simulations (approximately 1 week of wall clock time for the G model, less for the P model, running on an Apple G5 with dual 2 GHz processors), we cut down the candidate parameter set and repeated.

Table 2 – Values of parameters varied in the assimilations

Parameter	Initial	Best-fit	Units
G model			
α_{Chl}	0.460×10^{-5}	0.539×10^{-5}	$\text{g C } (\mu\text{mol phot g Chl})^{-1} \text{ m}^{-2}$
p^{ref}	3.00	0.958	day^{-1}
Q_{max}	0.143	0.109	(molar ratio, N:C)
Q_{min}	0.0340	0.0426	(molar ratio, N:C)
R^0	0.0250	0.0261	day^{-1}
$\phi^{\text{N,max}}$	0.300	0.280	g Chl/g N
P model			
α_{Chl}	0.957×10^{-5}	0.824×10^{-5}	$\text{g C } (\mu\text{mol phot g Chl})^{-1} \text{ m}^{-2}$
μ_{d}	4.45	3.19	day^{-1}
Q_0	0.0400	0.0414	(molar ratio, N:C)
R^0	0.176	0.144	day^{-1}
ξ_{Chl}	10.7	10.3	g Chl/g C
A_0	0.0105	0.0100	$\text{m}^3 (\text{mmol N day})^{-1}$

6. Results and discussion

For both models, initial paired assimilations varying all parameters did not converge to the same values, meaning that we could not uniquely constrain all parameters values. We were able to constrain six parameters for each model (Table 2). We arrived at these parameter subsets by systematically reducing the number of parameters varied until the assimilations converged (eliminating parameter ζ_N , the value of which is well known from thermodynamics, and a few others to which the solution was least sensitive). We did find similar best-fit simulations (and slightly lower errors) from assimilations varying more parameters, although those assimilations did not converge (paired assimilations with the same model yielded different parameter values).

Total errors with the initial guess parameter values were 68,000 for the G model and 6500 for the P model. Overall, the assimilation yielded good fits of both models to the data (Fig. 1). Total best-fit error (cost) was slightly lower for the G model (824) than for the P model (957). Over the initial lag phase (time < 5 days), the total (best-fit) error was 60.5 for the G model and 77.1 for the P model. During this period the greatest differences in cost were for NH_4 and C, which the G model simulated better, but the P model's cost for Chl was much lower (0.491 compared to 12.9 for the G model). The relative difference in total error over the whole experiment was $(957 - 824)/824 = 0.16$. For the initial lag phase it was $(77.1 - 60.5)/60.5 = 0.27$.

With the weights applied here, the G model agrees slightly better with the observations for both the initial lag phase and the whole experiment, and even more so for the former. Not too much can be made of this slight difference in error (cost), which depends on the weights applied to the various data types and on the choice of which data to include in the assimilation. However, the weights we applied are reasonable, and even with moderate changes to them the overall fits would be comparably good for the two models. It may be possible to constrain more parameters that we did in this study. However, further improvements in either model's fit to the data would necessarily be small, given the quality of the fits obtained.

For the G model, our best-fit is significantly better than the initial guess (using the parameter values from the original study), because we consider the whole data set, whereas Geider et al. (1998) started their simulation from day 5 of the experiment. For the P model, the improvement was only slight, because the simulation in Pahlow (2005) was already quite good. Our best-fit values of parameters for the P model were mostly close to the values of Pahlow (2005). The largest difference was for the growth rate, μ_d , for which the best-fit value is approximately 70% of the initial guess value. Differences in best-fit versus initial guess parameter values were greater for the G model.

Compared to the original parameter values from Geider et al. (1998), what differences in our best-fit values allowed such a good fit to the whole data set? The biggest difference is for the maximum rate of photosynthesis, P^{ref} , for which

our best-fit value is approximately one-third of the original value (Table 2). This slower rate, together with the higher (by about 12%) value of the Chl-specific absorption coefficient, α^{Chl} , yielded slower growth in the beginning of the experiment and a greater increase in the rate of photosynthesis as Chl content increased. The maximum N:C ratio, Q_{max} , was lower than its original value by approximately 25%, and the minimum N:C ratio, Q_{min} , was higher by approximately 25%. This narrowed the range over which the N:C ratio could vary, to more closely match the data. The higher value of Q_{max} also limits N assimilation during the initial lag phase, because N assimilation approaches zero as phytoplankton N:C approaches Q_{max} , which it does by day 2. The N:C ratio remains near its maximum value until approximately day 13. This means that phytoplankton must assimilate more C in order to assimilate more N. Initially, the low Chl content limits the rate of C assimilation (photosynthesis). The rate of Chl synthesis, which in the G model depends directly on N assimilation, therefore depends indirectly on C assimilation during the initial lag phase, mostly because of the lower best-fit value of Q_{max} .

The G model was designed to have chlorophyll synthesis limited by the rate of N assimilation, but not energetically (by the rate of photosynthesis) as in the P model. However, when fit to this data set, during the initial lag phase chlorophyll synthesis in the G model becomes energetically limited through the energetic requirement for assimilating N. This supports the idea that chlorophyll synthesis is in fact energetically limited during the initial lag phase.

6.1. N:C and Chl:N ratios

During the first 10 days of the experiment, the P model reproduces the decrease in the N:C ratio with time (Fig. 1), as phytoplankton assimilate more C. In the G model, N:C ratio remains at its maximum value throughout this period. The P model also correctly simulates the initial decrease in Chl:N ratio for the N-starved cells suddenly exposed to higher nutrients in this experiment (Fig. 1). Pahlow (2005) also simulated this decrease in Chl:N ratio and explained that it results from the direct dependence of Chl synthesis on C assimilation in the P model, rather than on N assimilation (as in the G model). Even in its best-fit solution, the G model simulated an increase in this ratio, counter to the observed trend. Still, the P model overestimates Chl:N throughout the first 10 days of the experiment, though not by as much as the G model.

6.2. Considerations for large scale simulations

The P model's formulation in terms of daily mean net rates is an advantage for large-scale simulations, because it allows longer timesteps. This also means that it cannot resolve day-night differences in nitrogen uptake, because it is not formulated in terms of instantaneous rates (which can vary with light). Such differences were observed in the experiment (Fig. 1) and were simulated by the G model. For simulating such short-term dynamics, the G model or other, similar models (Flynn et al., 2001) would be preferable to the P model. However, such high resolution observations are gen-

erally not available from the field, even at ocean time-series sites.

The P model also has one less adjustable parameter (or two less, if one counts the nitrogen uptake rate, $V_C^{N,ref}$ in the G model as a free parameter), which is an advantage in terms of keeping large-scale models as simple as possible. Even though with the weights we chose for the data, total error for the G model was slightly lower than for the P model, this slight difference is not significant when considering timescales longer than 1 day. This is because the G model resolves differences in the assimilation of nitrogen in light versus darkness, which the P model cannot because it is formulated in terms of growth averaged over the day–night cycle. Such differences could only be resolved by observations more frequent than once per day, which are rarely available at large scales.

7. Conclusions

The G model and P model can simulate the observed concentrations comparably well. The P model is clearly superior in its ability to simulate the dynamics of the N:C and Chl:N ratios, but overall it did not simulate the observed concentrations significantly better than the G model (based on the weights we applied), either for the initial lag phase or for the whole experiment.

Beyond this specific example, our main point is that comparisons of the ability of different models to simulate observations should be based on mathematically rigorous fitting of each model to the observations. Although such fitting is not completely objective, because of the necessity to define weights for various data types and to decide which data to include, it is still superior to hand-tuning. As shown here for the initial lag phase, rigorous fitting can also result in unexpected model behavior (the energetic limitation of chlorophyll synthesis in the G model), and can thereby provide insight into the dynamics of underlying processes.

Acknowledgment

We thank the anonymous reviewers and the editor for their thorough reading and constructive comments.

REFERENCES

- Annan, J.D., Hargreaves, J.C., Ohgaito, R., Abe-Ouchi, A., Emori, S., 2005. Efficiently constraining climate sensitivity with paleoclimate simulations. *SOLA* 1, 181–184.
- Anning, T., MacIntyre, H.L., Pratt, S.M., Sammes, P.J., Gibb, S., Geider, R.J., 2000. Photoacclimation in the marine diatom *Skeletonema costatum*. *Limnol. Oceanogr.* 45 (8), 1807–1817.
- Flynn, K.J., Davidson, K., Leftley, J.W., 1994. Carbon–nitrogen relations at whole-cell and free amino-acid levels, during batch growth of *isochrysis galbana* (prymnesiophyceae) under conditions of alternating light and dark. *Mar. Biol.* 118, 229–237.
- Flynn, K.J., Marshall, H., Geider, R.J., 2001. A comparison of two *n*-irradiance interaction models of phytoplankton growth. *Limnol. Oceanogr.* 46 (7), 1794–1802.
- Geider, R.J., MacIntyre, H.L., Kana, T.M., 1998. A dynamic regulatory model of phytoplankton acclimation to light, nutrients, and temperature. *Limnol. Oceanogr.* 43, 679–694.
- Hargreaves, J.A., Annan, J.D., 2002. Assimilation of paleo-data in a simple earth system model. *Clim. Dyn.* 19, 371–381.
- Harmon, R., Challenor, P., 1997. A markov chain monte carlo method for estimation and assimilation into models. *Ecol. Model.* 101, 41–59.
- Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H., Teller, E., 1953. Equations of state calculations by fast computing machines. *J. Chem. Phys.* 21, 1087–1091.
- Pahlow, M., 2005. Linking chlorophyll–nutrient dynamics to the redfield N:C ratio with a model of optimal phytoplankton growth. *Mar. Ecol. Prog. Ser.* 287, 33–43.
- Smith, S.L., Casareto, B.E., Niraula, M.P., Suzuki, Y., Hargreaves, J.C., Annan, J.D., Yamanaka, Y., 2006. Examining the regeneration of nitrogen by assimilating data from incubations into a multi-element ecosystem model. *J. Mar. Syst.*, doi:10.1016/j.jmarsys.2006.03.013.