

## Examining the regeneration of nitrogen by assimilating data from incubations into a multi-element ecosystem model

S. Lan Smith <sup>a,\*</sup>, B.E. Casareto <sup>b</sup>, M.P. Niraula <sup>c</sup>, Y. Suzuki <sup>c</sup>, J.C. Hargreaves <sup>d</sup>,  
J.D. Annan <sup>d</sup>, Y. Yamanaka <sup>a</sup>

<sup>a</sup> *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*

<sup>b</sup> *Laboratory of Aquatic Sciences Consultant Co., Tokyo, Japan*

<sup>c</sup> *Global Biogeochemistry Laboratory, Shizuoka University, Shizuoka, Japan*

<sup>d</sup> *Frontier Research Center for Global Change, Yokohama, Japan*

Received 29 August 2005; received in revised form 20 December 2005; accepted 20 March 2006

Available online 22 June 2006

### Abstract

We examined the flows of nitrogen in two batch incubations of plankton assemblages under controlled conditions, using mesopelagic seawater from different depths. Observations included concentrations of nutrients, organic matter (particulate and dissolved) and plankton (biomass by species). Because nitrogen flows were not observed directly (as is generally the case), we examined them using a multi-element ecosystem model developed to simulate these experiments. Dynamic changes in the observed bulk POC/PON ratio (C/N ratio of particulate organic matter, POM) were consistent with a quota model for phytoplankton. To relate these changes to the regeneration of nitrogen, we included such a quota model and all major planktonic groups observed.

By assimilating the data (all observations of carbon and nitrogen concentrations, including living carbon biomasses) into the model using the Monte Carlo Markov Chain, we accurately simulated the changes in bulk POC/PON ratio and its difference between incubations. Changes in phytoplankton composition dominated the changes in bulk POC/PON ratio. Bacteria dominated the regeneration of nitrogen in both incubations, and changes in bacterial regeneration strongly affected the bulk POC/PON ratio. After nitrate depletion, gross nitrogen regeneration varied inversely with observed POC/PON ratio and was constrained by this ratio. Thus, we have demonstrated a powerful new approach to constraining the flows of nutrients through ecosystems. Our results suggest that simultaneous observations of POC and PON could place an additional constraint on the supply of inorganic nitrogen in models of the oligotrophic ocean.

© 2006 Elsevier B.V. All rights reserved.

*Keywords:* Marine ecosystem models; Nitrogen regeneration; Data assimilation; Phytoplankton; Particulate organic matter ratios

### 1. Introduction

There is much interest in understanding the relationships between the structure of ecosystems and the material flows of nutrients and carbon. In general, concentrations are relatively easy to measure, but flows are either difficult or impossible to measure directly. This

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

E-mail address: [lanimal@jamstec.go.jp](mailto:lanimal@jamstec.go.jp) (S.L. Smith).

study aims to examine the flows of nitrogen in two incubation experiments, from which we obtained detailed observations of concentrations (nutrients, organic matter and living carbon biomasses estimated from cell counts). As is generally the case, there were no direct observations of material flows through the ecosystem. We hypothesize that the bulk POC/PON ratio is related to the flows of nitrogen through the ecosystem.

Available data is usually not sufficient to constrain values of all parameters in complex ecosystem models. To make the best use of the information contained in the data, it is therefore necessary to either modify the model until all parameters can be constrained (Spitz et al., 1998) or identify which subset of parameters is important to the given application and can be constrained (Harmon and Challenor, 1997). Vallino (2000) applied several methods to assimilate observations from mesocosm experiments into an ecosystem model and found that the best-fit values of parameters differed from values reported in the literature (based on culture experiments or allometric relationships). Previous studies had found the same discrepancy. Values obtained by data assimilation with one model cannot be applied in other models. We seek an accurate simulation of the ecological state and some measure of its uncertainty (Dowd and Meyer, 2003), which requires optimising parameter values, but we focus on relating the observations to the flows, rather than on parameter estimation per se.

We seek to relate the regeneration of nitrogen to the bulk POC/PON ratio (the C/N ratio of bulk particulate organic matter, POM). The dynamics of this ratio were consistent with a quota model of phytoplankton composition (Flynn, 2003). Bulk POC/PON ratios were low around the time of initial nutrient drawdown, when phytoplankton grew under nitrogen-replete conditions and higher after nitrate was depleted, when the only source of nitrogen was regeneration. We designed an ecosystem model specifically to simulate these experiments, embedding the quota model of Geider et al. (1998). Our hypothesis was that, by assimilating the detailed set of observed concentrations into our model, we could constrain the flows of nitrogen and thereby examine its regeneration.

## 2. Methods

### 2.1. Experiments

Incubations were conducted using deep-sea water from both 400 m and 700 m, sampled from Suruga Bay, Japan. Here we describe them briefly. For a detailed description, see Niraula et al. (2005). These experiments

were designed to examine the dynamics of blooms of planktonic assemblages under low grazing pressure. The use of unaltered mesopelagic water accomplishes this without the disturbing effects of filtration to remove zooplankton (which would be required for surface water samples). The difference in initial nutrient concentrations between the two depths allows an examination of the effects of nutrient concentration on growth.

Samples were collected using 10 L Niskin bottles mounted on a CTD-Carousel sampler system (SBE 9 Plus, Sea-Bird Electronics). For measurements of initial nutrient concentrations, samples were collected directly from the Niskin bottles into 100 mL polycarbonate bottles, which were immediately frozen until analyzed. Water was collected directly from the Niskin bottles into 10 L polycarbonate bottles for initial measurements and into 30 L polycarbonate bottles for incubations. All polycarbonate bottles were rinsed with 10% HCl solution followed by deionized water (MilliQ) prior to sampling, then with sampled water from the Niskin bottles prior to collection, after which they were covered to exclude sunlight until analysis or incubation. The bottles were incubated for 11 days (in duplicate for each depth) inside 1000 L water baths of circulating tap water, secured so that circulation of the tap water kept the bottles revolving. The experiment was conducted out doors with natural illumination attenuated by 90% using a hut made of mesh screen. Bottles were sampled approximately every 48 h. Water temperature and light intensity were measured continuously in situ (Alec Electronics Sensors MDS-MkV/T and MDS-MkV/L, respectively). Temperature ranged from 23.6 to 26.2 °C (mean: 24.5 °C). Maximum irradiance inside the water bath was 948  $\mu\text{E}$  (mean: 156  $\mu\text{E}$ ).

Chemical and biological quantities measured were: abundance by species for phyto- and zooplankton; abundance of picocyanobacteria, heterotrophic bacteria, nano-flagellates; concentrations of nutrients, POC, PON, DOC, DON (dissolved organic carbon, nitrogen), Psi (particulate silica) and Dsi (dissolved silicate). Filtrations for POM and DOM used pre-combusted GF/F filters. POM samples were analyzed using a Sumigraph NC-90A (Sumika Co.) to measure POC and PON simultaneously for each filtered sample. DOC and DON were measured (simultaneously for each sample) using a Sumigraph TOC-90A. Nutrients were measured using a four-channel BRAN-LUEBBE autoanalyzer with continuous flow system. Heterotrophic bacteria were counted under an epifluorescence microscope (Nikon-Eclipse) with previous staining by DAPI. Picocyanobacteria were counted under epifluorescence using a B filter. From cell counts and biovolumes

(calculated from observed dimensions), biomasses were estimated for bacteria and nanoflagellates, and by species for phytoplankton and zooplankton (Niraula et al., 2005). The phytoplankton biomass was dominated by diatoms. Bacteria dominated the heterotrophic biomass except for 1 day in the 700 m incubation, and biomasses of nanoflagellates and ciliates were significant.

The estimated uncertainties of measurement (from sample collection and analytical instruments),  $\sigma$ , were: 0.05  $\mu\text{M}$  for  $\text{PO}_4$ ; 0.10  $\mu\text{M}$  for  $\text{NO}_3$ ,  $\text{NH}_4$ , POC and PON, and for the biomasses of bacteria (BactC), small phytoplankton (PhySC) and heterotrophic nanoflagellates (HNFC); 1.0  $\mu\text{M}$  for the biomasses of small diatoms (DiaSC, <5  $\mu\text{m}$  size) and zooplankton (ZooC); 0.3  $\mu\text{M}$  for DON; 1.0  $\mu\text{M}$  for DOC; and 2.0  $\mu\text{M}$  for biomass of large diatoms (DiaLC, >5  $\mu\text{m}$  size).

The total amount of nitrogen observed was lowest midway through each incubation (Fig. 1), and its variations exceeded the uncertainty of measurement. The pattern for silicon (not shown) was similar. This suggests that the problem is with the POM and BSi (rather than DOM) measurements, because for silicon there is no pool corresponding to DON. However, the changes in nutrients, biomasses and organic matter far

exceeded the mass imbalance, allowing meaningful interpretation of those changes. Still, no mass-conservative model will be able to fit the observed imbalance, and we account for this in the assimilation (Section 2.5).

## 2.2. Model description

The ecosystem model (Fig. 2) includes 21 compartments. Its structure is similar to that of Smith et al. (2005), with the addition of the quota model of Geider et al. (1998) for phytoplankton.

For each of the three classes of phytoplankton (non-silicious phytoplankton, PhyS, small diatoms, DiaS, and large diatoms, DiaL), the carbon, nitrogen and chlorophyll concentrations are simulated separately, according to the model of Geider et al. (1998). For the three heterotrophic compartments (bacteria, B, nanoflagellates, HNF, and zooplankton Z) only the nitrogen concentrations are simulated and fixed stoichiometries are assumed.

Trophic relationships are determined according to the size of organisms. Heterotrophic nanoflagellates (HNF) graze bacteria (B) and small phytoplankton (PhyS). Zooplankton (Z) graze small phytoplankton (PhyS), small diatoms (DiaS) and heterotrophic nanoflagellates

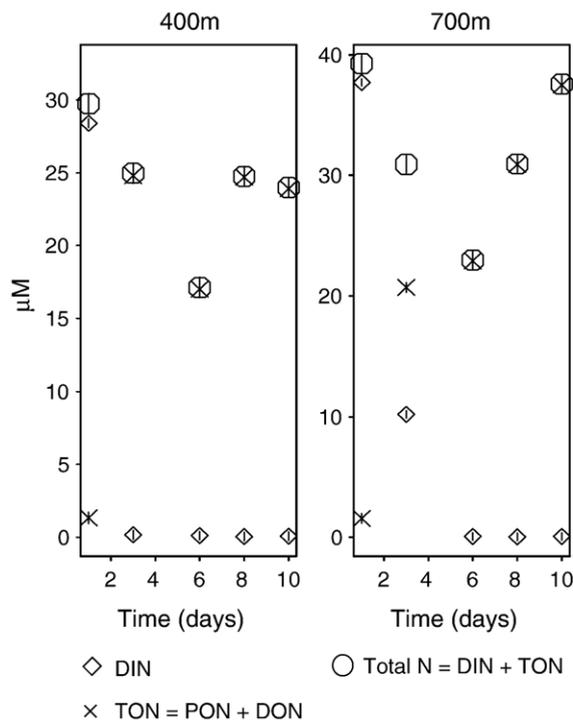


Fig. 1. Mass balance of nitrogen for each incubation (400 m and 700 m water). Diamonds are dissolved inorganic nitrogen (DIN), X's are total organic nitrogen (TON), which equals dissolved plus particulate organic nitrogen (DON+PON), and circles are total observed nitrogen (DIN+DON+PON). Vertical lines are error bars.

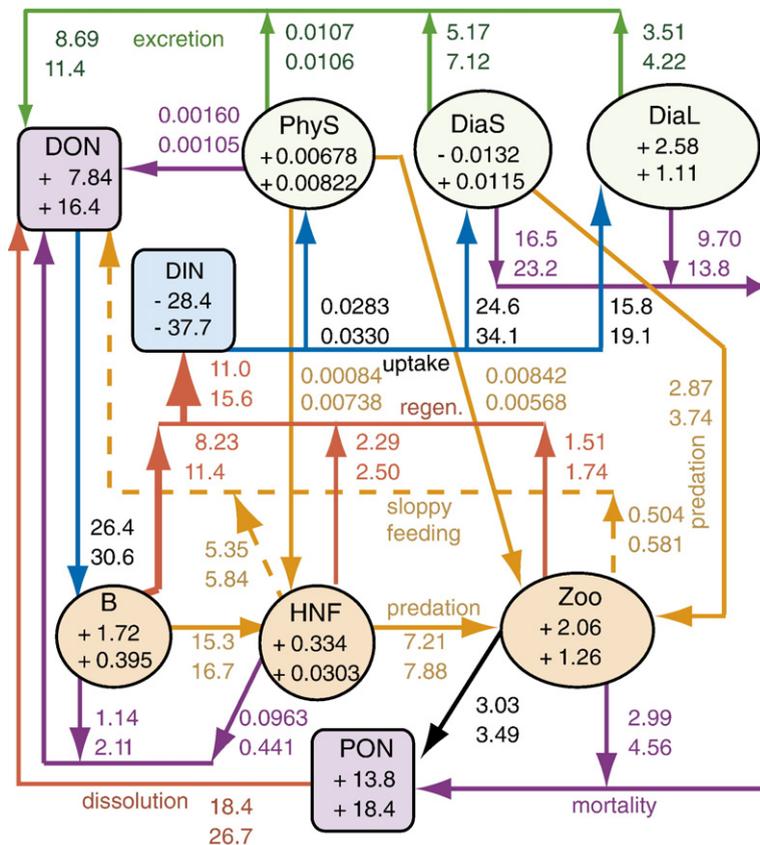


Fig. 2. Diagram of the carbon and nitrogen flows in the ecosystem model, with cumulative flows of nitrogen in the best-fit simulations of both incubations. Values are for the 400 m (upper) and 700 m (lower) case. Flows (beside arrows) are in  $\mu\text{M N}$ . In each compartment's box: the change in its standing stock over the course of the simulation (final–initial concentration, in  $\mu\text{M N}$ ). Compartment DIN includes both  $\text{NH}_4$  and  $\text{NO}_3$ .

(HNF). No zooplankton large enough to eat the large diatoms was present in the incubations, and they are not included in the model.

Two compartments are included for the carbon and nitrogen in dissolved organic matter (DOM), and two more for particulate organic matter (POM). For bacteria and DOM, the formulation is based on Anderson and Williams (1998), simplified to only one class of (non-refractory) DOM and with a first-order (in DOM) rate expression. Refractory DOM is assumed constant.

The model also includes nitrate ( $\text{NO}_3$ ), ammonium ( $\text{NH}_4$ ), silicic acid (DSi) and biogenic silica (BSi). Phosphate ( $\text{PO}_4$ ) is simulated implicitly, assuming a Redfield N/P ratio for all organic components. This is justified by the observed N/P drawdown ratio and the fact that  $\text{PO}_4$  was never depleted in the observations.

### 2.3. Parameters

Some changes were made to the phytoplankton model of Geider et al. (1998) to embed it into our ecosystem model. The original model, applied to

culture experiments, assumed no mortality and set the respiration rates to zero. We set the respiration rate coefficient at  $0.10 \text{ day}^{-1}$  from a study simulating ocean station BATS with this phytoplankton model (Lefevre et al., 2003). We set initial guesses for the maximum rates of photosynthesis and let them vary in the assimilation. Mortality rates were based on those of Fujii et al. (2002a), which were used in simulations of several ocean time series sites. All other parameters for phytoplankton (Table 1) are from Geider et al. (1998) or Fujii et al. (2002a). Parameters for bacteria and DOM (Table 2) are from Anderson and Williams (1998) or are based on their values. We applied growth efficiencies of 0.50 for heterotrophic nanoflagellates (Goldman and Caron, 1985) and zooplankton (Verity, 1991). Initial guesses for grazing rates were tuned to match the data, and these rates were varied in the assimilation. For POM degradation, we used the first-order rate reported for non-refractory POM by Fujii et al. (2002b) based on their decomposition experiments (neglecting the small refractory fraction of POM found in that study).

Table 1  
Parameters for phytoplankton

Parameter	Value	Units	Description
<i>i = PhyS, DiaS and DiaL, indicates a common value for all, otherwise specified separately</i>			
$P_{PhyS}^{ref}$	15.0 <sup>a</sup>	day <sup>-1</sup>	Max. growth rate at $T_{ref}$
$R_i$	0.10	day <sup>-1</sup>	Respiration rate at $T_{ref}$
$\alpha^{Chl}$	$3 \cdot 10^{-5}$	g C ( $\mu\text{mol phot g Chl}$ ) <sup>-1</sup> m <sup>-2</sup>	Chl-specific light adsorption coefficient
$Q_{min,i}$	0.034	(molar ratio, N/C)	Min. cell quota
$Q_{max,i}$	0.17	(molar ratio, N/C)	Max. cell quota
$V_{C,i}^{N,ref}$	$Q_{max,i} \cdot P_i^{ref}$	day <sup>-1</sup>	Max. possible DIN uptake rate
$\Phi_i$	1.5	( $\mu\text{mol N L}^{-1}$ ) <sup>-1</sup>	NH <sub>4</sub> inhibition of NO <sub>3</sub> uptake
$\gamma_{PhyS}$	0.20		Ratio of excretion: gross production
$\zeta_{N,i}$	2.0	mol C/mol N	Cost of assimilating NO <sub>3</sub> and NH <sub>4</sub>
$\Theta_i^{N,max}$	0.3	g Chl/g N	Max. ratio of Chl-to-N
$m_{PhyS}$	$2.0 \cdot 10^5$	day <sup>-1</sup> ( $\text{mol N L}^{-1}$ ) <sup>-1</sup>	Mortality rate coefficient at $T_{ref}$
$K_{NO_3,PhyS}$	0.5	$\mu\text{mol N L}^{-1}$	Half-saturation constant, NO <sub>3</sub> uptake
$K_{NH_4,PhyS}$	0.1	$\mu\text{mol N L}^{-1}$	half-saturation constant, NH <sub>4</sub> uptake
<i>For diatoms, d = DiaS and DiaL (indicating a common value for both)</i>			
$P_{DiaS}^{ref}$	20.0 <sup>a</sup>	day <sup>-1</sup>	Max. growth rate at $T_{ref}$
$\gamma_{DiaS}$	0.20		Ratio of excretion: gross production
$P_{DiaL}^{ref}$	13.0 <sup>a</sup>	day <sup>-1</sup>	Max. growth rate at $T_{ref}$
$\gamma_{DiaL}$	0.135		Ratio of excretion: gross production
$K_{NO_3,d}$	1.0	$\mu\text{mol N L}^{-1}$	Half-saturation constant, NO <sub>3</sub> uptake
$K_{NH_4,d}$	0.10	$\mu\text{mol N L}^{-1}$	Half-saturation constant, NH <sub>4</sub> uptake
$K_{Si,d}$	1.5	$\mu\text{mol N L}^{-1}$	Half-saturation constant, DSi uptake
$m_d$	$1 \cdot 10^5$ <sup>a</sup>	day <sup>-1</sup> ( $\text{mol N L}^{-1}$ ) <sup>-1</sup>	Mortality rate coefficient at $T_{ref}$

<sup>a</sup> Parameter varied in the assimilations, using this value as the initial guess.

Table 2  
Other parameters in the ecosystem model

Parameter	Value	Units	Description
<i>For zooplankton</i>			
$RCN_{HNF}, RCN_Z$	5.5	(molar ratio, C/N)	Fixed stoichiometry of HNF and Z
$g_{p2HNF}$	9.0 <sup>a</sup>	day <sup>-1</sup>	Grazing rates at $T_{ref}$
$g_{p2Z}$	3.0 <sup>a</sup>	day <sup>-1</sup>	Grazing rates at $T_{ref}$
$\alpha_{HNF}, \alpha_Z$	0.7		Assimilation efficiencies for HNF and Z
$\beta_{HNF}, \beta_Z$	0.5		Growth efficiencies for HNF and Z
$\delta_S$	0.25		Fraction of excretion as DOM
$\lambda_I$	1.4	( $\mu\text{M N}$ ) <sup>-1</sup>	Ivlev constant
$p_{min}$	0.020	$\mu\text{M N}$	Min. prey concentration for grazing
$m_{HNF}, m_Z$	$1 \cdot 10^5$	day <sup>-1</sup> ( $\text{mol N L}^{-1}$ ) <sup>-1</sup>	Mortality rate coefficients at $T_{ref}$
<i>For bacteria</i>			
$RCN_B$	5.1	(molar ratio, C/N)	Fixed stoichiometry of B
$\mu_B$	0.20	( $\mu\text{mol N day}$ ) <sup>-1</sup>	Max. uptake rate for DOC at $T_{ref}$
$\omega_B$	0.27 <sup>a</sup>		Bacterial growth efficiency (C-based)
$K_{NH_4,B}$	1.0	$\mu\text{mol N L}^{-1}$	Half-saturation constant, NH <sub>4</sub>
$m_B$	0.02	day <sup>-1</sup>	Bacterial mortality rate coefficient at $T_{ref}$
<i>Others</i>			
$RNP_{Org}$	16	(molar ratio, N/P)	Fixed stoichiometry for all organisms
$RDOC$	39.7	$\mu\text{M C}$	Refractory DOC concentration
$RDON$	2.30	$\mu\text{M N}$	Refractory DON concentration
$k_{d,POM}$	0.0375	day <sup>-1</sup>	Dissolution rate of POM at $T_{ref}$
$k_{r,POM}$	0.0	day <sup>-1</sup>	Regeneration rate of POM at $T_{ref}$
$E_a$	$47.3 \cdot 10^3$	J mol <sup>-1</sup>	Activation energy
$T_{ref}$	20.0	(°C)	Reference temperature

<sup>a</sup> Parameter varied in the assimilation, using this value as the initial guess.

#### 2.4. Initial conditions

The simulation began on the second day for each incubation, because some observed biomasses were below detection limits and because of the difficulty of simulating the lag phase at the beginning of the experiment (Geider et al., 1998). Nutrients, DOC and DON were initialized to their observed values. Carbon biomasses were initialized to observed values with added perturbations to any initial biomass smaller than its uncertainty of measurement (Section 2.5). Initial nitrogen and chlorophyll for phytoplankton were calculated assuming a C/N ratio of 6.6 and a Chl/N ratio of 0.15 (g Chl/g N), which was half the maximum value from Geider et al. (1998). Nitrogen biomasses for heterotrophs were set based on their observed carbon biomasses and their C/N ratios (Table 2). Detrital concentrations were initialized to the observed POM concentrations minus the sum of all POM in biomass. Observed irradiance and temperature were used to drive the model.

Each simulation lasted 11 days. Discrete approximations to the governing equations were solved using the fourth order Runge-Kutta scheme, with a timestep of 2 min. We verified that the model conserved mass of nitrogen and that the simulated uptake of dissolved inorganic carbon matched the change in organic carbon.

The same parameter values were applied for both incubations (using 400 m and 700 m seawater, respectively), except for the perturbations to initial conditions, which were varied separately for each incubation. Thus, the only differences in the simulations are in the initial conditions of nutrients, biomasses and organic matter. This study tested whether our model could simulate the observed differences based only on differences in those initial conditions.

#### 2.5. Assimilation method

We applied the Monte Carlo Markov Chain (MCMC) method (Harmon and Challenor, 1997; Hargreaves and Annan, 2002). It performs a random walk through parameter space by iterating the following procedure: generate a trial set of parameters by randomly perturbing the current parameter set, run a simulation with the trial set of parameters, calculate a “cost” based on the difference between model and data, and decide whether to step to the trial parameter set (based solely on a comparison of this cost with that of the current parameter set). A step to a better parameter set (lower cost) is always taken, and a step to a worse one is sometimes taken based on a randomised decision, which

depends on how much worse the simulation is. This allows the method to proceed towards better simulations and avoid local minima. After running the walk to (statistical) convergence, the resulting ensemble of runs samples the joint posterior probability distribution function defined by the data.

To calculate the cost function we used the common approach of using the reciprocal of the estimated uncertainty of measurement (Section 2.1),  $\sigma_i$ , as a weight for each data type  $i$ , except as discussed below for POM. Variable perturbations were also introduced (as parameters) for the initial concentrations (IC) of selected organism compartments, org (the choice of which is discussed below). To constrain these perturbations,  $\varepsilon_{\text{org}}$ , a penalty was added to the cost function, applying a weight of  $w_{\text{IC}}$  divided by the uncertainty of measurement,  $\sigma_{\text{org}}$ . The total cost is then:

$$\text{Cost} = \sum_n \left[ \sum_i \left( \frac{C_{i,n}^{\text{sim}} - C_{i,n}^{\text{obs}}}{\sigma_i} \right)^2 + \left( \frac{\text{RPOM}_n^{\text{sim}} - \text{RPOM}_n^{\text{obs}}}{\sigma_{\text{RPOM},n}} \right)^2 \right] + \sum_{\text{org}} \left( \frac{w_{\text{IC}} \varepsilon_{\text{org}}}{\sigma_{\text{org}}} \right)^2 \quad (1)$$

in which  $C_i^{\text{sim}}$  is the simulated value corresponding to observed value  $C_i^{\text{obs}}$  and  $n$  sums over the observation times. The sum over org covers each organism compartment for which the IC are perturbed. The weight for the penalty was  $w_{\text{IC}} = 2.0 \cdot 10^3$ . All chemical and biological observations listed above (Section 2.1) were included in the cost function, except for silicon. This was because Si/N drawdown ratios varied, which would require variable Si/N ratios for diatoms. No silicon limitation was expected because silicic acid concentration was always greater than 30  $\mu\text{M}$ .

For POM, however, the weighting was adjusted because of the imbalance in total observed nitrogen (Section 2.1), which our mass-conservative model could never simulate. We have assumed that the nutrient measurements are correct and that the error lies in the POM measurements. Terms for PON were added to the cost function only when the model underestimated them (equivalent to setting  $\sigma_{\text{PON}} = \infty$  when  $\text{PON}^{\text{sim}} > \text{PON}^{\text{obs}}$ ). Instead of including POC directly, a cost was added for the bulk POC/PON ratio, with  $\sigma_{\text{RPOM},n} = \sigma_{\text{POC}}/\text{PON}_n$  (using the observed PON concentration at each sampling time,  $n$ ). Thus, information from the POC/PON ratio was incorporated into the assimilation (assuming that observed POC/PON ratios were correct despite the mass imbalance).

The extent to which the assimilation wanders (to avoid local minima) depends on the value of the cost relative to the log of a random variable generated for each simulation. In order to achieve a realistic spread in parameter values, we can adjust the cost function to account for model inadequacy (Annan et al., 2005). We do this in the simplest manner possible, dividing the total cost by a factor of 500 (a value selected 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. Because we do not attempt a quantitative probabilistic interpretation of our results, the somewhat arbitrary nature of the scaling does not affect our conclusions.

## 2.6. Parameters optimised

To test the robustness of the assimilation method and decide which parameters to vary, we performed “paired assimilations” as described below. We could have based our iterative selection process on identical twin tests (ITTs) (Spitz et al., 2001), but the strong non-linearity of these simulations (because nutrients are severely depleted) raised the concern that ITTs using sub-optimal initial guess simulations as “truth” might not represent the behaviour near the optimal solution. In other words, it might be possible to recover parameter values using the best-fit simulated values, but not using the initial guess simulated values or vice-versa.

We chose a candidate parameter subset based on a sensitivity analysis. First, we calculated the first-order sensitivities and the Hessian matrix (matrix of second derivatives) of the cost function with respect to all parameters. The eigenvectors of this matrix are the directions in parameter space along which the cost function is most sensitive. We then chose the parameters having the greatest contributions to the sensitivity of the cost function, down to an arbitrary cutoff of 1% (of the sum of squares of all eigenvalues).

Because the cost function does not include material flows (for which there are no observations), nitrogen regeneration may be sensitive to parameters that do not strongly impact the cost function. It may be possible to constrain such parameters to which the cost function is only moderately sensitive. We therefore calculated the first-order sensitivity of the cumulative (gross) nitrogen regeneration to each parameter. If a parameter had a great impact on the regeneration of nitrogen and a moderate effect on the cost function, it was included in the candidate parameter set.

Varying only that subset of parameter values, we conducted paired assimilations, consisting of two assimilations of the observations, one starting from their initial guess values and one from 70% of those values. If the final answer is independent of the initial guesses (at least for the values applied), each parameter’s mean value after convergence (MVAC) should be the same for both assimilations. Because the MCMC samples the distribution of parameter values without refining the cost function to hone in on the best-fit (as does simulated annealing, for example), the MVACs are more meaningful than the best-fit values. If the two assimilations did not reach close to the same MVACs within approximately 10 days of wall clock time, we cut down the candidate parameter set and repeated.

At each step, for any parameter for which the paired assimilations did not converge to the same value, we calculated correlations with all other parameters in the ensemble of accepted values from the assimilations. Any such parameter that correlated strongly with other parameters was eliminated from the candidate parameter subset, with its value fixed at its initial guess. Priors for the parameters were assumed to be uniform, except that all parameters were restricted to positive values, and those expressing fractional quantities (e.g., bacterial growth efficiency) were restricted to a [0,1] range. The scale for varying parameters at each step (proposal density) was set at 1% of each parameter’s initial guess value, which yielded an acceptance rate of approximately 4% in the assimilations. This value is lower than ideal for efficient convergence, but does not affect the final answer.

## 3. Results

### 3.1. Final subset of parameters

From an original candidate subset of 15 model parameters (out of 54) and six initial conditions for organism compartments (out of 12 total: six organism compartments times two incubations), we could robustly constrain only four parameters and six initial conditions (three compartments times two incubations). Both incubations were simulated at each step of the assimilation, so that all six initial conditions were constrained simultaneously. Best-fit values were close to the MVACs in all cases (Table 3).

After 10 million simulations each (requiring approximately 2 weeks of wall clock time), the paired assimilations yielded minimum costs of 249.8 and 250.1, respectively (initial costs were 1026 and 1187, respectively). For the following analysis, the ensemble of runs for each assimilation was constructed by taking

Table 3  
Parameters varied in the assimilation

Parameter	Initial value	Best	Mean	S.D.	Units
<i>For phytoplankton</i>					
$P_{\text{DiaL}}^{\text{ef}}$	13.0	9.67	9.72	0.286	day <sup>-1</sup>
<i>For bacteria</i>					
$\omega_B$	0.270	0.343	0.343	0.00729	[dimensionless]
<i>For heterotrophic nanoflagellates and zooplankton</i>					
$g_{p2\text{HNF}}$	9.0	12.2	11.8	0.531	day <sup>-1</sup>
$g_{p2Z}$	3.0	3.29	3.20	0.150	day <sup>-1</sup>
<i>Perturbations to initial concentration for organism compartments, 400 m case</i>					
$\epsilon_{\text{DiaS}}^{400}$	0.025	0.0545	0.0566	0.00568	$\mu\text{mol C L}^{-1}$
$\epsilon_{\text{HNF}}^{400}$	0.0025	0.00259	0.00256	0.000798	$\mu\text{mol C L}^{-1}$
$\epsilon_{\text{Zoo}}^{400}$	0.025	0.0129	0.0150	0.00273	$\mu\text{mol C L}^{-1}$
<i>Perturbations to initial concentration for organism compartments, 700 m case</i>					
$\epsilon_{\text{DiaS}}^{700}$	0.025	0.0718	0.0707	0.00795	$\mu\text{mol C L}^{-1}$
$\epsilon_{\text{HNF}}^{700}$	0.0025	0.0000297	0.000125	0.000150	$\mu\text{mol C L}^{-1}$
$\epsilon_{\text{Zoo}}^{700}$	0.025	0.0172	0.0190	0.00600	$\mu\text{mol C L}^{-1}$

every 10,000th run from the last 5 million simulations, after the MCMC had converged. The ensemble means were statistically the same for total cost and N

regeneration (Fig. 3) and for all parameter values (not shown). The cost after convergence was much lower than with the original guess parameter values. The flows

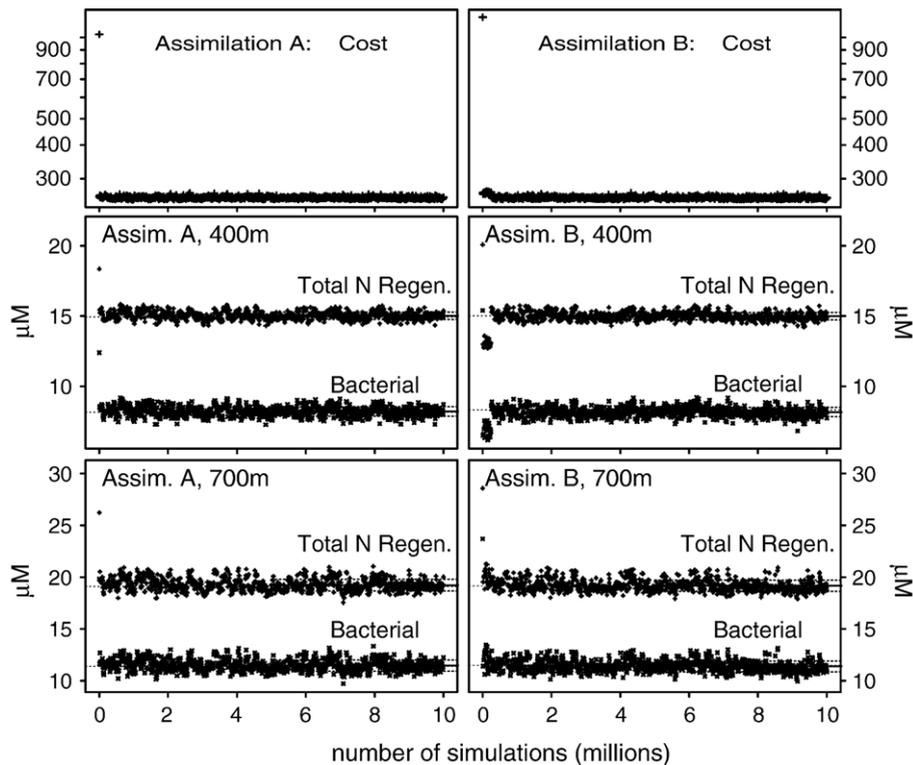


Fig. 3. Evolution of total cost (log scale) and simulated cumulative nitrogen regeneration in each incubation, for two assimilations (A: standard initial guesses for parameters varied in the assimilation and B: 70% of the initial guesses in A). Horizontal lines are mean (solid) and mean $\pm$ 1 S.D. (dotted) over the last half of the ensemble.

of nitrogen were well constrained, and we would draw the same conclusions using the results of either assimilation. Below we analyze the results of assimilation A, which yielded the lower cost.

Although we systematically eliminated parameters from our candidate subset until we could constrain all the parameters that were varied, due to computational constraints, we did not systematically add parameters to determine the maximum number that could be constrained by the data. Our goal was to constrain the flows of nitrogen and examine its regeneration, specifically in relation to the bulk POC/PON ratio. Even if it is possible to constrain more parameters, we do not believe that this would significantly enhance our conclusions (see the sensitivity analysis below).

### 3.2. Values for parameters

The MVAC and best-fit values for maximum photosynthesis rates (Table 3) were higher than those

in Geider et al. (1998), which were either 3.0 or 5.1 day<sup>-1</sup>. This is because these batch incubations started with abundant nutrients and very low phytoplankton concentrations. The model of Geider et al. (1998) was originally applied to culture experiments under conditions of balanced growth (when relative rates of change of phytoplankton C, N and Chl were equal). It also set all respiration and mortality rates to zero (except for a slow respiration rate in one case), so that the photosynthesis rates were net rates. Our (gross) rates should be greater to account for losses to respiration and mortality. Our rates are several times those in Geider et al. (1998) and are greater than those used by Lefevre et al. (2003), even though we use the same respiration rate coefficient. One reason is that our second-order rate expression for mortality results in greater losses than their first-order rate expression. The faster rates in this study are required to simulate the dynamics of these batch incubations. The observed imbalance in total nitrogen, which may result from

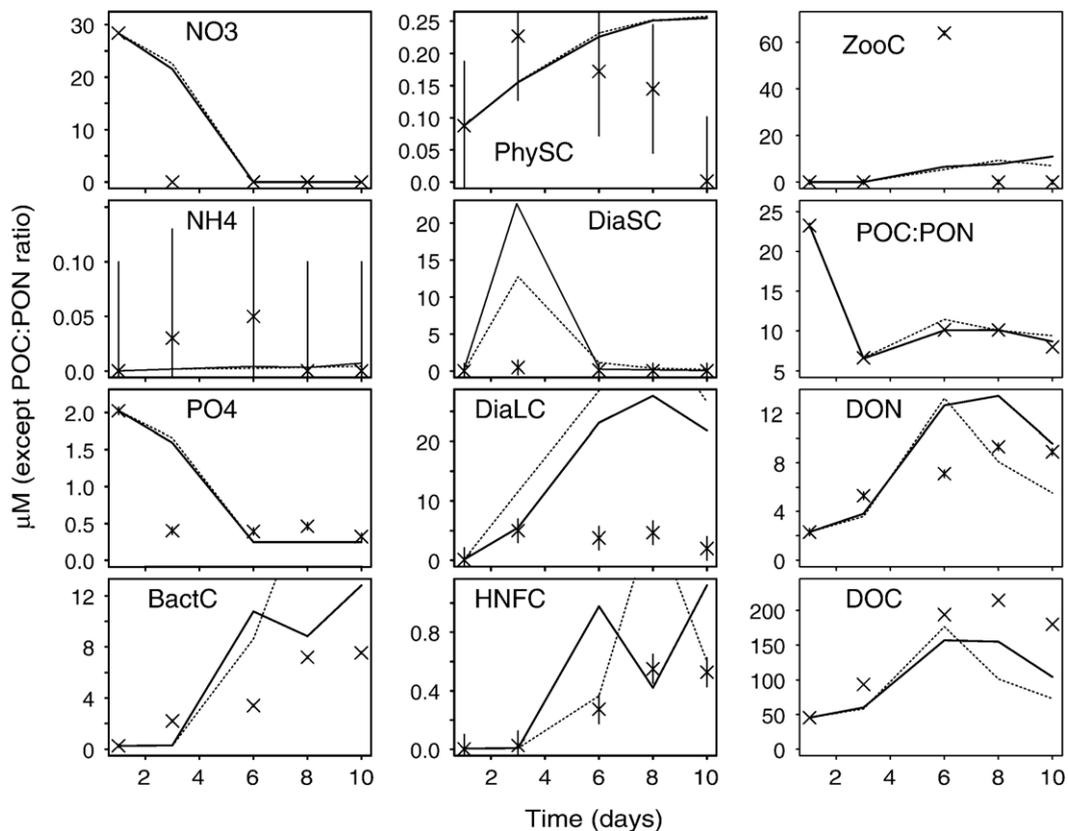


Fig. 4. Observations (X's) and simulations with the best-fit (solid lines) and initial guess (dotted lines) parameter sets for the 400 m incubation. Vertical axes are concentrations in  $\mu\text{M}$ , except for the POC/PON ratio which is dimensionless. Symbols for carbon biomasses end in C: BactC (bacteria), DiaSC (small diatoms), DiaLC (large diatoms), HNFC (heterotrophic nanoflagellates) and ZooC (zooplankton). Others are as defined in the text. Vertical lines are error bars.

biomass growing on the walls of the containers, also likely contributed to the high values for these rates.

### 3.3. Best-fit simulations

For the most part, the best-fit simulations are visibly closer to the observations than those using the initial guesses for parameter values (Figs. 4 and 5). The simulated drawdown of nitrate is too slow. The trends in biomass over time for diatoms are qualitatively well simulated, although the biomasses are overestimated. The model overestimates PON and POC (Fig. 6), especially midway through each incubation, because of the lack of a penalty for overestimating these data (Section 2.5). This is because the observed PON and POC concentrations (and likely the biomasses) are too low (Section 2.1). The model cannot track the changes in total nitrogen.

Previous paired assimilations (not shown) which varied the mortality rate for diatoms similarly overestimated diatom biomass (but not all parameters converged). We can only speculate that the missing

nitrogen, silicon and (presumably) carbon could have been in biomass growing on the walls of the container, or in material that was not uniformly suspended in the water. We cannot determine how much of this discrepancy resulted from errors in the model versus errors in the observations.

In both the data and the simulations for both incubations, biomass of small diatoms peaks early, but biomass of large diatoms is higher at the end. The half-saturation coefficients for uptake were the same for small and large diatoms, consistent with the findings of Niraula et al. (2005). Changes in the simulated C/N ratio of diatoms dominated the changes in bulk POC/PON ratio over time.

The best-fit simulations of POC/PON ratio track the observations closely, consistent with the high weight given to the ratio (based on the uncertainties of measurement for POC and PON). POC/PON ratios were lower early in the incubations when nutrients were replete and higher after nitrate was depleted. Our model simulates these trends and the difference between the two incubations (Fig. 6).

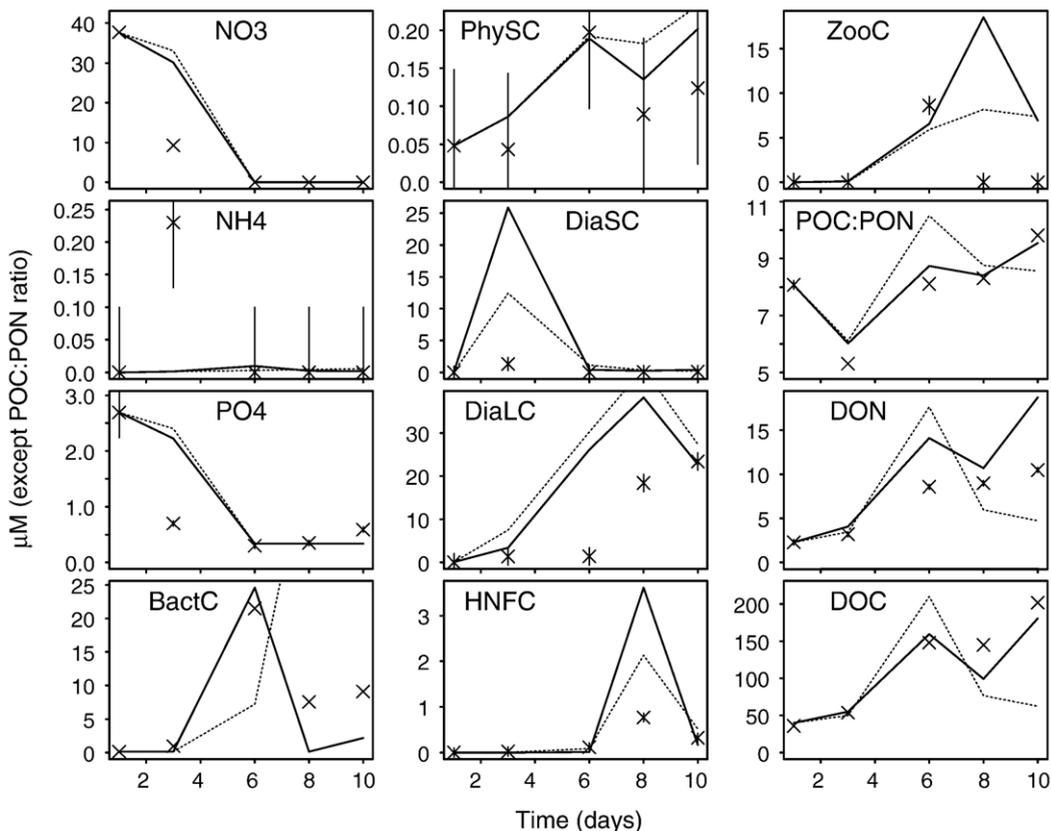


Fig. 5. Observations (X's) and simulations with the best-fit (solid lines) and initial guess (dotted lines) parameter sets for the 700 m incubation (notation is the same as in Fig. 4).

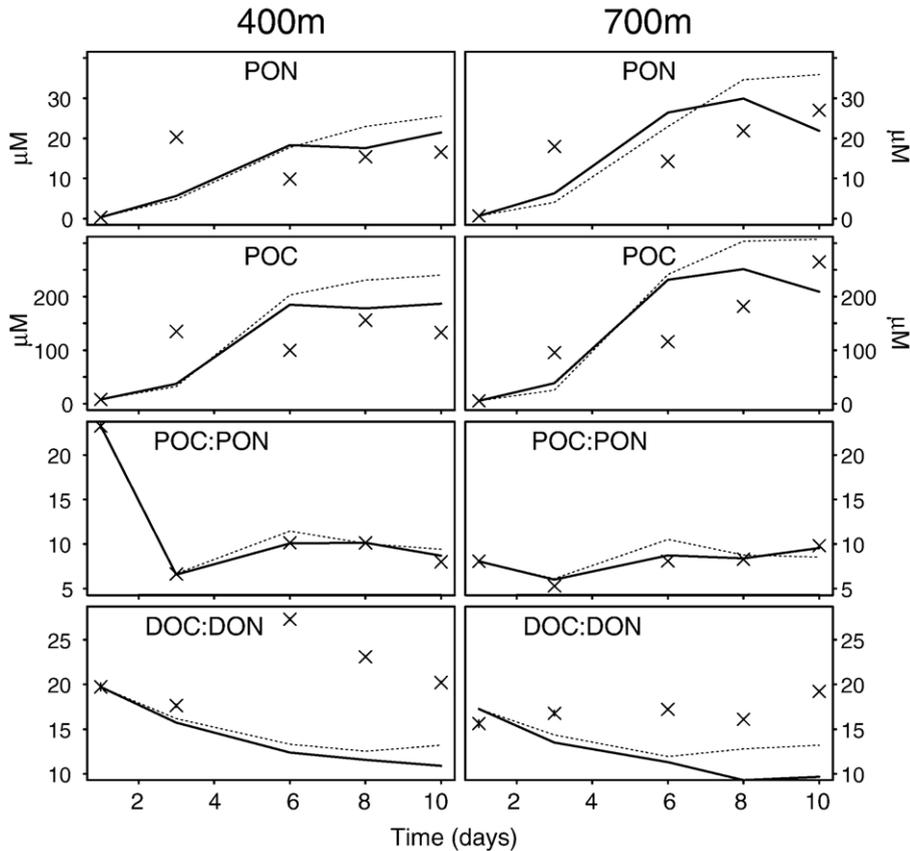


Fig. 6. Bulk POC, PON, POC/PON ratio and DOC/DON ratio for the 400 m (left) and 700 m (right) incubations. X's are observations. Solid (dotted) lines are the best-fit (initial guess) simulations.

### 3.4. Dissolved organic matter

The model overestimates DON and underestimates DOC concentrations for both incubations, resulting in DOC:DON ratios that are consistently too low (Fig. 6). Our model assumes that phytoplankton excrete DOM with the same stoichiometry as their growth, and that other sources and sinks of DOM are determined by mass balance among the compartments. The assimilation was not able to find an accurate simulation of the observed DOC/DON ratios, showing that these assumptions are not consistent with the observations.

### 3.5. Regeneration

Simulated regeneration of nitrogen is greater in the 700 m case (Fig. 3). Bacteria are the primary regenerators of nitrogen in both incubations. Regeneration by bacteria in the best-fit simulation of the 700 m incubation is 39% greater than in the 400 m case (Fig. 2). This is because the simulated DOC:DON ratios in the 400 m case, although still well below the observa-

tions, were higher than those in the 700 m case (Fig. 6). With their C-based growth efficiency in this model, bacteria consuming more C-rich DOM use more N for growth and therefore regenerate less. Simulated gross uptake of nitrogen was also greater in the incubation using 700 m water (Fig. 7). The quantiles (5% and 95%) of simulated values reveal that this difference was robust (present in the vast majority of simulations in the ensemble).

### 3.6. Sensitivity of flows

To assess the extent to which parameters not varied in the assimilation may have effected our results, we calculated the sensitivities of total nitrogen regeneration to each model parameter. Results are summarized in Table 4. Each parameter was varied by  $\pm 25\%$ . The flows were sensitive to heterotrophic parameters (growth and assimilation efficiencies and grazing rates).

Of the 17 parameters that most affected the regeneration of nitrogen (the union of the two 15-member subsets that affected the flows in the 400 m

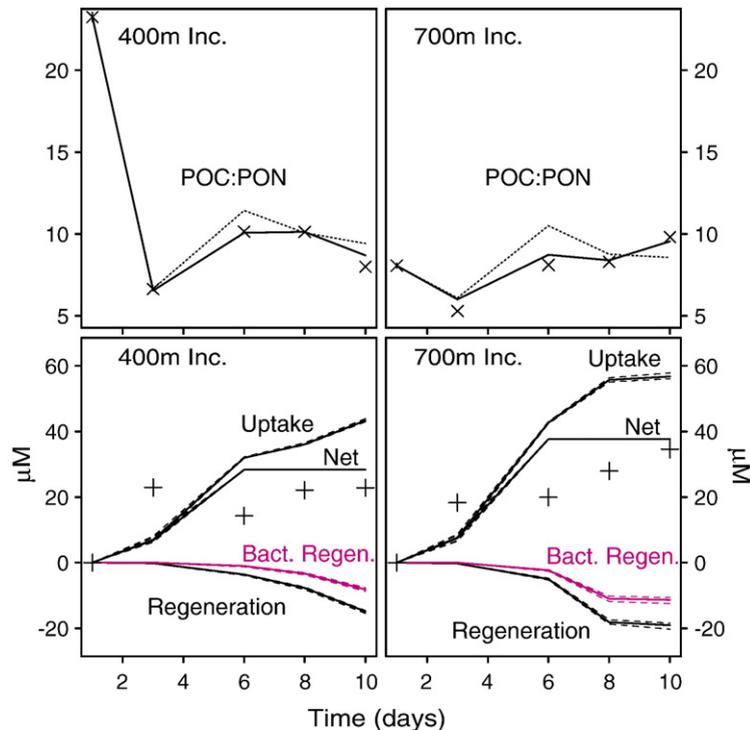


Fig. 7. Above: bulk POC/PON ratios; X's are observations and solid (dotted) lines are the best-fit (initial guess) simulations. Below: simulated gross flows of nitrogen (uptake and regeneration) and their sum (net organic nitrogen production, which is compared to observations); solid lines are the best-fit simulation, and dotted lines are the 5% and 95% quantiles from the ensemble of simulated values.

and 700 m cases, respectively), only three were varied in the assimilation. This is because the sensitivities of the cost function (which determines whether parameters can be constrained by the assimilation) differ from those of the nitrogen flows. However, the regeneration of nitrogen was without exception greater in the incubation using 700 m water. Bacterial regeneration was greater than half of total regeneration in nearly all cases and contributed a minimum of 36% of total regeneration.

This first-order analysis ignores the combined effects of differences in multiple parameters. Previous test assimilations which included more parameters yielded similar results for nitrogen flows and bulk POC/PON ratio, even though it was not possible to constrain the values of all parameters. It is therefore unlikely that moderate changes in the fixed parameter values would alter our conclusions.

#### 4. Discussion

Because the ecosystems in these incubations differ from those in the surface ocean, the values of parameters determined in this study are not necessarily applicable to

the real ocean. However, the assimilation constrained the flows of nitrogen (Figs. 3 and 7), and we can draw conclusions based on the similarities and differences between the two incubations that were consistently simulated.

The major advantages of using mesocosm experiments for this kind of study is that (at least ideally) they provide a local mass balance, controlled conditions and frequent, detailed measurements. The disadvantage of the effects of the container walls (Vallino, 2000) seems likely to have contributed to the mass imbalance in these experiments, however. The short duration of the time series and the low initial biomasses also meant that uncertainty in the initial conditions was significant. This required that we include perturbations to initial conditions for biomasses in the assimilation, which reduced the number of model parameters that could otherwise be constrained. When applied with longer time series of field observations (Spitz et al., 1998, 2001), data assimilation can constrain more parameters. Such an approach cannot, however, examine the short-term dynamics of ecosystems during the exponential growth phase, which are important for phytoplankton blooms.

Table 4  
Sensitivities of N regeneration to model parameters

Parameter <sup>a</sup>	Varied <sup>b</sup>	400 m incubation		700 m incubation	
		Regen.	% Bact.	Regen.	% Bact.
$\mu_B$	No	-18.8 (-8.46)	62.1 (40.7)	-27.4 (-16.6)	70.8 (56.6)
$\beta_{\text{HNF}}$	No	-9.26 (-17.3)	50.4 (61.8)	-15.6 (-24.4)	63.3 (71.2)
$m_d$	No	-12.9 (-16.6)	53.2 (56.3)	-18.8 (-25.3)	62.8 (65.8)
$k_{d,\text{POM}}$	No	-18.2 (-10.6)	61.5 (45.5)	-22.3 (-16.0)	62.5 (56.0)
$m_Z$	No	-11.0 (-16.1)	46.0 (61.6)	-18.1 (-24.4)	58.8 (72.1)
$g_{p2\text{HNF}}$	Yes	-8.84 (-15.4)	41.8 (66.0)	-16.6 (-22.3)	58.0 (74.4)
$\lambda_I$	No	-16.0 (-10.4)	64.2 (44.8)	-23.3 (-17.9)	73.7 (58.6)
$g_{p2Z}$	Yes	-16.5 (-6.69)	65.8 (35.6)	-23.9 (-18.6)	74.5 (58.0)
$\alpha_Z$	No	-17.3 (-12.5)	52.8 (57.6)	-21.9 (-16.9)	56.4 (64.2)
$\gamma_{\text{DiaS}}$	No	-16.9 (-11.4)	59.3 (51.1)	-22.2 (-18.2)	63.4 (59.4)
$P_{\text{DiaL}}^{\text{ref}}$	Yes	-13.5 (-15.5)	51.0 (58.6)	-18.3 (-22.1)	57.0 (65.3)
$Q_{\text{max,DiaS}}$	No	-12.5 (-13.6)	57.1 (49.4)	-20.4 (-17.2)	64.7 (54.5)
$\alpha_{\text{mat hit H NF}}$	No	-16.1 (-13.8)	45.5 (65.5)	-20.7 (-17.5)	52.2 (68.7)
$\beta_Z$	No	-15.7 (-8.32)	66.8 (36.8)	-23.3 (-20.3)	75.8 (55.6)
$R_i^0$	No	-15.9 (-14.0)	52.6 (57.0)	-20.1 (-18.1)	57.5 (62.0)
$p_{\text{min}}$	No	-13.5 (-15.5)	51.7 (57.2)	-18.7 (-19.6)	59.2 (60.1)
$\zeta_{N,i}$	No	-14.2 (-12.3)	53.9 (54.8)	-19.1 (-19.7)	61.3 (63.4)

Cumulative N regeneration (Regen.) and the percentage of it mediated by bacterial (% Bact.) are given for perturbations to each parameter listed (relative to best-fit parameter values). Values without parentheses are for a +25% perturbation. Those within parentheses are for a -25% perturbation.

<sup>a</sup> These parameters constitute the union of the two subsets of 15 parameters each to which total N regeneration was most sensitive for each incubation. They are in decreasing order of sensitivity for the 700 m case.

<sup>b</sup> Whether the parameter was varied (included) in the assimilation.

#### 4.1. POC/PON ratios

Because the ecosystem model includes all major planktonic groups present, we could compare our simulated bulk POM (detritus plus organisms) directly with the observed, bulk POM. In our simulations, changes in phytoplankton composition dominated the changes in bulk POC/PON ratio, as was found by Mongin et al. (2003) in a study simulating station BATS.

The phytoplankton model can only produce biomass with high C/N ratios when nitrogen supply is limited. POM stoichiometry therefore constrains nitrogen regeneration in the assimilations. The higher bulk POC/PON ratios for most of the latter half (days 6 and 8) of the 400 m incubation limit regeneration more than the lower values for the 700 m incubation (Fig. 6). On the last sampling day (day 10), the POC/PON ratio is lower for the 400 m case, meaning that there was an increase in regeneration at the end of the 400 m incubation (Fig. 7). The simulated C/N ratio of phytoplankton is very sensitive to changes in nitrogen supply, because of the energetic cost of assimilating nitrogen (parameter  $\zeta_N$ ). The more nitrogen phytoplankton assimilate, the more organic carbon they respire, further lowering their C/N ratio (Geider et al., 1998). This sensitivity of phytoplankton composition to nitrogen supply allowed the assimilation to constrain the flows of nitrogen.

#### 4.2. DOC/DON ratios and bacterial regeneration

The greater bacterial regeneration in the 700 m incubation is consistent with the observation of more DOM and bacterial biomass in that case. We suspect that this difference between the two incubations resulted from differences in phytoplankton assemblages: specifically, the dominance of the small diatom *Thalassiosira minima* early in the 700 m incubation. Copious amounts of DOM have been observed when this species dominates in marine waters (B. Casareto, unpublished observations).

Because our model underestimated the DOC/DON ratios in both incubations, we expect that it overestimates bacterial regeneration of nitrogen. High DOC/DON ratios (exceeding POC/PON ratios) are often observed in the oligotrophic ocean (Williams, 1995) and in incubations (Vallino, 2000). We are not aware of any study that has accurately simulated this “excess” DOC using a mechanistic model.

Still, the relative difference in simulated DOC/DON ratios (higher in the 400 m case) agrees with the observations (Fig. 6). This is important to the successful simulations of POC/PON ratios. With their C-based growth efficiency in the model, bacteria consuming more C-rich DOM must retain more N for growth and regenerate less, so that phytoplankton

grow with higher C/N ratios in the 400 m incubation. A study simulating the Hawaii Ocean Time-Series Stn. ALOHA (Smith et al., 2005) found that this same effect was key for simulating the high TOC/TON remineralization ratio (from an independent data-based estimate). Because stoichiometry of DOM excretion by phytoplankton is proportional to that of POM production, this results in a feedback in our simulations, for which bacterial regeneration is the primary source of DIN. Simulations (not shown) with an oceanic mixed layer model, however, found that this feedback was not important because other sources of nitrogen dominated.

Thus, bacteria, through variations in nitrogen regeneration, competed indirectly with phytoplankton for ammonium, even though bacteria never became nitrogen-limited in these simulations. In our model, bacteria become nitrogen-limited when  $\text{DOC}^{\text{nr}}/\text{DON}^{\text{nr}} > \text{RCN}_B/\omega_B$ , which for the best-fit value of  $\omega_B$  (0.34), =  $5.1/0.34=15$ . Simulated  $\text{DOC}^{\text{nr}}/\text{DON}^{\text{nr}}$  ranged from 8 to 12 in our best-fit simulations. For comparison, we can calculate the non-refractory concentrations of DOM and  $\text{DOC}^{\text{nr}}/\text{DON}^{\text{nr}}$  ratios in our incubations, based on the bulk observations and assumed refractory concentrations. The minimum  $\text{DOC}^{\text{nr}}/\text{DON}^{\text{nr}}$  was 15 and the maximum was 32. Therefore, if simulated bulk DOC/DON approached the observations, bacteria would likely be nitrogen-limited during part of the simulations. The best-fit value of  $\omega_B$  would almost certainly be lower, which would give a higher threshold for bacterial N limitation. Our simulations of bacterial regeneration are consistent with experiments by Goldman et al. (1987), who measured regeneration of nitrogen by bacteria in incubations for DOC/DON ratios as high as 10 (the highest value examined).

## 5. Conclusions

We have successfully implemented the MCMC method to constrain flows in a complex multi-element ecosystem model. Our ecosystem model, by explicitly simulating carbon and nitrogen cycles, was able to consistently simulate the changes in bulk POC/PON ratio over time, and the differences between the two incubations. The variable C/N ratio of phytoplankton was the key to simulating the observed changes in POC/PON ratio and allowed our assimilation to constrain the regeneration of nitrogen based on that ratio. This was possible despite the lack of observed mass balance of nitrogen. Bacteria dominated the regeneration of nitrogen in both incubations. Differences in bacterial regeneration between the two

incubations were the primary reason for the differences in C/N ratio of phytoplankton and bulk POC/PON ratio.

Ecosystem models including quota models for the variable C/nutrient ratios of phytoplankton have recently been applied to Stn. BATS near Bermuda (Mongin et al., 2003) and Stn. ALOHA near Hawaii (Christian, 2005), yielding improved simulations compared to models using fixed Redfield ratios. Our results suggest that, with such models, simultaneous observations of POC and PON could be used to constrain simulated flows of nitrogen in the oligotrophic ocean (e.g., during blooms vs. stratified conditions). This will require that the models include all major planktonic groups, so that simulated bulk POC and PON can be compared with observations.

## Acknowledgments

We thank the staff of the Shizuoka Prefectural Fisheries Laboratory for hosting and assisting with the experiments, Richard Geider for advice about embedding the phytoplankton model into our ecosystem model, and Akari Furuta for the DOM measurements. We thank the reviewers of this and previous manuscripts for their thorough reviews and Yvette Spitz for a helpful discussion.

## Appendix A. Model description

The model of Geider et al. (1998), which simulates the C, N and Chl of biomass separately, is applied for each phytoplankton class. Stoichiometries of all other organism compartments are fixed, and only their nitrogen concentrations are simulated. For each class of non-living organic matter, two elemental fractions are simulated: carbon and nitrogen. Thus, the stoichiometries of dissolved organic matter (DOM) and detrital (particulate) organic matter (DetM) can vary. Their elemental fractions are denoted by  $\text{DOX}$  and  $\text{DetX}$ , where  $X=C$  or  $N$ .

Other quantities are either terms for processes (defined below) or parameters. Variables for processes are:  $P_i^{\text{hot}}$  for gross primary production by phytoplankton  $i$ ;  $L_i$  for light-limitation of  $i$ ;  $R_x$  for respiration by organism  $x$  or regeneration of compartment  $x$ ;  $E_i$  for excretion of DOM by phytoplankton  $i$ ;  $M_O$  for mortality of organism  $O$ ;  $G(i,Z)$  for grazing of prey  $i$  by zooplankton  $Z$ ;  $E_{gz}$  for egestion of fecal pellets by  $Z$ ;  $\text{Slf}_Z$  for sloppy feeding (producing DOM) by  $Z$ ;  $D_{\text{DetX}}$  for dissolution of DetM to form DOM;  $\Theta_T$  for the temperature dependence of all rates; and  $\text{RCN}_x$  and

$RNP_x$  for the molar C/N and N/P ratio, respectively, of compartment  $x$ .

Zooplankton processes are formulated mostly as in the NEMURO (North Pacific Ecosystem Model for Understanding Regional Oceanography) model (Fuji et al., 2002a; Eslinger et al., 2000), and the carbon cycle is treated as in Smith et al. (2005). The formulation for bacteria and DOM cycling is based on that of Anderson and Williams (1998). The mass balance equations for each compartment follow.

For the C component of small, non-silicious phytoplankton,  $PhyS_C$ :

$$\frac{\partial PhyS_C}{\partial t} = \left\{ P_{PhyS}^{phot} - R_{PhyS}^C \right\} PhyS_C - E_{PhyS}^C - \frac{\{M_{PhyS} + G(PhyS, HNF) + G(PhyS, Z)\}}{Q_{PhyS}} \quad (A.1)$$

where  $P_{PhyS}^{phot}$  is the gross rate of organic carbon production (photosynthesis),  $R_{PhyS}^C$  is the C respiration rate,  $-E_{PhyS}^C$  is the rate of DOC excretion,  $M_{PhyS}$  is the mortality rate (N-based),  $G(PhyS, HNF)$  is the rate of grazing by heterotrophic nanoflagellates (HNF),  $G(PhyS, Z)$  is the rate of grazing by zooplankton (Z), and  $Q_{PhyS}$  is the N/C quota for  $PhyS$ . For the N component:

$$\frac{\partial PhyS_N}{\partial t} = \frac{\{V_{C,PhyS}^N - R_{PhyS}^N\}}{Q_{PhyS}} PhyS_N - E_{PhyS}^N - \{M_{PhyS} + G(PhyS, HNF) + G(PhyS, Z)\} \quad (A.2)$$

where  $V_{C,PhyS}^N$  is the rate of uptake of dissolved inorganic nitrogen per mole of C biomass,  $R_{PhyS}^N$  is the rate of N respiration, and  $E_{PhyS}^N$  is the rate of excretion of DON. For the Chl component:

$$\frac{\partial PhyS_{Chl}}{\partial t} = \left\{ \frac{\rho_{PhyS}^{Chl} \cdot V_{C,PhyS}^N}{\Theta_{PhyS}^C} - R_{PhyS}^{Chl} \right\} PhyS_{Chl} - \frac{\{M_{PhyS} + G(PhyS, HNF) + G(PhyS, Z)\} \Theta_{PhyS}^C}{Q_{PhyS}} \quad (A.3)$$

where  $\rho_{PhyS}^{Chl}$  is the Chl-synthesis regulating term,  $R_{PhyS}^{Chl}$  is the rate of respiration of Chl, and  $\Theta_{PhyS}^C$  is the ratio of Chl/C for  $PhyS$ .

Equations for diatoms, DiaS and DiaL, are exactly analogous, except for grazing terms. Heterotrophic nanoflagellates (HNF) graze bacteria (B) and small phytoplankton ( $PhyS$ ). Zooplankton (Z) graze small phytoplankton ( $PhyS$ ), small diatoms (DiaS) and

heterotrophic nanoflagellates (HNF). The model does not include zooplankton large enough to eat the large diatoms, DiaL, and so their only loss is mortality. Processes are listed below.

The rate of photosynthesis for each phytoplankton class,  $i$ , is:

$$P_i^{phot} = P_i^{max} \cdot \left( 1 - \exp \left\{ \frac{-\alpha_i^{Chl} \Theta_i^C I^0}{P_i^{max}} \right\} \right) \quad (A.4)$$

in which  $P_i^{max}$  is the maximum possible photosynthetic rate:

$$P_i^{max} = P_i^{ref} \cdot \left( \frac{Q_i - Q_{min,i}}{Q_{max,i} - Q_{min,i}} \right) \Theta_T \quad (A.5)$$

$\alpha_i^{Chl}$  is the Chl-specific light adsorption coefficient,  $\Theta_i^C$  is the Chl-to-C ratio of phytoplankton  $i$ ,  $I$  is the light intensity,  $Q_i$  is the N/C quota of phytoplankton  $i$ , and  $Q_{min,i}$  and  $Q_{max,i}$  are the minimum and maximum quotas, respectively. Uptake of DIN is  $V_{C,i}^N = V_{C,i}^N f_{upt,i}$ , where  $f_{upt,i}$  is the term for limitation of uptake by external nutrient concentrations. For small phytoplankton,  $PhyS$ , and diatoms,  $d$  (=DiaS, DiaL), the external nutrient limitations are:

$$f_{upt,PhyS} = \min \left( \frac{NO_3}{K_{NO_3,PhyS} \cdot e^{-\Phi_{PhyS} \cdot NH_4} + NO_3} + \frac{NH_4}{K_{NH_4,PhyS} + NH_4} \right) \quad (A.6)$$

$$f_{upt,d} = \min \left( \frac{NO_3}{K_{NO_3,d} \cdot e^{-\Theta_d \cdot NH_4} + NO_3} + \frac{NH_4}{K_{NH_4,d} + NH_4} \right) \times \left( \frac{DSi}{K_{DSi,d} + DSi} \right) \quad (A.7)$$

The maximum possible rate of DIN uptake is:

$$V_{C,i}^{N,max} = v_{C,i}^{N,ref} \Theta_T \left[ \frac{Q_{max,i} - Q_i}{Q_{max,i} - Q_{min,i}} \right]^{0.05} \quad (A.8)$$

where  $v_{C,i}^{N,ref}$  is the maximum possible rate at temperature  $T_{ref}$ . The term for Chl-synthesis is:

$$\rho_i^{Chl} = \Theta_i^{N,max} P_i^{phot} / (\alpha_i^{Chl} \Theta_i^C I) \quad (A.9)$$

in which parameter  $\Theta_i^{N,max}$  is the maximum possible Chl-to-N ratio for phytoplankton  $i$ .

Rates of phytoplankton processes depend on temperature:

$$\Theta_T = \exp \left\{ -E_a R_{\text{gas}} \cdot \left( \frac{1}{T_{\text{ref}}} - \frac{1}{T} \right) \right\} \quad (\text{A.10})$$

where  $E_a$  is the energy of activation,  $R_{\text{gas}}$  is the gas constant and  $T_{\text{ref}}$  is a reference temperature (at which the rate is specified).

Respiration rates of the C, N and Chl in phytoplankton,  $i$  (=PhyS, DiaS or DiaL), [M nitrogen day<sup>-1</sup>] are:  $R_i^C = r_i^C \Theta_T i C - \zeta_N V_{C,i}^N$ ,  $R_i^N = r_i^N \Theta_T i N$  and  $R_i^{\text{Chl}} = r_i^{\text{Chl}} \Theta_T i^{\text{Chl}}$ , where the  $r_i$ 's are the parameters for each specific rate of respiration at temperature  $T_{\text{ref}}$ , and  $iC$ ,  $iN$  and  $i\text{Chl}$  are respectively the carbon, nitrogen and Chl concentrations of phytoplankton  $i$ . Excretion of DON by phytoplankton  $i$  [M nitrogen day<sup>-1</sup>] is  $E_i = \gamma_i \cdot P_i^{\text{phot}}$ . Mortality of organisms (except bacteria), Org (=PhyS, DiaS, DiaL, HNF or Z), is  $M_{\text{org}} = m_{\text{org}} \cdot \Theta_T \cdot \text{Org}^2$ , in [M nitrogen day<sup>-1</sup>], where  $m_{\text{org}}$  is the mortality rate coefficient for each Org.

Grazing of prey,  $p$ , by heterotrophs,  $h$  (=HNF, Z) [M nitrogen day<sup>-1</sup>] is:

$$G(p, h) = g_{p2h} \cdot \Theta_T \cdot \{ \max 0.0, [1.0 - \exp\{\lambda_1 \cdot (p_{\text{min}} - p)\}] \} \cdot h \quad (\text{A.11})$$

in which  $g_{p2h}$  is the maximum grazing rate of  $p$  by  $h$  (at  $T_{\text{ref}}$ ),  $\lambda_1$  is the Ivlev coefficient and  $p_{\text{min}}$  is the minimum concentration for grazing of  $p$ .

For heterotrophic nanoflagellates, HNF, and zooplankton, Z:

$$\frac{\partial \text{HNF}}{\partial t} = G(\text{B}, \text{HNF}) + G(\text{PhyS}, \text{HNF}) - G(\text{HNF}, \text{Z}) - \text{Ege}_{\text{HNF}}^N - R_{\text{HNF}}^N - \text{Slf}_{\text{HNF}}^N - M_{\text{HNF}} \quad (\text{A.12})$$

$$\frac{\partial \text{Z}}{\partial t} = G(\text{HNF}, \text{Z}) + G(\text{PhyS}, \text{Z}) + G(\text{DiaS}, \text{Z}) - \text{Ege}_Z^N - R_Z^N - \text{Slf}_Z^N - M_Z \quad (\text{A.13})$$

For each heterotroph,  $h$  (=HNF or Z), mortality,  $M_h$  defined above for phytoplankton.  $\text{Ege}_h$  represents egestion of fecal pellets (or DOM for HNF),  $R_h$  represents respiration (regeneration) and  $\text{Slf}_h$  represents sloppy feeding (including all excretion of DOM), all in [M nitrogen day<sup>-1</sup>]. Each is calculated as a fraction of non-assimilated C or N, with the fractions being:  $f_h^{\text{Ege}} = (1 - \delta_S)(1 - \alpha_h)$ ,  $f_h^R = (\alpha_h - \beta_h)$  and  $f_h^{\text{Slf}} = \delta_S(1 - \alpha_h)$ , where

$\alpha_h$  is the assimilation efficiency,  $\beta_h$  is the growth efficiency and  $\delta_S$  is the fraction of sloppy feeding, =  $\text{Slf}_h / (\text{Slf}_h + R_h)$ . If C in the prey is sufficient to support growth with growth efficiency  $\beta_h$ :

$$\sum_{p_h} [\text{RCN}_{p_h} G(p_h, h)] > \text{RCN}_h \beta_h \sum_{p_h} G(p_h, h) \quad (\text{A.14})$$

then non-assimilated N is:

$$N_h^{\text{na}} = (1 - \beta_h) \sum_{p_h} G(p_h, h) \quad (\text{A.15})$$

where  $p_h$  sums over all prey for  $h$ . Otherwise, if the C in the prey is insufficient to support growth with growth efficiency  $\beta_h$ , less nitrogen is assimilated. In this case, non-assimilated N and C are:

$$N_h^{\text{na}} = \sum_{p_h} G(p_h, h) - \frac{1}{\text{RCN}_h} \sum_{p_h} [\text{RCN}_{p_h} G(p_h, h)] \quad (\text{A.16})$$

$$C_h^{\text{na}} = \max \left\{ \sum_{p_h} [\text{RCN}_{p_h} G(p_h, h)] - \text{RCN}_h \beta_h \sum_{p_h} G(p_h, h), 0.0 \right\} \quad (\text{A.17})$$

Egestion, respiration and sloppy feeding of element  $X$  (=N or C) are then:  $\text{Ege}_h^X = f_h^{\text{Ege}} \cdot X_h^{\text{na}}$ ,  $R_h^X = f_h^R \cdot X_h^{\text{na}}$  and  $\text{Slf}_h^X = f_h^{\text{Slf}} \cdot X_h^{\text{na}}$ .

Mass balances for nutrients (except silicic acid) follow. Here sums over  $i$  count all phytoplankton classes ( $i$ =PhyS, DiaS or DiaL). Sums over  $h$  count HNF and Z.

$$\frac{\partial \text{NH}_4}{\partial t} = \sum_i \left\{ R_i - (1 - f_{\text{new},i}) \cdot P_i^{\text{phot}} \right\} + \sum_h R_h^N + R_{\text{DON}} + R_{\text{DetN}} - R_{\text{Nit}} \quad (\text{A.18})$$

$$\frac{\partial \text{NO}_3}{\partial t} = - \sum_i \left\{ f_{\text{new},i} \cdot P_i^{\text{phot}} \right\} + R_{\text{Nit}} \quad (\text{A.19})$$

$$\frac{\partial \text{PO}_4}{\partial t} = \frac{1}{\text{RNP}} \frac{\partial \text{NO}_3}{\partial t} + \frac{\partial \text{NH}_4}{\partial t} \quad (\text{A.20})$$

$R_{\text{Nit}} = k_{\text{Nit}} \cdot \Theta_T \cdot \text{NH}_4$  is nitrification (oxidation of  $\text{NH}_4$  to  $\text{NO}_3$ ) [M nitrogen day<sup>-1</sup>].  $f_{\text{new},i}$  is the fraction of production as  $\text{NO}_3$  for each phytoplankton,  $i$  (=PhyS, DiaS or DiaL):

$$f_{\text{new},i} = \frac{\left\{ \frac{\text{NO}_3}{K_{\text{NO}_3,i} + \text{NO}_3} \cdot e^{-\theta_i \cdot \text{NH}_4} \right\}}{\left\{ \frac{\text{NO}_3}{K_{\text{NO}_3,i} + \text{NO}_3} \cdot e^{-\theta_i \cdot \text{NH}_4} + \frac{\text{NH}_4}{K_{\text{NH}_4,i} + \text{NH}_4} \right\}} \quad (\text{A.21})$$

Regeneration,  $R_{\text{DOX}}$ , for DOM depends on bacteria (below). For POM, where  $X$  (=C or N) represents the elemental fraction, regeneration is  $R_{\text{DetX}} = k_{r,\text{DetM}} \cdot \Theta_T \cdot \text{DetX}$  and dissolution is  $R_{\text{DetX}} = k_{d,\text{DetM}} \cdot \Theta_T \cdot \text{DetX}$ , where  $k_{r,\text{DetM}}$  and  $k_{d,\text{DetM}}$  are the respective rate coefficients. DetM is produced by mortality ( $o$  sums over DiaS, DiaL and Z below), and by egestion.

$$\frac{\partial \text{DetC}}{\partial t} = \sum_o \{ \text{RCN}_o \cdot M_o \} + \text{Ege}_{\text{Zoo}}^{\text{C}} - D_{\text{DetC}} - R_{\text{DetC}} \quad (\text{A.22})$$

$$\frac{\partial \text{DetN}}{\partial t} = \sum_o M_o + \text{Ege}_{\text{Zoo}}^{\text{N}} - D_{\text{DetN}} - R_{\text{DetN}} \quad (\text{A.23})$$

Mass balances for dissolved silicic acid and biogenic silica (BSi) follow.

$$\frac{\partial \text{DSi}}{\partial t} = - \sum_d \left\{ \left( P_d^{\text{phot}} - R_d - E_d \right) \cdot \text{RSiN}_d \right\} + R_{\text{BSi}} \quad (\text{A.24})$$

$$\frac{\partial \text{BSi}}{\partial t} = \sum_d \{ [M_d + G(d, h)] \cdot \text{RSiN}_d \} - R_{\text{BSi}} \quad (\text{A.25})$$

where  $d$  sums over DiaS and DiaL. BSi is produced by mortality of diatoms, and all BSi ingested by zooplankton is assumed to be egested in fecal pellets. Dissolution of BSi is:  $R_{\text{BSi}} = k_{r,\text{BSi}} \cdot \Theta_T \cdot \text{BSi}$  [M silicon  $\text{day}^{-1}$ ], where  $k_{r,\text{BSi}}$  is the rate coefficient.

For bacteria and DOM, the formulation is based on Anderson and Williams (1998), simplified to only one class of (non-refractory) DOM by assuming that the labile fraction is consumed instantaneously by bacteria. We use a first-order (in DOM) rate expression (parameters in Table 3). Refractory DOM is assumed constant. For bacteria, B:

$$\frac{\partial \text{B}}{\partial t} = \text{GHP}_B - M_B - G(\text{B}, \text{HNF}) \quad (\text{A.26})$$

where  $\text{GHP}_B$  is the gross heterotrophic production (gross growth rate) of bacteria,  $M_B$  is their mortality rate and  $G(\text{B}, \text{HNF})$  is their rate of grazing by heterotrophic nanoflagellates (HNF). The rate of uptake of DOM,  $U_{\text{DOM}}$ , is based on the carbon fraction, DOC:

$$U_{\text{DOC}} = \mu_B \cdot \Theta_T \cdot \text{DOC}^{\text{nR}} \cdot \text{B} \quad (\text{A.27})$$

in which  $\mu_B$  is the maximum rate at 0 °C,  $\Theta_T$  is the factor for temperature dependence and B is the concentration of bacteria. Thus, the rate coefficients,  $\mu_B$ , have units of

[ $\text{day}^{-1} (\text{mol N of bacterial biomass})^{-1}$ ]. Potential uptake of ammonium is formulated with a Michaelis–Menten type rate:

$$U_{\text{NH}_4}^{\star} = \frac{\mu_B}{\text{RCN}_B} \cdot \Theta_T \cdot \frac{\text{NH}_4}{(\text{K}_{\text{NH}_4, \text{B}} + \text{NH}_4)} \cdot \text{B} \quad (\text{A.28})$$

If the total uptake of nitrogen (from DOM and ammonium uptake) is sufficient ( $(\omega_B \cdot U_{\text{DOC}}) / \text{RCN}_B < (U_{\text{DON}} + U_{\text{NH}_3}^{\star})$ ), bacteria grow with fixed (carbon-based) growth efficiency,  $\omega_B$ , so that  $\text{GHP}_B = (\omega_B \cdot U_{\text{DOC}}) / \text{RCN}_B$ . Regeneration rates for DOM depend on bacteria:  $R_{\text{DOC}} = U_{\text{DOC}} - \text{GHP}_B \cdot \text{RCN}_B$  and  $R_{\text{DON}} = U_{\text{DON}} - \text{GHP}_B$ .

If  $R_{\text{DON}} < 0$ , the negative regeneration is the uptake of  $\text{NH}_4$  ( $U_{\text{NH}_4} = -R_{\text{DON}}$ ), provided the uptake is less than the potential uptake ( $-R < U^{\star}$ ).

If the total uptake of N is not sufficient for bacteria to grow with growth efficiency  $\omega_B$ ,  $-R_{\text{DON}} > U_{\text{NH}_4}^{\star}$  and nitrogen limits the growth of bacteria. In this case,  $\text{GHP}_B = U_{\text{DON}} + U_{\text{NH}_4}^{\star}$ . Regeneration is then calculated as above, so that there is no net regeneration of the nitrogen by bacteria. For mortality,  $M_B$ , we use the same first-order rate expression as Anderson and Williams (1998) (with temperature-dependence added):  $M_B = m_B \cdot \Theta_T \cdot \text{B}$ , in which  $m_B$  is the maximum mortality rate coefficient and  $G(\text{B}, \text{Z})$  is grazing of bacteria by Z (formulated as the other grazing rates). All bacterial mortality is assumed to become DOM.

The equations for DOM are:

$$\frac{\partial \text{DOC}^{\text{nR}}}{\partial t} = \sum_P [\text{RCN}_P \cdot E_P] + \sum_h \text{Slf}_h^{\text{C}} + D_{\text{DetC}} + \text{RCN}_B \cdot M_B - U_{\text{DOC}} \quad (\text{A.29})$$

$$\frac{\partial \text{DON}^{\text{nR}}}{\partial t} = \sum_P E_P + \sum_h \text{Slf}_h^{\text{N}} + D_{\text{DetN}} + M_B - U_{\text{DON}} \quad (\text{A.30})$$

Uptake of DON,  $U_{\text{DON}}$ , is proportional to the (variable) stoichiometry of non-refractory DOM:  $U_{\text{DON}} = (\text{DON}^{\text{nR}} / \text{DOC}^{\text{nR}}) \cdot U_{\text{DOC}}$ . Simulated bulk  $\text{DOX} = \text{DOX}^{\text{nR}} + \text{RDOX}$ , for  $X = \text{C}$  or  $\text{N}$ .

## References

- Anderson, T.R., Williams, P.J., 1998. Modeling the seasonal cycle of dissolved organic carbon at station e1 in the English Channel. *Estuar. Coast. Shelf Sci.* 46, 93–109.
- 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.

- Christian, J.R., 2005. Biogeochemical cycling in the oligotrophic ocean: Redfield and non-Redfield models. *Limnol. Oceanogr.* 50, 646–657.
- Dowd, M., Meyer, R., 2003. A Bayesian approach to the ecosystem inverse problem. *Ecol. Model.* 168, 39–55.
- Eslinger, D.L., Kashiwai, M., Kishi, M.J., Megrey, B.A., Ware, D.M., Werner, F.E., 2000. Model task team workshop report-final report of the international workshop to develop a prototype lower trophic level model for comparison of different marine ecosystems in the north pacific. Tech. Rep., vol. 15. PICES (North Pacific Marine Science Organization).
- Flynn, K.J., 2003. Modeling multi-nutrient interactions in phytoplankton: balancing simplicity and realism. *Prog. Oceanogr.* 56, 249–279.
- Fujii, M., Nojiri, Y., Yamanaka, Y., Kishi, M.J., 2002a. A one-dimensional ecosystem model applied to time-series station knot. *Deep-Sea Res., II* 49, 5441–5461.
- Fujii, M., Murashige, S., Ohnishi, Y., Yuzawa, A., Miyasaka, H., Suzuki, Y., Komiyama, H., 2002b. Decomposition of phytoplankton in seawater: Part 1. Kinetic analysis of the effect of organic matter concentration. *J. Oceanogr.* 58, 433–438.
- 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.
- Goldman, J.C., Caron, D.A., 1985. Experimental studies on an omnivorous microflagellate: implications for grazing and nutrient regeneration in the marine microbial food chain. *Deep-Sea Res.* 32, 899–915.
- Goldman, J.C., Caron, D.A., Dennett, M.R., 1987. Regulation of gross growth efficiency and ammonium regeneration in bacteria by substrate C:N ratio. *Limnol. Oceanogr.* 32, 1239–1252.
- 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.
- Lefevre, N., Taylor, A.H., Gilbert, F.J., Geider, R.J., 2003. Modeling carbon to nitrogen and carbon to chlorophyll a ratios in the ocean at low latitudes: evaluation of the role of physiological plasticity. *Limnol. Oceanogr.* 48, 1796–1807.
- Mongin, M., Nelson, D.M., Pondaven, P., Brzezinski, M.A., Tregeur, P., 2003. Simulation of upper ocean biogeochemistry with a flexible-composition phytoplankton model: C, N and Si cycling in the western Sargasso Sea. *Deep-Sea Res., I* 50, 1445–1480.
- Niraula, M.P., Casareto, B.E., Hanai, T., Smith, S.L., Suzuki, Y., 2005. Development of carbon biomass using incubations of unaltered deep-sea water. *J. Eco-Eng.* 17, 121–131.
- Smith, S.L., Yamanaka, Y., Kishi, M.J., 2005. Attempting consistent simulations of Stn. ALOHA with a multi-element ecosystem model. *J. Oceanogr.* 61, 1–23.
- Spitz, Y.H., Moisan, J.R., Abbott, M.R., Richman, J.G., 1998. Data assimilation and a pelagic ecosystem model: parameterization using time series observations. *J. Mar. Syst.* 16, 51–68.
- Spitz, Y.H., Moisan, J.R., Abbott, M.R., 2001. Configuring an ecosystem model using data from the Bermuda Atlantic Time Series (BATS). *Deep-Sea Res., II* 48, 1733–1768.
- Vallino, J.J., 2000. Improving marine ecosystem models: use of data assimilation and mesocosm experiments. *J. Mar. Res.* 58, 117–164.
- Verity, P.G., 1991. Measurement and simulation of prey uptake by marine planktonic ciliates fed plastidic and aplastidic nanoplankton. *Limnol. Oceanogr.* 36, 729–750.
- Williams, P.J.J., 1995. Evidence for the seasonal accumulation of carbon-rich dissolved organic material, its scale in comparison with changes in particulate material and the consequential effect on net C/N assimilation ratios. *Mar. Chem.* 51, 17–29.