Exploring Mechanisms of P Content Heterogeneity in Cultured Phytoplankton using Agent Based Modeling

A Thesis Presented

By

Neil Dakin Fredrick

to

The Department of Civil and Environmental Engineering

in partial fulfillment of the requirements
for the degree of

Master of Science

in

Civil and Environmental Engineering

in the field of

Environmental Engineering

Northeastern University
Boston, Massachusetts

August, 2012
Abstract
There can be significant intraspecific individual-level heterogeneity in the internal P content of phytoplankton, which can affect population-level behaviors, including the growth rate. There are several mechanisms that can create this heterogeneity, including biological variability in states and behavior. Here we use modeling to explore the contribution of various mechanisms to the heterogeneity in phytoplankton grown in a laboratory culture. An agent-based model simulates individual cells and their nutrient content. Heterogeneity is introduced by randomizing states (e.g. P content) and parameters (e.g. maximum uptake rate) of daughter cells at division. The model was calibrated to observations of internal P content of individual cells of *Thalassiosira pseudonana* measured using synchrotron X-ray fluorescence (SXRF). Then, simulations were performed with individual mechanisms of heterogeneity turned off. Comparison of the coefficient of variation (CV) of this and the basecase simulation (all mechanisms turned on) provide an estimate of the relative contribution of the mechanism. The results show that the mechanism with the largest contribution is variability in maximum internal cell quota, which when removed results in a CV of 0.21 compared to a CV of 0.38 with all mechanisms.
1. Introduction

Phytoplankton play an important role in the ecology and biogeochemistry of freshwater and marine systems. Thus, understanding their composition and behavior is important. The elemental composition of phytoplankton was originally considered to be relatively uniform (Redfield, 1934). However, subsequent research has shown that it can vary between species (Sarthou et al., 2005) or clones (Geider and Roche, 2002) and within a species under different conditions (Droop, 1975; Leonardos and Geider, 2004). The composition can also vary within a species at the same conditions. Recent observations of internal phosphorus content in individual cells have shown significant intraspecific heterogeneity in laboratory and field populations, with coefficients of variations (CV) ranging from 0.08 to 1.11 (Table 1).

This may have important consequences, because heterogeneity in nutrient content results in heterogeneity in behavior, and that can have significant effects on community ecology (Bolnick et al., 2011). Specifically, for phytoplankton the growth rate is generally considered to be a function of the internal nutrient content (Droop, 1975). Heterogeneity in P content can affect the population growth rate, because the equation is nonlinear, a case of Jensen’s inequality. If the population growth rate is determined based the population average cell quota (i.e. average the quota then determine one growth rate for the population), then it will be higher than if the growth rate is determined for each individual and then averaged to get the population growth rate.

There are several mechanisms that can create heterogeneity in the internal nutrient content of a population of phytoplankton (summarized in Table 2). Asynchronicity in the cell cycle, biological variability and feedbacks between mechanisms can create heterogeneity in the laboratory cultures explored here. Other mechanisms do not apply to the conditions investigated here. Microscale patchiness is not important because there are no zooplankton. Marcoscale mixing is assumed not to apply to the mixed reactor conditions in the lab. Genetic differences applies to populations with multiple strains, which is not the case in a laboratory mono strain culture.

Theory suggests that heterogeneity of P content is important and there are multiple mechanisms that can cause it. However, observation of intraspecific heterogeneity are limited, as are the validated models (Bucci et al., 2012b) that predict it. What needs to be determined is how much heterogeneity there is what mechanisms cause it, and what affect it has on population-level growth rates and behaviors.

Agent-based modeling (ABM), also known as individual-based modeling (IBM), can be used to simulate population heterogeneity and investigate the contribution of different mechanisms to it. A previous study used this approach to explore the heterogeneity in a field population (Bucci et
al., 2012b). Here we use the same general approach to investigate the heterogeneity observed in laboratory culture. We develop a model, which includes a number of mechanisms of heterogeneity and apply it to the data of Nunez-Milland et al. (2010), who grew the marine diatom *Thalassiosira pseudonana* in batch culture and measured P content using synchrotron X-ray fluorescence (SXRF). We calibrated the model to the observations. Then, simulations are performed with different mechanisms turned off to estimate their contribution to the heterogeneity.

2. Model Description

2.1 Overview

The ABM simulates one type of phytoplankton (*Thalassiosira pseudonana*) in culture. The reactor is assumed to be completely mixed (uniform concentration of extracellular phosphorus), so the agents do not see any spatial variation in the environment. Cell growth is a function of photosynthesis and respiration. Photosynthesis depends on the P content and light, with no effect from temperature and other nutrients. Phytoplankton take up and excrete P. The model includes mortality and dead cells are tracked by the model. They excrete nutrient but do not perform uptake, photosynthesis, respiration and growth. Details of the processes and modeling approach are provided below. The model code and a description following the standard ODD protocol for documenting ABMs (Grimm et al., 2010) are provided in the SI.

2.2 Photosynthesis, Respiration, Growth

In between divisions, a cell’s size changes due to photosynthesis and respiration (Bucci et al., 2012b):

\[
\frac{dm}{dt} = (\mu_P - \mu_R)m = \mu_G m
\]

where \(m\) is the cell mass (mol C/cell), \(\mu_P\) (1/d) is the photosynthesis rate, \(\mu_R\) (1/d) is the respiration rate, and \(\mu_G\) (1/d) is the growth rate. The photosynthesis rate is a function of light and the phosphorus quota (Droop, 1975; Hellweger and Kianiard, 2007a; Hellweger et al., 2008)

\[
\mu_P = \mu_{P,MAX} L_I \left(1 - \frac{q_0}{q}\right) \quad \text{for } q > q_0
\]

\[
\mu_P = 0 \quad \text{for } q \leq q_0
\]

where \(\mu_{P,MAX}\) (1/d) is the maximum photosynthesis rate, \(L_I\) is the light limitation term (light(1) or dark(0)), \(q_0\) (mol P/mol C) is the subsistence quota and \(q\) (mol P/mol C) is the cell quota. If the quota drops below the subsistence quota then the cell does not perform photosynthesis.
2.3 Cell Division

Diatoms have a silicate outer cell wall called a frustule, which provides protection from predators (Hamm, 2005). In many diatoms, the frustule makes the cell division process complex, because the daughter cell’s valve forms inside the mother cell and therefore must be a smaller diameter (Jewson, 1992). This gradually reduces the average diameter of the population over time (Round et al., 1990). The diameter is increased periodically through the formation of an auxospore (a phase of sexual reproduction) that does not have a rigid silicate shell (Mann, 1993; Round et al., 1990). This size reduction mechanism has been modeled before for *Cyclotella meneghinana* (Bucci et al., 2012b). However, size reduction has not been observed for *T. pseudonana*. Paasche (1973) grew *T. pseudonana* in culture and there were no significant changes in the mean cell diameter. Hildebrand et al. (2007) also did not observe size reduction in this species and provide evidence that this is because the girdle-band portion of the cell is expandable (allows the area where the valves are created to be wider than the ends of the cells). Therefore, size reduction is not included in the model. Cell division is based on the size state variable \( m \) and a minimum cell size parameter \( m_0 \). When a cell’s size is double the minimum, it divides (Hellweger and Kianirad, 2007a; Hellweger et al., 2008).

2.4 Biological Variability

The model includes biological variability by randomizing the states (e.g. cell size) or parameters (e.g. max uptake rate) of daughter cells at division. Biological variability in states can result from unequal division (Mihalcescu et al., 2004). In the model, the mass and P are divided between the daughter cells using the split fractions \( S_f-m \) and \( S_f-qm \) that are randomly varied from the ideal value (0.5) using the function proposed by Kreft (Bucci et al., 2012b; Kreft et al., 1998). A value is drawn from a modified normal distribution, truncated to lower and upper bounds. The function limits the value to within two standard deviations of the mean. Additional upper and lower limits can be specified to avoid unrealistic values (e.g. a split fraction that is less than zero or greater than one). There can also be significant variability in behavior among cells in a population (Musat et al., 2008). The model simulates this by randomizing the value of parameters at division. Here, 10 parameters are randomized (Table 4). These values are randomized from a global mean value using the function of Kreft et al. (1998).

2.5 Mortality

Mortality of phytoplankton has been observed to be a function of environmental stress (Brussaard et al., 1997). This was shown in *T. pseudonana* by Franklin et al. (2012), where up to 25 percent of the cells in a population died after 3 weeks of nutrient limitation. Death is simulated stochastically based on a mortality rate, which increases as the cells internal phosphorus quota decreases (environmental stress):
\[ \mu_D = \mu_{D,\text{BASE}} \left( \frac{q_D}{q} \right) \]  

(3)

where \( \mu_D \) (1/d) is the death rate, \( \mu_{D,\text{BASE}} \) (1/d) is the base death rate, and \( q_D \) (mol P/mol C) is the death quota (\( \mu_D = \mu_{D,\text{BASE}} \) when \( q = q_D \)). The model simulates mortality as a discrete event where each cell has a probability of dying at each time step (\( \mu_D \Delta t \)). A random number (between 0 and 1) is drawn and compared to the probability. If the random number is smaller than the probability the cell dies.

2.6 P quota, uptake, excretion

An individual cell’s quota changes due to uptake, excretion and growth dilution (Bucci et al., 2012b):

\[ \frac{dq}{dt} = V - W - \mu_G q \]  

(4)

where \( V \) (mol P/mol C/d) is the uptake rate, \( W \) (mol P/mol C/d) is the excretion rate, and \( \mu_G q \) accounts for growth dilution. The uptake formulation, is based on Michaelis-Menten kinetics, applied on a surface-area basis with an additional term from (Thingstad, 1987) to limit uptake at high quotas:

\[ V = V_{A,\text{MAX}} \frac{A}{m} \frac{PO_4}{K_M + PO_4 \frac{q_{\text{MAX}}}{{\text{MAX}}} - q} \]  

(6)

where \( V_{A,\text{MAX}} \) (mol P/\( \mu m^2 \)/d) is the maximum specific uptake rate per area, \( A \) (\( \mu m^2 \)) is the surface area, \( K_M \) (mol P/m^3) is the half-saturation constant, \( PO_4 \) (mol P/m^3) is the extracellular concentration of phosphate, and \( q_{\text{MAX}} \) (mol P/mol C) is the maximum quota. The surface area is calculated from the cell size (\( m \)) assuming a cylindrical shape (Reynolds, 2006), and specified cell diameter (\( d \)) and density (\( \rho_c \)). The excretion rate is a first order function of the internal P content:

\[ W = k_w q \]  

(7)

where \( k_w \) (1/d) is the excretion constant. Excretion is in the form of phosphate, the same form as extracellular nutrient.

2.7 Extracellular Phosphate

The concentration of extracellular phosphate in the reactor changes as a function of inflow/outflow, and uptake and excretion by the phytoplankton cells (Bucci et al., 2012b):
\[
\frac{dP_{O4}}{dt} = \frac{Q}{Vol} P_{O4,IN} - \frac{Q}{Vol} P_{O4} + \frac{1}{Vol} \sum_{cells} [(W - V)mS_R]
\]  

(8)

where \(Vol\) (m\(^3\)) is the volume of the reactor, and \(P_{O4,IN}\) (mol/ m\(^3\)) is the concentration of phosphate in the inflow. \(S_R\) is the super-individual number (see section 2.9)

### 2.8 Outflow and Dilution

The model includes inflow/outflow (for chemostat operation) and dilution, which add/remove extracellular phosphate and remove cells. The mass balance for external P is shown in equation 8. The removal of cells can be simulated in two ways (Hellweger and Bucci, 2009). One approach is stochastic, where each agent has a probability of being removed by outflow or removed in the dilution during every time-step. The probability is equal to the dilution rate \((Q/Vol)\) multiplied by the time-step \((\Delta t)\) for wash out or the dilution ratio for dilution. The alternative approach is to decrease the \(S_R\) (see section 2.9) of the agents (e.g. \(\Delta S_R = Q/vol S_R\Delta t\)), but keep all the agents in the model. Here outflow and dilution is modeled using the reduction of \(S_R\) approach.

### 2.9 Agent Accounting

The very large number of cells in the culture (up to about 2.5 billion in the application presented here) makes it impractical to explicitly simulate each one. For this case, using a 3.8 Ghz processor, it would take 32 years to perform one simulation with that many agents. It would also require 100 terabytes of RAM. The model therefore uses super individuals, where one agent in the model represents many individual cells in the real culture (Bucci et al., 2012b; Woods, 2005). Each agent has a \(S_R\), which specifies the number of real cells it represents. The model has two groups of super individuals: living and dead cells. Each group has a range for its number of super individuals. To keep the number of agents within these specified ranges, the model splits or combines cells. If the number of agents goes below the minimum, the model splits the cell with the largest \(S_R\) into two cells, each with a \(S_R\) that is half of the original cell. The two cells retain all the other state variables of the original cell. If the number of agents exceeds the maximum then the model combines the two cells with the lowest \(S_R\). The combined cell’s \(S_R\) is the sum of the individual cells and, its mass and P mass is determined by a weighted average of the two original cells.

### 2.10 Implementation

The model is implemented in the Netlogo ABM framework version 5.0 (Wilensky, 1999). A constant time step is used. For each time step, first every agent calculates its uptake, excretion
and updates its internal P; this is done without changing the concentration of extra-cellular P. Second after all of the agents have finished updating their P content, the model calculates the new concentration of extracellular P. Then the agents will calculate their growth, divide if necessary and run the mortality routine. If any of the earlier steps causes the number of agents to be outside of the acceptable ranges then the split/combine routine is run. The time step used is 0.0005 day (43 second).

2.11 PLM Model

The model includes an equivalent population level model (PLM) for comparison (Bucci et al., 2012a).

3. Model Application

3.1 Experimental Data

Observations of internal P content were obtained by Nunez-Milland et al. (2010) using SXRF. Three cultures were run for different methods of measuring the P content of cells. Here the data from the cultures corresponding to the nickel and gold grids are used. Thalassiosira pseudonana (CCMP 1335) was grown in batch culture in a 1 liter vessel. Cells were grown in media with 1.5 μM phosphate, which are nutrient replete conditions. The cultures were grown with a light:dark cycle of 14:10 hours respectively, and were stirred manually daily. Cells were sampled during the exponential growth phase and analyzed using SXRF, which allows for the P content of individual cells to be measured.

3.2 Simulations Performed

A number of simulations were performed. The first one, the base case (M1), is calibrated to the observations. This simulation has all of the mechanisms of heterogeneity turned on, and the variability for each state and parameter was optimized by minimizing the error between model and data (see Section 3.5).

Then, several simulations were performed to investigate the contribution of various mechanisms of heterogeneity to the observed heterogeneity (Table 3). Simulations X1-X12 are the same as the basecase simulation but with one mechanism of heterogeneity switched off. The contribution of each mechanism is determined by comparing the CV of the X run to that of the M1 run. This approach is similar in spirit, but different in two details, from that used by Bucci et al. (2012b). First, they grouped parameters (i.e. randomization for several parameters was turned on/off
together in one simulation) while here individual parameters are considered. Second, the runs in Bucci et al. (2012b) generally had only one mechanism turned on. Here the contribution of mechanisms is investigated using simulations where individual mechanisms are turned off. The use of runs that remove a single mechanism allows for determining the contribution of a single mechanism, while keeping the overall heterogeneity that is created by having an asynchronized population of cells.

3.3 Initial and Boundary Conditions

The model starts with a population of living cells (50,000 cells/ml) with uniform states (i.e. same $m$, $q$). If a mechanism of heterogeneity is turned on then the cells start with that parameter randomized. The concentration of extracellular P in the media (at the start and for each dilution) is 1.5 $\mu$M. The model is allowed to spin up for 4 days, then there is a 1 to 100 dilution (i.e. 99 percent cells are removed and medium is replaced with new medium). This point corresponds to the beginning of the experiment, and the cell density matches the one measured in the lab. The simulation is then run for 3.5 days, stopping during the light cycle. This is consistent with the experimental procedure where the samples were taken during the exponential growth phase. The run time is sufficient to remove the effects of the initial conditions and to reach a stable P content distribution (about 8 generations). The ranges of super individuals for living and dead cells are 1000-2000 and 100-200 respectively.

3.4 Parameter Values: Means and Variabilities

The means for the phytoplankton parameters were calibrated manually to match the SXRF data, and the resulting values are within the range of the literature (Table 4). Then, an optimization algorithm was used to find the best variability for each of the 10 parameter and two split fractions (quantified as CV, Table 4). The objective of the algorithm is to minimize the model error, which is quantified using the RMSE between observed and modeled individual cell quotas. The algorithm proceeds in a step-wise manner. At each step, the model is run 10 times (to average out variability between runs) and the RMSE is averaged. At the start a base-line RMSE is determined from the starting conditions. Then the algorithm proceeds down the list (randomized) of parameters and increases or decreases (random) the CV by a small amount. The resulting RMSE is compared to the previous lowest RMSE. If the new RMSE is lower the change is kept and the base-line RMSE is updated. The algorithm continues until there are no further improvements in RMSE. The parameters for the PLM are set to the mean of the corresponding parameters in the ABM.

4. Results and Discussion

4.1 Optimization Results
The optimization algorithm was run for 9,600 model runs (80 passes through 12 parameters and split fractions with 10 runs each), 9 times (3 different starting points, 3 different random number seeds). The results from the algorithm show that 8 of the runs find a similar final RMSE, while one run gets stuck at a significantly higher RMSE (Figure 1). This can happen because the model is stochastic and the RMSE will vary among repeat runs. The optimization uses a simple hill-climbing method so when the optimization algorithm is in a plateau the stochastic results can cause the algorithm can get stuck in a phantom local minimum. More sophisticated routines are available (Xi et al., 2004), which can avoid this problem. However, for the purpose of this investigation the current routine’s results are acceptable. The CV values for the optimization run with the lowest error were adopted for the basecase simulation (Table 4). The optimization algorithm found variability for the parameters that was within the range of values used in other models (Bucci et al., 2012a; Bucci et al., 2012b; Hellweger and Kianirad, 2007b; Kreft et al., 1998) and observed in lab studies (Mihalcescu et al., 2004; Musat et al., 2008).

4.2 Model-data Comparison

The results from the basecase simulation (M1) are compared to the observations (SXRF) in Figure 2A, which shows individual cell quotas in femtomole P per cell. The model reproduces the observations relatively well, with an RMSE of 0.27. There is a group of cells with similar observed P contents between 4 and 4.5 fmol/cell and this feature is not reproduced by the model. The time from culture inoculation and SXRF observation was relatively short (3 days) and these cells may represent a synchronized sub-population. The simulation had a CV of P content of 0.38, which is somewhat higher than the data’s CV of 0.36.

The model predicts that dead cells account for about 3 percent of the population, and they are mostly at the lower end of the P distribution (Figure 2B). The fraction of dead cells is within observed ranges for diatoms in culture. Brussaard et al. (1997) found the percentage of dead cells ranges from 2 to 28 percent depending on the amount of nutrient limitation. The higher percentage of dead cells at the lower end of the P distribution is also consistent with observations (Brussaard et al., 1997; Franklin et al., 2012). In the model, this is due to a higher death rate for cells with lower internal P quota and excretion after cell death.

4.3 Sources of Heterogeneity

To investigate the contribution of various mechanisms of heterogeneity, a number of simulations with different mechanism switched off were performed. The relative contributions were determined by comparing the CV of each run with a mechanism removed and comparing it to the base case (Figure 4 and Table 3). These results show that the main mechanism for creating heterogeneity is variability in the maximum internal cell quota ($q_{\text{MAX}}$). The variation in the state of internal P ($SF-qm$) at cell division has a smaller contribution. The $q_{\text{MAX}}$ run has a CV of 0.19 and the $SF-qm$ run has a CV of 0.34. Heterogeneity created by the other mechanisms is within
the error (determined as the standard deviation over ten runs). This shows that variability in $q_{MAX}$ has the largest contribution to the overall heterogeneity of the population.

### 4.4 Effect of Heterogeneity

Most of the current methods to model phytoplankton, with ecological (Flynn, 2003; Geider and Roche, 2002) or water quality (Cerco et al., 2004; Hellweger and Lall, 2004) objectives, or in textbooks (Chapra, 1997; Schnoor, 1996), use a population-level modeling (PLM) approach. The PLM approach divides large areas into smaller sections, in which the conditions are uniform. The assumption of uniform conditions in PLM models can cause a significant error (Hellweger and Kianirad, 2007a), but the use of ABM allows for inter-population variability, which the current PLM’s cannot consider. The PLM of the laboratory culture has a growth rate that was 14 percent higher (PLM average growth rate of 0.77 d$^{-1}$ and ABM average growth of 0.68 d$^{-1}$).

### 5. Summary and Outlook

This paper explored the mechanisms that underlie heterogeneity in internal P content in cultured phytoplankton. An ABM was developed, which includes a number of mechanisms of heterogeneity, including biological variability in states and behavior. The model was calibrated for the conditions in the laboratory culture with averages determined manually and the variation of parameters and states determined by an optimization algorithm. The laboratory data had a CV for internal P content of 0.36 and the base-case simulation (all mechanisms turned on) had a CV of 0.38. Then, Simulations were performed with individual mechanisms turned off to determine the contributions of individual mechanisms to the overall heterogeneity in the population. The simulations demonstrate that most of the mechanisms have little effect on heterogeneity, with their CV’s falling within one standard deviation of the basecase. However two mechanisms, variation in maximum internal cell quota ($q_{MAX}$) state of internal P ($SF_{qm}$), each contribute significantly to heterogeneity so that turning them off affects the populations CV. Variation in internal maximum cell quota ($q_{MAX}$) is responsible for most of the heterogeneity.

The contributions of the various mechanisms is a function of the experimental conditions (e.g. batch culture with nutrient replete conditions), and may not be applicable to all conditions. The contributions of various mechanisms may be different under nutrient limiting conditions or in natural systems where there are additional mechanisms contributing to the heterogeneity (e.g. microscale patchiness (Bucci et al., 2012b)).

The conclusion that the variability in the maximum internal cell quota ($q_{MAX}$) is responsible for most of the heterogeneity is biologically plausible. Phytoplankton, including diatoms, can store P in excess of their immediate growth requirements (Droop, 1975; Rhee, 1973). This luxury
uptake is mostly in the form of polyphosphates. For example, Perry (1976) demonstrated that *T. pseudonana* can accumulate between 19-43% of cell P in polyphosphates under N limitation. Since there is not an absolute requirement for the amount of P taken up by this mechanism, it makes sense that it can be highly variable (and thus *qmax* is highly variable). In contrast, for example, there is probably very little a cell could do to change minimum P quota. Much of the essential P resides in nucleic acids found in ribosomal RNA (Geider and Roche, 2002; Leonardos and Geider, 2004) and relatively little variation is possible in this pool. However, recent observations show that phytoplankton, including *T. pseudonana* can substitute non-phosphorus (sulfur and nitrogen containing) lipids for phospholipids (Martin et al., 2010). Recent work on sediment cycling of P has underlined the importance of such polyphosphate pools: a significant fraction of P that finds its way into calcium phosphate in sediments is derived from diatom polyphosphates (Diaz et al., 2008), such that variability in these pools could genuinely drive global elemental cycles.

Future research may build on these results. A more comprehensive picture of heterogeneity in cell composition may be obtained by exploring various fractions of P (e.g. polyphosphates see above) and also considering other elements (i.e. carbon, nitrogen). Another potential expansion would be to bring in data on heterogeneity in behavior. For example, Musat et al. (2008) measured inorganic carbon and nitrogen assimilation by bacteria. Levy et al. (2012) measured the protein expression, morphology, and growth rate of individual cells of yeast. An ideal data set would include measured growth rates of individual cells and corresponding internal nutrient contents.

**Supplementary Information**

The ODD for the model can be found in Appendix A and the model code can be found in Appendix B.

**Acknowledgments**  This work was supported by grants from NSF (CBET 0730061 and 0913071) to Benjamin S. Twining, NSF (CBET 0730239 and IOS 1121233) and NEU Provost Grant to Ferdi L. Hellweger, and NSF (IOS 1121513) and Burroughs Wellcome Fund Collaborative Research Travel Grant (1010175) to John A. Berges.
References


Olson, R.J., Vaulot, D., Chisholm, S.W., Chisholm, S.M., 1986. Effects of Environmental Stresses on the Cell Cycle of Two Marine Phytoplankton Species. Plant Physiology 80, 918-925.


pp. 177–192.


Table 1. Observed intraspecific heterogeneity in P content in phytoplankton.

<table>
<thead>
<tr>
<th>Species</th>
<th>Lab/Field</th>
<th>Method</th>
<th>CV</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prochlorococcus</td>
<td>Lab</td>
<td>TEM-X-ray (a)</td>
<td>0.17-0.27(b)</td>
<td>(Heldal et al., 2003)</td>
</tr>
<tr>
<td>Synechococcus</td>
<td>Lab</td>
<td>TEM-X-ray (a)</td>
<td>0.20-0.86(b)</td>
<td>(Heldal et al., 2003)</td>
</tr>
<tr>
<td>Thalassiosira pseudonana</td>
<td>Lab</td>
<td>SXRF (c)</td>
<td>0.27-0.49(b)</td>
<td>(Nunez-Milland et al., 2010)</td>
</tr>
<tr>
<td>Microcystis aeruginosa</td>
<td>Field</td>
<td>XRMA (d)</td>
<td>0.18-0.47(b)</td>
<td>(Krivtsov et al., 2005)</td>
</tr>
<tr>
<td>Dinophysis norvegica</td>
<td>Field</td>
<td>NMP (f)</td>
<td>0.30(b)</td>
<td>(Gisselson et al., 2001)</td>
</tr>
<tr>
<td>Microcystis aeruginosa</td>
<td>Field</td>
<td>XRMA (d)</td>
<td>0.14-0.37(b)</td>
<td>(Sigee and Levado, 2000)</td>
</tr>
<tr>
<td>Ceratium hirendenella</td>
<td>Field</td>
<td>XRMA (d)</td>
<td>0.23-0.30(e)</td>
<td>(Sigee et al., 1999)</td>
</tr>
<tr>
<td>Asterionella formosa</td>
<td>Field</td>
<td>XRMA (d)</td>
<td>0.30-2.1(b)</td>
<td>(Krivtsov et al., 2000)</td>
</tr>
<tr>
<td>Anabaena spiroides</td>
<td>Field</td>
<td>XRMA (d)</td>
<td>0.08-0.13(c)(g)</td>
<td>(Clay et al., 1991)</td>
</tr>
<tr>
<td>Spirulina platensis</td>
<td>Field</td>
<td>XRMA (d)</td>
<td>0.11-0.27(b)</td>
<td>(El-Bestawy et al., 1996)</td>
</tr>
<tr>
<td>Cyclotella meneghiniana</td>
<td>Field</td>
<td>XRMA (d)</td>
<td>0.14-0.64(b)</td>
<td>(El-Bestawy et al., 1996)</td>
</tr>
<tr>
<td>Cyclotella meneghiniana</td>
<td>Field</td>
<td>SXRF(c)</td>
<td>1.11(b)</td>
<td>(Bucci et al., 2012b)</td>
</tr>
</tbody>
</table>

(a) Transmission electron microscope X-ray microanalysis
(b) Biomass-based (mol P/mol C)
(c) Synchrontron X-ray fluorescence
(d) X-ray microanalysis
(e) Cell based (mol P/ cell)
(f) Nuclear microprobe
(g) Vegetative cells
Table 2. Mechanisms of Internal Nutrient Heterogeneity in Phytoplankton

<table>
<thead>
<tr>
<th>Mechanism</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Microscale Patchiness</td>
<td>The concentration of extracellular nutrients can vary at the microscale. Patches with higher concentrations are created by zooplankton excretion or other processes, and cells randomly encounter these patches which leads to heterogeneity (Bucci et al., 2012b; Lehman and Scavia, 1984).</td>
</tr>
<tr>
<td>Macroscale Mixing</td>
<td>Macroscale advection and dispersion causes populations at one location made up of cells with different life histories (i.e. exposure to external nutrient concentration) and therefore different nutrient contents (Hellweger and Kianirad, 2007a).</td>
</tr>
<tr>
<td>Genetic Differences</td>
<td>In nature, phytoplankton populations are composed of many different species and strains or clones that have different behaviors (Franklin et al., 2012), which leads to heterogeneity.</td>
</tr>
<tr>
<td>Cell Cycle Asynchronicity</td>
<td>Phytoplankton populations are generally not synchronized, but consist of cells at different stages in the cell cycle and size. This translates to heterogeneity in the absolute nutrient content (mol P/cell). The biomass-based nutrient content (mol P/mol C) is not directly affected by this, but it can be indirectly effected if P and C uptake are asynchronous. The cell cycle can also affect nutrient uptake (Brzezinski, 1992).</td>
</tr>
<tr>
<td>Biological variability</td>
<td>Cell division creates two daughter cells that are never exactly the same in states (i.e. size) (Mihalcescu et al., 2004) or behavior (i.e. nutrient uptake) (Musat et al., 2008). Differences in states creates heterogeneity immediately. The differences in behavior can create heterogeneity of the internal nutrient quota in a population subsequently.</td>
</tr>
<tr>
<td>Feedbacks</td>
<td>Once there is heterogeneity a number of feedback mechanisms can modify it. Internal nutrient quotas can affect nutrient uptake, excretion, and mortality (Rhee, 1973; Tilman and Kilham, 1976) (Brussaard et al., 1997; Parslow et al., 1984). Nutrient status also affects the growth rate and cell cycle (Droop, 1975; Olson et al., 1986), which can lead to feedback loops.</td>
</tr>
</tbody>
</table>
Table 3. Summary of simulations performed

<table>
<thead>
<tr>
<th>Run</th>
<th>Description</th>
<th>Heterogeneity (CV)(a)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data</td>
<td>0.36</td>
<td></td>
</tr>
<tr>
<td>M1</td>
<td>Base case</td>
<td>0.38 (0.0066)</td>
</tr>
</tbody>
</table>

*No variability in...*  

<table>
<thead>
<tr>
<th>Run</th>
<th>Description</th>
<th>Heterogeneity (CV)(a)</th>
</tr>
</thead>
<tbody>
<tr>
<td>X1</td>
<td>$\mu_{P,MAX}$</td>
<td>0.38 (0.0070)</td>
</tr>
<tr>
<td>X2</td>
<td>$q_0$</td>
<td>0.37 (0.0035)</td>
</tr>
<tr>
<td>X3</td>
<td>$\mu_R$</td>
<td>0.37 (0.0029)</td>
</tr>
<tr>
<td>X4</td>
<td>$m_0$</td>
<td>0.36 (0.0032)</td>
</tr>
<tr>
<td>X5</td>
<td>$\mu_{D,BASE}$</td>
<td>0.37 (0.0053)</td>
</tr>
<tr>
<td>X6</td>
<td>$q_D$</td>
<td>0.36 (0.0049)</td>
</tr>
<tr>
<td>X7</td>
<td>$V_{Max}$</td>
<td>0.36 (0.0057)</td>
</tr>
<tr>
<td>X8</td>
<td>$K_M$</td>
<td>0.37 (0.0060)</td>
</tr>
<tr>
<td>X9</td>
<td>$q_{MAX}$</td>
<td>0.21 (0.0018)</td>
</tr>
<tr>
<td>X10</td>
<td>$k_w$</td>
<td>0.36 (0.0048)</td>
</tr>
<tr>
<td>X11</td>
<td>$Sf=qm$</td>
<td>0.34 (0.0063)</td>
</tr>
<tr>
<td>X12</td>
<td>$Sf-m$</td>
<td>0.36 (0.0051)</td>
</tr>
</tbody>
</table>

(a) Mean (SD)
Table 4. Model parameters for phytoplankton (*Thalassiosira pseudonana*)

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Units</th>
<th>Mean (CV)</th>
<th>Literature (1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu_{P,MAX}$</td>
<td>1/day</td>
<td>1.6 (0.40)</td>
<td>2.7-2.9 (a)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2.0-2.4 (b)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2.4 (c)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2.4-2.8 (d)</td>
</tr>
<tr>
<td>$q_0$</td>
<td>mol P/mol C</td>
<td>1.52e-3 (0.50)</td>
<td>1.5e-3 (e)(j)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1.8e-3 (f)(j)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1.9e-3 (g)(j)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4.5e-3 (h)(j)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1.8e-4 - 2.7e-4(i)(j)</td>
</tr>
<tr>
<td>$\mu_R$</td>
<td>1/day</td>
<td>0.15 (0.50)</td>
<td>0.1-0.2 (k)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.15 (0.05-0.25)(l)</td>
</tr>
<tr>
<td>$m_0$</td>
<td>mol C /cell</td>
<td>1.8e-13 (0.05)</td>
<td>1.8e-13 - 3.2e-13(n)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3.6e-13 (o)</td>
</tr>
<tr>
<td>$\mu_{D,BASE}$</td>
<td>1/day</td>
<td>0.2 (0.40)</td>
<td>0.1 (0.05-0.25)(p)</td>
</tr>
<tr>
<td>$q_D$</td>
<td>mol P/mol C</td>
<td>1.6e-2 (0.45)</td>
<td></td>
</tr>
<tr>
<td>$V_{A,MAX}$</td>
<td>mol P/µm²·day</td>
<td>2.64e-15 (0.40)</td>
<td>1.31e-15 - 2.06e-15(q)(t)(u)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2.0e-15 (r)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>9.15e-15 (s)(t)(u)</td>
</tr>
<tr>
<td>$K_M$</td>
<td>mol P /m³</td>
<td>6e-4 (0.45)</td>
<td>5.8e-3 (v)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>5.0e-4 - 7.0e-4(w)</td>
</tr>
<tr>
<td>$q_{MAX}$</td>
<td>mol P/mol C</td>
<td>1.5e-2 (0.55)</td>
<td>$q_{MAX}/q_0$=5-13(x)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>6.8-20(y)</td>
</tr>
<tr>
<td>$k_w$</td>
<td>1/day</td>
<td>0.5 (0.60)</td>
<td>0.67-3.3(x)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0-2.9(z)</td>
</tr>
<tr>
<td>$d$</td>
<td>µm</td>
<td>4.13 (0.00)</td>
<td>4.0-5.0(A)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3.8-6.0(B)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2.5-14(C)</td>
</tr>
<tr>
<td>$\rho_c$</td>
<td>mol C/µm²</td>
<td>2.95(0.00)</td>
<td>2.95(D)</td>
</tr>
<tr>
<td>$S_f$-qm</td>
<td></td>
<td>0.5(0.75)</td>
<td></td>
</tr>
<tr>
<td>$S_f$-m</td>
<td></td>
<td>0.5(0.10)</td>
<td></td>
</tr>
</tbody>
</table>

(1) Normal font indicates any phytoplankton. **Bold** font indicates *T pseudonana*. **Italics** represents the clone used in the culture (3H)

(a) Olsen and Paasche (1986), 20°C, Salinity 24, $\mu_{MAX}$ (1/day) (*T pseud* (W))
(b) Guillard and Ryther (1962), salinity 0.5-32, $\mu_{MAX}$ (1/day) (*T pseud* various clones)
(c) Guillard et al. (1973), salinity 30, $\mu_{MAX}$ (1/day) (*T pseud* (3H))
(d) Paasche (1973), salinity 24, $\mu_{MAX}$ (1/day) (*T pseud* (3H))
(e) Perry (1976), 7.5e-16 mol p / cell at growth rate of 0.408/d (*T pseud* (66-A))
(f) Fuhs (1969), 0.90 fmol P/cell (*T pseud* (3H))
(g) Fuhs et al. (1971), 0.95 fmol P/cell (*T pseud* (3H))
(h) Parslow et al. (1984), 2.2 fmol P/cell(*T pseud* (3H))
(i) Wynne and Rhee (1986), 0.9 -1.34 e-16 mol P/cell(*T pseud* (3H))
(j) Converted using average cell mass 2.49 e-13 mol C/cell
(l) Thomann and Mueller (1987), Typical value (range)
(m) Hellweger and Kianirad (2007a)Converted using $m_{AVE}$=2ln(2)$m_0$
(n) Perry (1976), 0.25 and 0.45 pmol C/cell, at growth rates of 0.984/d and 0.408/d respectively \((T\,\text{pseud}\,(66-A))\)
(o) Reynolds (2006), 5.9 pg C/cell \((T\,\text{pseud})\)
(p) Vasalou et al. (2009), Typical value (range), p. 211, \(\mu_{D,\text{MAX}}\)
(q) Perry (1976), 0.443-0.699E-13 mol P/cell-day \((T\,\text{pseud})\)
(r) Fuhs et al. (1971), 2.0e-15 mol P/\(\mu\)m²d \((T\,\text{pseud}\,(3H))\)
(s) Parslow et al. (1984), 3.10E-15 mol P/cell-day \((T\,\text{pseud}\,(3H))\)
(t) \(V_{\text{A,MAX}}=V_{\text{Max}}/A_{\text{AVE}}\)
(u) \(A_{\text{AVE}}=33.8797\ \mu\text{m}^2/\text{cell}\)
(v) Fuhs et al. (1971), 0.58e-6 g-at. L⁻¹p \((T\,\text{pseud}\,(3H))\)
(w) Perry (1976), Phosphorous and Nitrogen Limited Cultures 0.5-0.7 \(\mu\)M \((T\,\text{pseud}\,(3H\ 5.8-7.0)\ (13-1\ 7.0)\ (66-A\ 6.0-7.0))\)
(x) Fisher and Lean (1992)
(y) Bidle and Falkowski (2004)
(z) Olsen (1989)
(A) Belcher and Swale (1977), 4-5 \(\mu\)m \((T\,\text{pseud})\)
(B) Lowe and Busch (1975), 3.8-6.0 \(\mu\)m \((T\,\text{pseud})\)
(C) Guillard and Ryther (1962) 2.5-14 \(\mu\)m \((T\,\text{pseud})\)
(D) Reynolds (2006) cell volume 20 \(\mu\)m³ and cell Carbon (pg/cell) \((T\,\text{pseud})\)
Figure 1. Results of 9 optimization runs. The lines show the error of the best parameter set found by the algorithm, and the points are the error of each run performed.
Figure 2.  A. Observed and simulated (M1) distribution of cell quota of *Thalassiosira pseudonana* in laboratory culture. Data are from (Nunez-Milland et al., 2010). B. Percentage of dead cells for corresponding ranges of internal P content.
Figure 3. Observed and simulated distribution of cell quota of *Thalassiosira pseudonana* in laboratory culture. Symbols are data, and lines are model results. The simulations are named in the legend based on the mechanism removed (e.g. “qmax” has no variability in $q_{MAX}$)(Table 3).
Figure 4. Coefficients of variation of runs with individual mechanisms of heterogeneity removed. Error bars are +/- one standard deviation based on the variability in N=10 simulations (Figure 3 and Table 3).