Governing Equations

Updated 25 Aug 2016 to add O_2 constraint on reaction r_{PG}

Advection-Dispersion-Reaction Transport Model

The maximum entropy production problem for Siders Pond is constrained by the following vertical 1D advection-dispersion-reaction equation,

$$\frac{\partial \mathbf{c}(t,x)}{\partial t} = D(x)\frac{\partial^2 \mathbf{c}(t,x)}{\partial x^2} + \left(\frac{D(x)}{A(x)}\frac{\partial A(x)}{\partial x} + \frac{\partial D(x)}{\partial x} - \frac{q(x)}{A(x)}\right)\frac{\partial \mathbf{c}(t,x)}{\partial x} - \frac{\mathbf{c}(t,x)}{A(x)}\frac{\partial q(x)}{\partial x} + \frac{q_L(x)\mathbf{c}_L(x)}{A(x)} + \mathbf{r}(t,x,\boldsymbol{\varepsilon}), \tag{1}$$

with the following Neumann and Robin boundary conditions,

$$\text{At pond surface: } \left.\frac{\partial \mathbf{c}(t,x)}{\partial x}\right|_{x=0} = 0 \tag{2}$$

$$\text{At pond bottom: } \left(-D(x)A(x)\frac{\partial \mathbf{c}(t,x)}{\partial x} + q(x)\mathbf{c}(t,x)\right)_{x=d} = q(d)\mathbf{c}_B \tag{3}$$

where $\mathbf{c}(t,x)$ is a state vector of concentration variables (mmol m^{-3}) at time t and depth x, and includes phototrophic catalyst, $\mathbb{S}_A$ (see below), and $\mathbf{r}(t,x,\boldsymbol{\varepsilon})$ is a vector of reaction rates (mmol m^3 d^{-1}) in the given metabolic network whose values are determined by maximizing entropy production via adjustment of control variable vector, $\boldsymbol{\varepsilon}$. Siders Pond cross-sectional area (m^2) as a function of depth, $A(x)$, was determined from bathymetry plots, and the vertical volumetric flow rate (m^3 d^{-1}), $q(x)$, as well as lateral inputs (m^2 d^{-1}), $q_L(x)$, were determined from previous studies that quantified saltwater intrusion and freshwater entrainment that drives the vertical salinity gradient (Caraco 1986). It was assumed that the lateral inputs flow at an exponentially increasing rate towards the pond surface. The vertical dispersion coefficient (m^2 d^{-1}), $D(x)$, was numerically determined by fitting model output to observed salinity profile obtained from our 24 hr field sampling (see Fig GE-1 below). Numerical results support the previous conclusion that the meromictic nature of Siders Pond is a dynamic balance between fresh and saltwater inputs. The characteristic timescale for transport in Siders Pond is approximately 1 yr, as evident by the transient salinity dynamics simulated over 1000 days starting at a uniform salinity of 18 PSU (Fig GE-2). Numerical solution of the above partial differential equations (PDEs) was obtained using an adaptive collocation package based on B-splines (Pew et al. 2016).

Maximum Entropy Production (MEP) Model

Reaction rates in Eq. (1) over t and x are determined by maximizing entropy production over an interval of time (see Vallino et al. 2014) and either locally (at x) or globally over the domain $[0, d]$ (Vallino 2011). Since we have previously develop MEP network models for chemolithoautotrophs, chemolithoheterotrophs and chemoorganoheterotrophs (Vallino 2010, Algar & Vallino 2014), we have focused our initial efforts on phototrophs, as they are critical in Siders Ponds (and elsewhere of course), and phototrophy had not been explored within the MEP context.

To develop and test the MEP phototroph model, we consider a PO_4^{3-} limited system (as is Siders Pond) consisting of only phototrophs. Phototrophs are represented by two metabolic reactions. The first reaction,

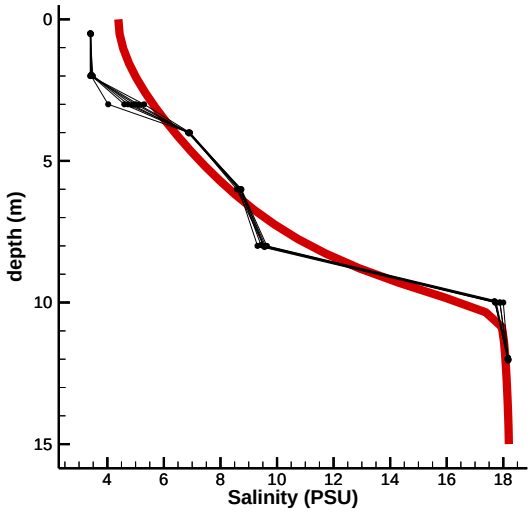

Fig. GE-1. Fit of modeled salinity (red line) to observed salinity (black circles connected by lines).

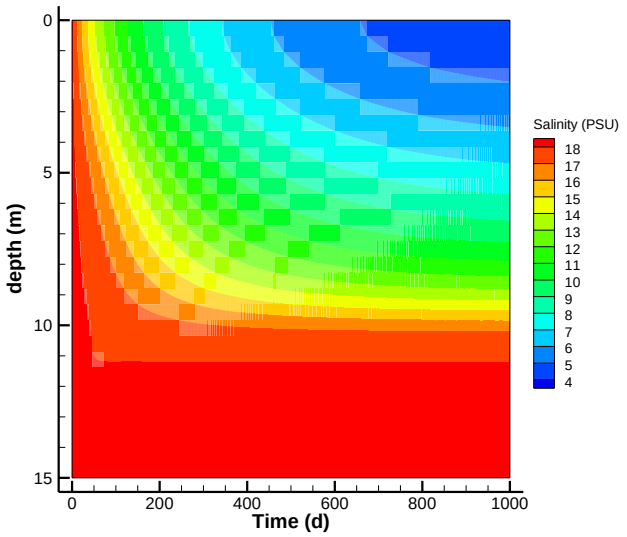

Fig. GE-2. Transient simulation of salinity starting with uniform salinity of 18 PSU.

r_{PP}, is the fixation rate of dissolved inorganic carbon (H_2CO_3) into simple sugar (CH_2O) driven by the dissipation of photosynthetic active radiation, $h\nu$, as given by,

$$r_{PP}\colon\ \varepsilon_{PP}H_2CO_3 + n_\gamma h\nu \xrightarrow{\mathbb{S}_P} \varepsilon_{PP}(C_P + O_2), \tag{4}$$

where ε_{PP} is the efficiency by which light is converted into chemical potential, C_P is the intracellular carbon reserve stored as CH_2O by algae, $\mathbb{S}_P$, but C_P does not contribute to catalytic activity, and n_γ is the mmol of photons captured per mmol of reaction extent, which is set to

$$n_\gamma = -\frac{\Delta_r G_{CH_2O}}{\Delta_r G_\gamma}. \tag{5}$$

The free energy of reaction for CH_2O synthesis, $\Delta_r G_{CH_2O}$ $(J\ mmol^{-1})$, and photon energy transfer, $\Delta_r G_\gamma$, are,

$$\Delta_r G_{CH_2O} = \Delta_r G^\circ_{CH_2O} + RTln\left(\frac{[C_P][O_2]}{[\mathrm{DIC}]}\right) \tag{6}$$

$$\Delta_r G_\gamma = -\eta_I h\nu N_A, \tag{7}$$

where η_I is the thermodynamic efficiency for the conversion of radiation to work, given by (Candau 2003),

$$\eta_I = 1 - \frac{4}{3}\frac{T_{env}}{T_{sun}} + \frac{1}{3}\left(\frac{T_{env}}{T_{sun}}\right)^4, \text{ where } T_{sun} = 5800\ K, T_{env} = 298\ K, \text{ so } \eta_I \cong 0.93, \tag{8}$$

$h\nu$ is Plank's constant times the frequency of light (we used just green light for simplicity), and N_A is Avogadro's number in mmol.

Consequently, n_γ is defined in Eq. (5) so that as the efficiency of solar energy capture, ε_{PP}, approaches 1, $\Delta_r G_{PP}$ for reaction r_{PP} approaches 0, as given by,

$$\Delta_r G_{PP} = \varepsilon_{PP}\Delta_r G_{CH_2O} + n_\gamma \Delta_r G_\gamma = -(1-\varepsilon_{PP})\Delta_r G_{CH_2O}. \tag{9}$$

And as ε_{PP} approaches 0, all light energy is just dissipated as heat producing entropy at the rate $-r_{PP}(t,x)\Delta_r G_{PP}/T$.

The reaction rate of carbon fixation, r_{PP}, is given by the following generalized kinetic model we have used in previous work (Vallino 2011, Algar & Vallino 2014) that consists of three terms, a maximum rate, a kinetic driver and a thermodynamic driver, as given by

$$r_{PP} = \frac{\Delta I_P}{n_\gamma}\left(\frac{c_{CO_2}}{c_{CO_2} + \kappa^* \varepsilon_{PP}^2}\right)\left(1 - e^{\Delta_r G_{PP}/RT\chi_{PP}}\right). \tag{10}$$

Here, the maximum rate, $\frac{\Delta I_P}{n_\gamma}$, is dictated by the rate of light interception by the catalytic fraction of the phototroph, $\mathbb{S}_P$. The differential equation over depth for light absorption is given by,

$$\frac{dI(t,x)}{dx} = -\left(k_W + k_P\left(c_{C_P}(t,x) + c_{\mathbb{S}_P}(t,x)\right)\right)I(t,x) = -\Delta I_T, \tag{11}$$

which is solved for $I(t,x)$ as part of the PDEs (1-3) above. The light absorbed by only the catalytic portion of the phototroph is,

$$\Delta I_P = k_P c_{\mathbb{S}_P}(t,x) I(t,x). \tag{12}$$

The light absorbed by phototrophs including that intercepted by carbon storage is then,

$$\Delta I_{PC} = k_P\left(c_{C_P}(t,x) + c_{\mathbb{S}_P}(t,x)\right)I(t,x) \tag{13}$$

The second metabolic reaction captures the synthesis of phototrophs, with elemental composition $\mathrm{CH}_{\alpha_P}\mathrm{O}_{\beta_P}\mathrm{N}_{\gamma_P}\mathrm{P}_{\delta_P}\mathrm{Si}_{\zeta_P}\mathrm{Fe}_{\xi_P}$, from stored carbon reserves and nutrients in the environment, as given,

$$\begin{aligned} r_{PG}\colon C_P + \varepsilon_{PG}\gamma_P \mathrm{NH_3} + \varepsilon_{PG}\delta_P H_3PO_4 + (1-\varepsilon_{PG})\mathrm{O_2} &\xrightarrow{\mathbb{S}_P} \varepsilon_{PG}x_{PG}\mathbb{S}_P + (1-\varepsilon_{PG}x_{PG})\mathrm{H_2CO_3} \\ &+ \varepsilon_{PG}y_{PG}\mathrm{H_2O} \end{aligned} \tag{14}$$

where the control variable, ε_{PG}, is the efficiency of converting store carbon to phototroph biomass versus its oxidation to CO_2 and water, and x_{PG} and y_{PG} are stoichiometric coefficients determined from conservation of H and O. The kinetics of this reaction is given by,

$$r_{PG} = \nu^* \varepsilon_{PG}^2 c_{\mathbb{S}_P}\left(\frac{c_{C_P}}{c_{C_P} + \kappa^* \varepsilon_{PG}^4}\right)\left(\frac{\frac{c_{PO4}}{\delta_P}}{\frac{c_{PO4}}{\delta_P} + \kappa^* \varepsilon_{PG}^4}\right)\left(\frac{c_{O_2}}{c_{O_2} + \kappa^* \varepsilon_{PG}^4}\right)\left(1 - e^{\Delta_r G_{PG}/RT\chi_{PG}}\right), \tag{15}$$

that is also based on kinetics previously developed. The standard reaction free energy, $\Delta_r G^\circ_{PG}$, is given by

$$\Delta_r G^{\circ}_{PG} = \varepsilon_{PG}\Delta_r G^{\circ}_{\mathbb{S}\ biosynthesis} - (1-\varepsilon_{PG})\Delta_r G^{\circ}_{CH_2O} \tag{16}$$

where $G^{\circ}_{\mathbb{S}\ biosynthesis}$ is the standard reaction free energy for biomass synthesis that we have previously calculated (Vallino 2010), and $\Delta_r G_{PG}$ is readily calculated based on activity coefficients with the appropriate exponents as dictated by Eq. (14).

For testing the above model formulation, we included as state variables salinity, dissolved oxygen, DIC, PO_4^{3-}, C_P and $\mathbb{S}_P$, so that the state vector of concentrations in Eq. (1) is,

$$\mathbf{c}^T(t,x) = \left[c_{Sal}, c_{DO}, c_{DIC}, c_{PO4}, c_{C_P}, c_{\mathbb{S}_P}\right]^T. \tag{17}$$

The reaction rate vector in Eq. (1) is given by,

$$\mathbf{r}(t,x) = \begin{bmatrix} r_{Sal} \\ r_{DO} \\ r_{DIC} \\ r_{PO4} \\ r_{C_P} \\ r_{\mathbb{S}_P} \end{bmatrix} = \begin{bmatrix} 0 \\ \varepsilon_{PP} r_{PP}(t,x) - (1-\varepsilon_{PG}) r_{PG}(t,x) \\ -\varepsilon_{PP} r_{PP}(t,x) + (1-\varepsilon_{PG} x_{PG}) r_{PG}(t,x) \\ -\varepsilon_{PP}\delta_P r_{PG}(t,x) \\ \varepsilon_{PP} r_{PP}(t,x) - r_{PG}(t,x) \\ \varepsilon_{PG} x_{PG} r_{PG}(t,x) \end{bmatrix} \tag{18}$$

Entropy Production Terms

Entropy production as a function of depth, $\dot{\sigma}(t,x)$ (J m^{-1} K^{-1} d^{-1}), associated with both reactions, but NOT including light dissipated by abiotic particles, is given by,

$$\dot{\sigma}_{PP}(t,x) = \frac{A(x)}{T}\Delta_r G_{CH_2O}\left(\frac{\Delta I_{PC}(t,x)}{n_\gamma} - \varepsilon_{PP}(t,x) r_{PP}(t,x)\right) \tag{19}$$

$$\dot{\sigma}_{PG}(t,x) = -\frac{A(x)}{T} r_{PG}(t,x)\Delta_r G_{PG} \tag{20}$$

Total entropy production per unit depth, $\dot{\sigma}_T(t,x)$ (J m^{-1} K^{-1} d^{-1}), which is the quantity to be maximized, includes light dissipation by abiotic particles, is given by,

$$\dot{\sigma}_T(t,x) = \frac{A(x)}{T}\left(\Delta_r G_{CH_2O}\left(\frac{\Delta I_T(t,x)}{n_\gamma} - \varepsilon_{PP}(t,x) r_{PP}(t,x)\right) - r_{PG}(t,x)\Delta_r G_{PG}\right) \tag{21}$$

Optimization Problem

To determine how the control variables $\varepsilon_{PP}(t,x)$ and $\varepsilon_{PG}(t,x)$ vary over time and space, we setup a receding horizon optimal control problem (Vallino et al. 2014) in which $\dot{\sigma}_T(t,x)$ is maximized over successive intervals of time, Δt, and is either maximized locally, as

$$\max_{\varepsilon_{PP},\varepsilon_{PG}} \int_{t_i}^{t_i+\Delta t} \dot{\sigma}_T(\tau, x_j)\, d\tau \quad \forall j \text{ in the control grid } [x_1, x_2, \ldots, x_n], \tag{22}$$

or globally over the depth of the water column, d, as,

$$\max_{\varepsilon_{PP},\varepsilon_{PG}} \int_{\chi=0}^{\chi=d} \int_{t_i}^{t_i+\Delta t} \dot{\sigma}_T(\tau,\chi)d\tau\, d\chi \tag{23}$$

where either optimization is subject to the constraints imposed by Eqs. (1-3) and Eq. (18). However, when we constructed the global optimization problem as given by Eq. (23), an obvious problem arose. In this phototroph only problem, entropy production cannot exceed the rate at which light is being dissipated in the water column as given by Eqs. (11) and (21). If the water column is sufficiently deep for a given level of turbidity given by k_W, then all the light will be dissipated before it reaches the bottom of the pond, as is the case for Siders Pond. Consequently, the addition of phytoplankton to the system cannot increase the global entropy production, because it is already maximized by light absorption by water. However, if entropy production is maximized locally as in Eq. (22), then increased phototroph concentrations in the surface layers will increase local entropy production. ***This leads us to the conclusion that for at least simple phototrophic systems, entropy production must be maximized locally***, which is not the case for chemically based systems as discussed in previous work (Vallino 2011).

Model Simulations

Here we present results for the simple phototroph model described above. As mentioned above, entropy is maximized locally as given by Eq. (22). For the numerical solution, two computational grids are used, one for solution of the PDEs given by Eqs. (1-3) and (18), and another, lower density, grid, $g_{OC}(t_i,x_j)$ where values of $\varepsilon_{PP}(t,x)$ and $\varepsilon_{PG}(t,x)$ are set by the optimization use to solve Eq. (22). For local optimization, entropy is maximized as a given depth, x_j, that corresponds to the spatial grid points in $g_{OC}(t_i,x_j)$. Once a solution is found for a particular depth, the next depth in $g_{OC}(t_i,x_j)$ is then optimized. Once all levels have been optimized, the procedure repeats until either all levels no longer change to a specified precision, or a maximum number of iterations is reached. The receding horizon optimal control uses two time intervals, one that corresponds to an "infinite", or long, time scale, Δt_∞, and a shorter "finite" update interval, Δt_F. While the spacial grid points in $g_{OC}(t_i,x_j)$ are set at fixed values, the time grid points span the current infinite interval. Once a solution to the infinite interval is obtained, the solution is only propagated over the shorter finite interval, Δt_F. Once the first interval is completed, the next interval is begun, and the solution marches forward a sufficient number of finite intervals until the desired simulation time is reached. See Vallino et. al (2014) for more details. For the optimization of Eq. (22), we used L-BFGS-B, which was developed to solve large-scale bound constrained optimization problems (Morales & Nocedal 2011). We used finite differences to construct the gradient needed by L-BFGS-B, but we parallelized the code to take advantage of multi CPU architectures using OpenMP, but we well be moving to MPI to take advantage of massively parallel architectures.

For the simulation results presented below, a finite interval of 2 days and an infinite interval of 4 days was used, where the optimal control grid, $g_{OC}(t_i,x_j)$, consisted of 16 time points spaced exponentially over the infinite interval and 10 depth points at 0, 1, 2, 3, 4, 6, 8, 10, 12, and 15 m. The simulations were run over three finite intervals, or 6 days corresponding to June 25 to July 1. The simulations were started from uniform initial conditions over depth, so these results are only being used for testing and demonstration. ***We note that all of the simulated outputs below result purely from the maximization of local entropy production under the imposed constraints. There are few adjustable parameters in this model.***

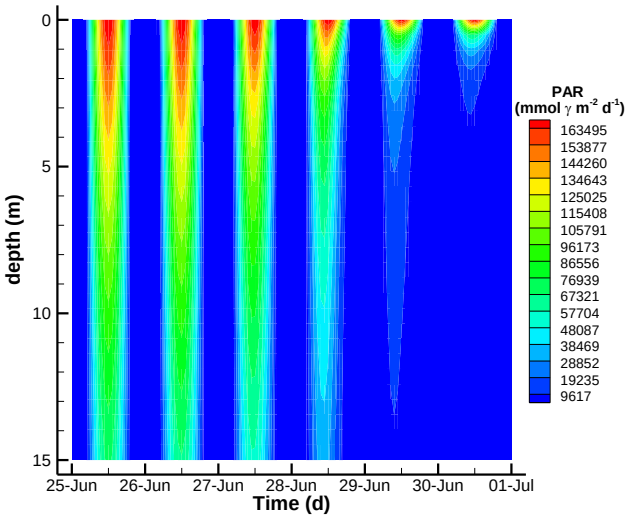

Fig. GE-3. Intensity of photosynthetic active radiation over time and depth. As phototroph population increases, more light is intercepted and dissipated as entropy in the surface layers of the pond.

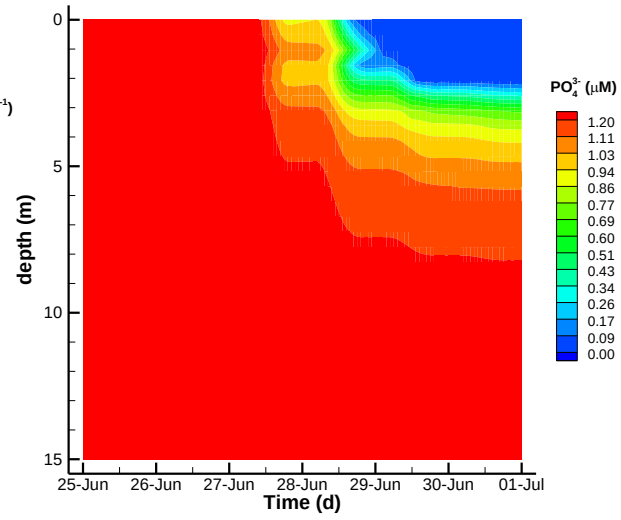

Fig. GE-4. Simulated phosphate concentration over time and depth. Phosphate becomes limiting in the surface layers as phototroph population increases.

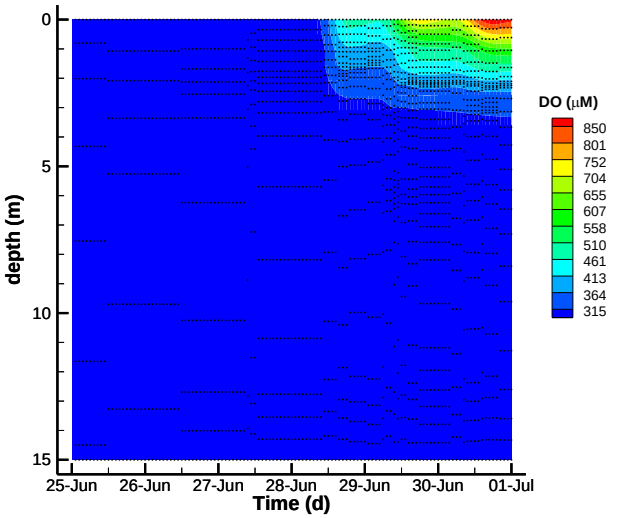

Fig. GE-5. Dissolved oxygen concentrations over time and depth. As expected, the surface layers become super saturated in O_2 as phototrophs grow. However, air-water gas exchange is currently not implemented in the model. Also included in this simulation are the grid points (black dots) used in the PDE solution grid that change both in location and in number as steep gradients develop in space. Time stepping by BACOLI95 is also dynamic, but not shown here.

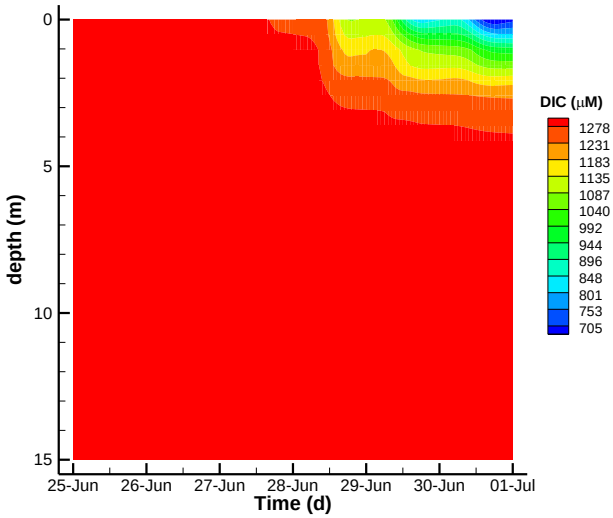

Fig. GE-6. Dissolved inorganic carbon over time and depth. Because of CO_2 fixation by phototrophs, decreases in DIC correspond to where and when phototrophs are growing. Note, the current version of the model does not model carbonate chemistry, which will be added to later versions of the model.

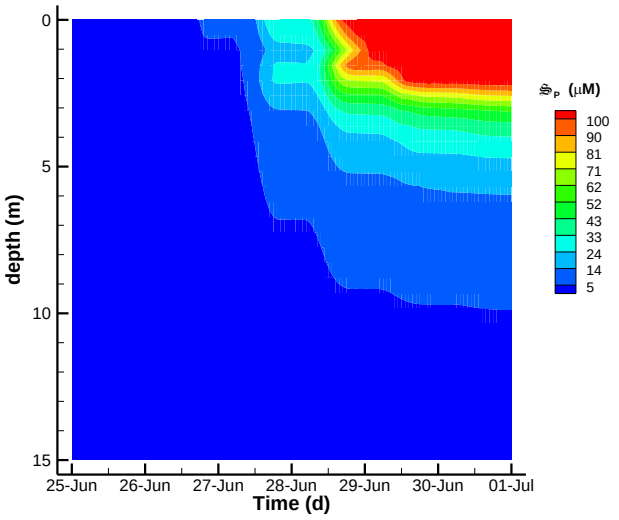

Fig. GE-7. Concentration of phototroph catalytic biomass, $\mathbb{S}_P$, over time and depth. It take a few days before biomass accumulates to significant levels from the initial conditions of 0.1 μM on 25 Jun.

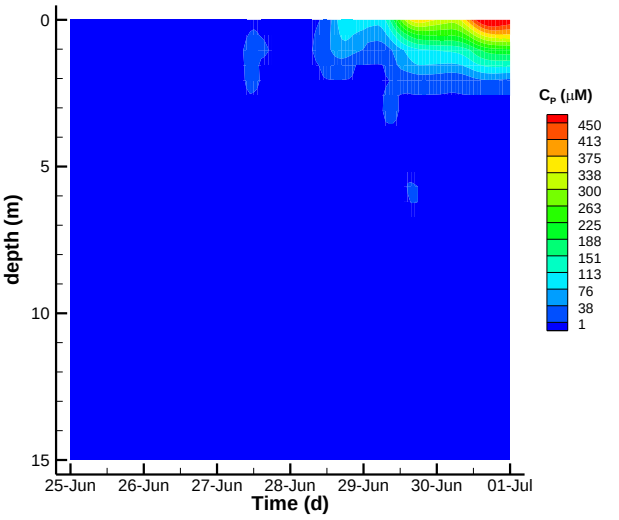

Fig. GE-8. Concentration of stored carbon reserves by phototrophs produced by reaction r_{PP}. This carbon is used for synthesis of phototroph catalytic biomass by reaction r_{PG}. Note, some diel oscillations in concentration are evident, and peak concentrations appear to occur after peak solar radiation at noon.

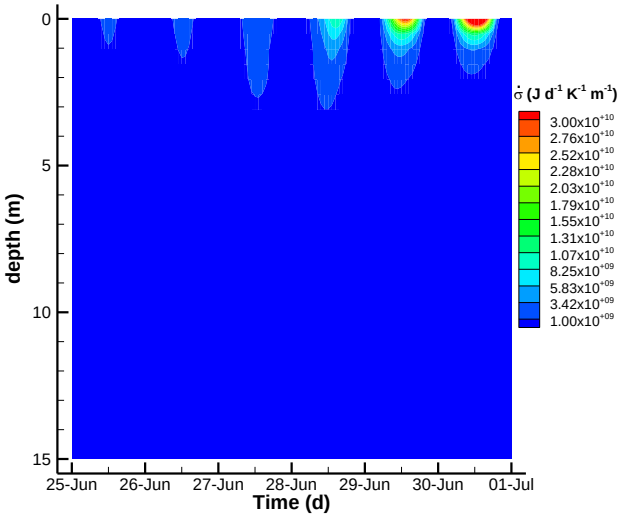

Fig. GE-9. Entropy production rate that resulted from local optimization. As biomass accumulates at the surface, more light is locally absorbed and dissipated as heat. Compare to PAR intensity shown in Fig. GE-3.

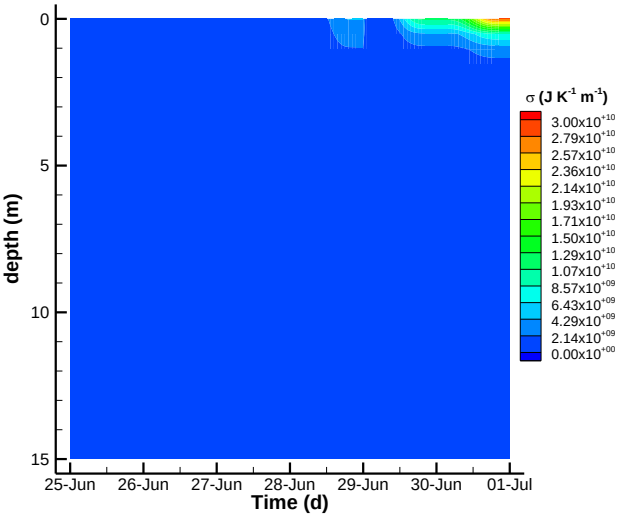

Fig. GE-10. Entropy produced over each finite interval. At the start of each interval total entropy produced is reset to zero, and the integrated value at the end of the interval is what is maximized at a given depth in Eq. (22).

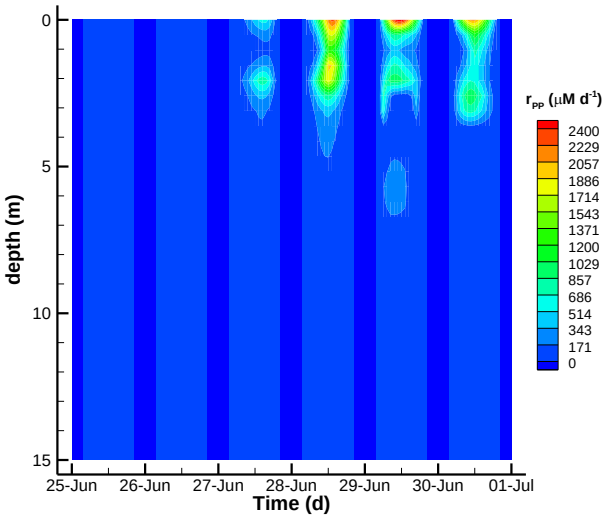

Fig. GE-11. Rate of CO_2 fixation by reaction $\boldsymbol{r_{PP}}$. Since this reaction can only occur when light is present, at night or in the shade the rates go to zero. It is not clear at this time why a bimodal distribution over depth occurs in the simulation.

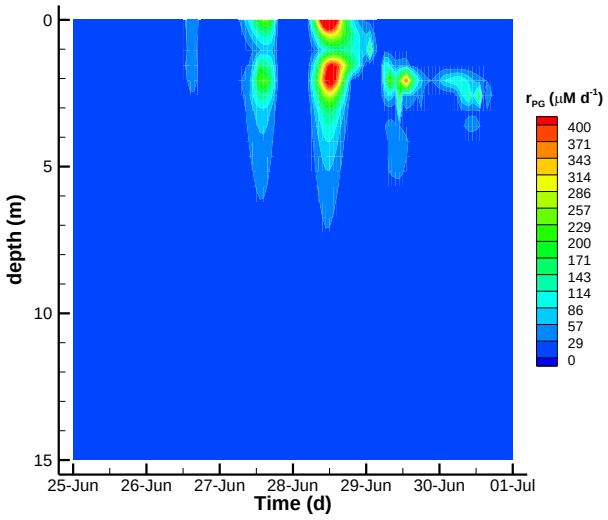

Fig. GE-12. Rate of phototroph catalytic biomass synthesis from internal C storage as given by reaction $\boldsymbol{r_{PG}}$. As PO_4^{3-} becomes limiting in the surface water, the rate of reaction is reduced, but can occur at deeper locations in the water column where P and $\boldsymbol{C_P}$ are still available.

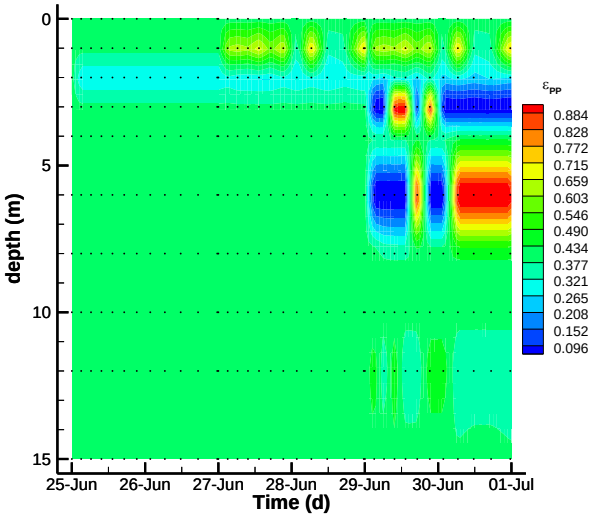

Fig. GE-13. Control variable $\boldsymbol{\varepsilon_{PP}}$ that determines how much of the intercepted light is used to fix CO_2 versus being dissipated as heat. This figure also shows the grid used by the optimal control problem (black dots). Only the control variables are changed by the optimal control solution in order to maximize entropy production. All other variables result from adjusting $\boldsymbol{\varepsilon_{PP}}$ and $\boldsymbol{\varepsilon_{PG}}$

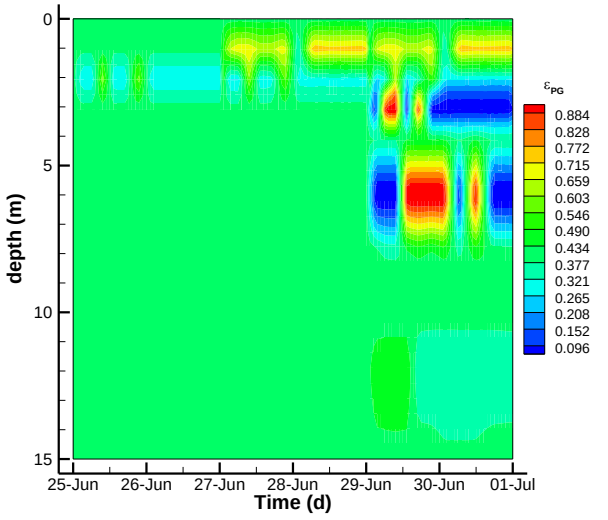

Fig. GE-14. Control variable $\boldsymbol{\varepsilon_{PG}}$ that determines how much of the fixed carbon is converted to phototroph biomass versus being oxidized into CO_2 and water. The four graphs on this page (Figs. GE-11-14) are the variables that best map to our metagenomics and metatranscriptomics data that have been collected over similar spatial and temporal scales in Siders Pond. The switching shown in these figures should correspond to changes in gene abundance and messenger levels if the model is accurately capturing the phototrophic community. However, these are just example simulations at this point.

References

Algar, C. K., and J. J. Vallino. 2014. Predicting microbial nitrate reduction pathways in coastal sediments. *Aquat.Microb.Ecol.* **71**:223-238, doi:10.3354/ame01678.

Candau, Y. 2003. On the exergy of radiation. *Solar Energy* **75**:241-247, doi:10.1016/j.solener.2003.07.012.

Caraco, N. 1986. Phosphorus, iron, and carbon cycling in a salt stratified coastal pond. Boston University, Boston.

Morales, J. L., and J. Nocedal. 2011. Remark on "algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound constrained optimization". *ACM Trans. Math. Softw.* **38**:1-4, doi:10.1145/2049662.2049669.

Pew, J., Z. Li, and P. Muir. 2016. Algorithm 962: BACOLI: B-spline Adaptive Collocation Software for PDEs with Interpolation-Based Spatial Error Control. *ACM Trans. Math. Softw.* **42**:1-17, doi:10.1145/2818312.

Vallino, J. J. 2010. Ecosystem biogeochemistry considered as a distributed metabolic network ordered by maximum entropy production. *Philosophical Transactions of the Royal Society B: Biological Sciences* **365**:1417-1427, doi:10.1098/rstb.2009.0272.

Vallino, J. J. 2011. Differences and implications in biogeochemistry from maximizing entropy production locally versus globally. *Earth Syst.Dynam.* **2**:69-85, doi:10.5194/esd-2-69-2011.

Vallino, J. J., C. K. Algar, N. F. González, and J. A. Huber. 2014. Use of Receding Horizon Optimal Control to Solve MaxEP-Based Biogeochemistry Problems. Pages 337-359 *in* R. C. Dewar, C. H. Lineweaver, R. K. Niven, and K. Regenauer-Lieb, editors. Beyond the Second Law - Entropy production and non-equilibrium systems. Springer Berlin Heidelberg, doi:10.1007/978-3-642-40154-1_18.