

Manuscript prepared for Geosci. Model Dev.
with version 2014/09/16 7.15 Copernicus papers of the L^AT_EX class copernicus.cls.
Date: 22 August 2015

Addressing the Challenge of Nonphysical Solution Negativity in Coupling a Reactive Transport Code with a Global Land Surface Model for Mechanistic Biogeochemistry Representation

Guoping Tang¹, Fengming Yuan¹, Gautam Bisht^{1,2}, Glenn E. Hammond³, Peter C. Lichtner⁴, Jitendra Kumar¹, Richard T. Mills^{1,5}, Xiaofeng Xu^{1,6}, Ben Andre⁷, Forrest M. Hoffman¹, Scott L. Painter¹, and Peter E. Thornton¹

¹Oak Ridge National Laboratory, Oak Ridge, Tennessee, United States

²Lawrence Berkeley National Laboratory, Berkeley, California, United States

³Sandia National Laboratories, Albuquerque, New Mexico, United States

⁴OFM Research, Redmond, Washington, United States

⁵Intel Incorporation, Portland, Oregon, United States

⁶University of Texas at El Paso, El Paso, Texas, United States

⁷National Center for Atmospheric Research, Boulder, Colorado, United States

Correspondence to: Peter E. Thornton (thorntonpe@ornl.gov)

This manuscript has been authored by UT-Battelle, LLC under Contract No. DE-AC05-00OR22725 with the U.S. Department of Energy. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes. The Department of Energy will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan(<http://energy.gov/downloads/doe-public-access-plan>).

Abstract. Reactive transport codes (e.g., PFLOTTRAN) are increasingly used to improve the representation of biogeochemical processes in terrestrial ecosystem models (e.g., the Community Land Model, CLM). As CLM and PFLOTTRAN use explicit and implicit time stepping, implementation of CLM biogeochemical reactions in PFLOTTRAN can result in negative primary species concentrations, which is not physically meaningful. The objective of this work is to address the nonphysical solution negativity to obtain accurate, efficient, and robust solutions. We illustrate the implementation of a reaction network with the CLM-CN decomposition, nitrification, denitrification, and plant nitrogen uptake reactions and test the implementation at arctic, temperate, and tropical sites. We examine use of scaling back the update during each iteration (SU), log transformation (LT), and down-regulating the reaction rate to account for reactant availability limitation to enforce nonnegativity. Both SU and LT guarantee nonnegativity but with implications. When a very small scaling factor occurs due to either consumption or numerical overshoot, and the iterations are deemed converged because of too small an update, SU can introduce excessive numerical error. LT involves multiplication of the Jacobian matrix by the concentration vector, which increases the condition number, decreases the time step size, and increases the computational cost. Neither SU nor LT prevents zero concentration. When the concentration is close to machine precision or zero, a small positive update stops all reactions for SU, and LT can fail due to a singular Jacobian matrix. The consumption rate has to be downregulated such that the solution to the mathematical representation is positive. A first-order rate downregulates consumption and is nonnegative, and adding a residual concentration makes it positive. For zero-order rate (the reaction rate is not a function of a reactant), representing the availability limitation of each reactant with a Monod substrate limiting function provides a smooth transition between a zero-order rate when the reactant is abundant and first-order rate when the reactant becomes limiting. When the half saturation is small, marching through the transition may require small time step sizes to resolve the sharp change within a small range of concentration values. Our results from simple tests and CLM-PFLOTTRAN simulations caution against use of SU and indicate that accurate, stable, and relatively efficient solutions can be achieved with LT and downregulation with Monod substrate limiting function and residual concentration.

35 **1 Introduction**

Land surface models calculate the fluxes of energy, water, and green house gases across the land-atmosphere interface for the atmospheric general circulation models for climate simulation and weather forecasting. Evolving from the first generation “bucket”, second generation “biophysical”, and third generation “physiological” models (Sellers et al., 1997; Seneviratne et al., 2010), current
40 land surface models, e.g., the Community Land Model (CLM), implement comprehensive thermal, hydrological, and biogeochemical processes (Oleson et al., 2013). The important role of soil bio-geochemistry is suggested by the confirmation that the increase of CO₂, CH₄, and N₂O in the atmosphere since the preindustrial times is the main driving cause of climate change, and inter-dependent water, carbon and nitrogen cycles in terrestrial ecosystems are very sensitive to climate
45 changes (IPCC, 2013). CLM4.5 incorporates CLM-CN and CENTURY, and adds methane models for carbon and nitrogen cycles (Oleson et al., 2013). In addition to ~ 250 soil biogeochemical models developed in the past ~ 80 years (Manzoni and Porporato, 2009), increasingly mechanistic models continue to be developed to increase the fidelity of process representation for improving climate prediction (e.g., Wang et al., 2012; Riley et al., 2014).

50 As land surface models usually hardcode the reaction network (pools/species, reactions, rate formulae), substantial effort is often required to revise the source code for testing alternative biogeochemical models, and incorporating new process understanding. To mitigate this issue, Fang et al. (2013) demonstrates the use of a reaction-based approach to facilitate implementation of CLM-CN and CENTURY models, and incorporation of phosphorus cycle. While Fang et al. (2013) adds a
55 reaction solver developed from the reactive transport (geochemical) model literature for algebraic and ordinary differential equations, Tang et al. (2013b) solves the advection diffusion equation in CLM using operator splitting with the Crank-Nicolson scheme for diffusion and a forward-in-time upstream discretization for advection. In contrast, a reactive transport code, TOUGH-REACT, is used to develop multi-phase (gas, aqueous, solid) mechanistic carbon and nitrogen cycle models
60 with many speciation and microbial reactions (Maggi et al., 2008; Gu and Riley, 2010; Riley et al., 2014). Coupling a reactive transport code (e.g., Steefel et al., 2014) with CLM facilitates testing and implementation of increasingly mechanistic biogeochemical models in terrestrial ecosystem models by taking advantages of the developments in reactive transport modeling.

The nonphysical solution negativity arises in the coupling of a reactive transport code with CLM
65 as an essential aspect of these land surface models is the ability to simulate competition for nutrients (e.g., mineral nitrogen, phosphorus, etc.) among plants and microbes. For example, the CLM-CN decomposition cascade downregulates the demand based on the available nitrogen (N) (Oleson et al., 2013; Thornton and Rosenbloom, 2005). Specifically, marching from time step k to $k+1$ with a supply rate S^k and consumption rate D^k using the forward difference, $[N]^{k+1} = [N]^k + (S^k - D^k)\Delta t$.
70 ($[]$ is used to denote concentration.) CLM replaces D^k with $\min(D^k, [N]^k/\Delta t)$ so that $[N]^{k+1} \geq S^k\Delta t \geq 0$. As a result, $[N]$ is nonnegative in CLM (Tang and Riley, 2015).

Geochemical codes generally use implicit time stepping such as the backward Euler method, $([N]^{k+1} - [N]^k)/\Delta t = S^{k+1} - D^{k+1}$. Solving this nonlinear equation with the Newton-Raphson method to iterate from p to $p + 1$, the residual $f = ([N]^{k+1,p} - [N]^k)/\Delta t - S^{k+1,p} + D^{k+1,p}$, the derivative $f' = 1/\Delta t - \partial S^{k+1,p}/\partial[N]^{k+1,p} + \partial D^{k+1,p}/\partial[N]^{k+1,p}$, the update $\delta = f/f'$, and $[N]^{k+1,p+1} = [N]^{k+1,p} - \delta$. Depending on $[N]^k$, Δt , $S^{k+1,p}$, $D^{k+1,p}$, and the derivatives, $[N]^{k+1,p+1}$ can be negative, which is not physical, and can cause numerical instability and errors (Shampine et al., 2005). With instability, the simulation may abort before reaching target simulation duration. Excess numerical error may jeopardize the improvements in process representation. Avoiding numerical instability and error may result in small time step sizes or high computational cost. Therefore, it is necessary to address the nonphysical solution negativity for accurate, robust, and efficient numerical simulation.

Enforcing nonnegativity is a common challenge in science, engineering, and business, including, for example, image processing, optimization, and ecosystem and geochemical modeling (Antonelli et al., 2009; Broekhuizen et al., 2008; Bruggeman et al., 2007; Burchard et al., 2003, 2005; Chen and Plemmons, 2009; Pierre and Schmitt, 2000; Shampine et al., 2005). The challenge increases in geochemical modeling because the concentration can be very low. For example, the threshold concentration for H_2 is 1.5 nM (10^{-9} M) for dechlorinators and 5~20 nM for methanogens (Fennell and Gossett, 1998). The EPA maximum contaminant level for drinking water for dioxin is 30 ppq (3.5×10^{-13} M). The redox potential Eh needs to be decreased to -0.35 V (corresponding to an O_2 concentration $< 10^{-22}$ M Hungate, 1975) for methanogens to grow (Jarrell, 1985). With very low concentration, a small consumption or numerical overshoot can lead to negative concentration.

Two methods are used to avoid negative concentration in geochemical codes. One is to use the logarithm concentration as the primary variable (Bethke, 2007; Hammond, 2003; Parkhurst and Appelo, 1999). Log transformation (LT) is well suited for solving the mass action equations because it converts these nonlinear equations into linear equations. However, LT converts linear advection and diffusion equations into nonlinear equations, which may increase the computational cost (Hammond, 2003). The other approach scales back the update in each of the Newton-Raphson iteration (SU) to enforce nonnegativity (Bethke, 2007; Hammond, 2003). Both methods are available in PFLOTRAN (Lichtner et al., 2015) (with SU as the default) and some other geochemical codes (e.g., Geochemist's Workbench, Bethke, 2007). However, to our knowledge, the implications of these approaches have not been thoroughly examined, particularly for application to CLM.

Competition for nutrients (for instance, labile carbon, phosphate, O_2 , and H_2) is common in terrestrial ecosystems, and is increasingly incorporated in process-rich models. As the terrestrial ecosystem models are often run under a variety of conditions around the globe for hundreds of years with a time step as small as half an hour, proper treatment of the nonphysical solution negativity is necessary for accurate, efficient, and robust simulation of biogeochemical processes using reactive transport codes for earth system models. The objective of this work is to implement CLM subsurface biogeochemical reactions in CLM-PFLOTRAN with the focus on addressing the nonphysical

solution negativity to obtain accurate, efficient, and robust solutions. We implement the CLM-CN
110 decomposition (Bonan et al., 2012; Oleson et al., 2013; Thornton and Rosenbloom, 2005); nitrification;
denitrification (Dickinson et al., 2002; Parton et al., 2001, 1996); and plant nitrogen uptake
reactions in CLM-PFLOTRAN and test the implementation at arctic, temperate, and tropical sites.
In addition to SU and LT, we examine ways to downregulate consumption rate to account for the
limitation of reactant availability on reaction rate. This work focuses on addressing the nonphysical
115 solution negativity for biogeochemistry, comprehensive CLM-PFLOTRAN coupling in heat transfer
(including freeze and thaw), hydrology, and biogeochemistry will be presented in future publications.
While we use CLM-PFLOTRAN to implement and test simple carbon and nitrogen reactions
at a few sites, we hope that what we develop here will be relevant for a wide range of applications.

2 CLM-PFLOTRAN biogeochemistry

120 The terrestrial ecosystem models generally include biogeochemical reactions for carbon and nitro-
gen cycles, in particular, the organic matter decomposition, nitrification, denitrification and methane
production and oxidation. The kinetics are usually described by a first-order rate modified by re-
sponse functions for environmental variables (temperature, moisture, pH, etc.) (Bonan et al., 2012;
Boyer et al., 2006; Schmidt et al., 2011). In this work, we use the CLM-CN decomposition (Bo-
125 nan et al., 2012; Oleson et al., 2013; Thornton and Rosenbloom, 2005), nitrification, denitrification
(Dickinson et al., 2002; Parton et al., 2001, 1996), and plant nitrogen uptake reactions (Fig. 1) as an
example. The reactions and rate formulae are detailed in Appendix A.

In CLM-PFLOTRAN, CLM instructs PFLOTRAN to solve the partial differential equations for
energy (including freezing and thawing), water flow, and reaction and transport in the surface and
130 subsurface. This work focuses on the biogeochemistry. Specifically, we focus on addressing the
nonphysical solution negativity for the geochemical reactions, with CLM solving the energy and
water flow equations and handling the solute transport (mixing, advection, diffusion, and leaching).

In each time step, CLM provides production rates for Lit1C, Lit1N, Lit2C, Lit2N, Lit3C, Lit3N
for litter fall; CWDC, and CWDN for coarse woody debris production, NH_4^+ and NO_3^- for nitro-
135 gen deposition and fixation; and plant N demand (rate); and specifies liquid water content, matrix
potential and temperature for PFLOTRAN; PFLOTRAN solves the ordinary differential equations
for the kinetic reactions, the mass action equations for the equilibrium equations, and provides the
final concentrations back to CLM.

PFLOTRAN does not track individual reaction rates such as total nitrogen mineralization rate. We
140 add optional hypothetical species (diagnostic variables), e.g., PlantA, PlantN, N_2Od , and DeniN,
to track NO_3^- uptake by plants as in reactions (R13,R14); N_2O production from nitrification reaction
(R11) due to net mineralization (Eq. A5); and denitrification (R12). At the end of each time step,
CLM uses the change of these concentrations to calculate the specific rates. To simplify the calcu-

lations, we reset these concentrations to 10^{-10} at the beginning of each time step instead of storing
145 the values of the previous time step.

The reactions and rates in Appendix A are implemented using the “reaction sandbox” concept
in PFLOTRAN (Lichtner et al., 2015). For each reaction, we specify a rate and a derivative of the
rate with respect to any components in the rate formula, given concentrations, temperature, moisture
content, grid cell volume, and other environmental variables. PFLOTRAN accumulates these rates
150 and derivatives into a residual vector and a Jacobian matrix, and the global equation is discretized
in time using the backward Euler method and solved using the Newton-Raphson method. Ignoring
equilibrium reactions and transport for simplicity of discussion in this work, PFLOTRAN solves the
ordinary differential equation,

$$dc/dt = \mathbf{R}(\mathbf{c}), \quad (1)$$

155 with \mathbf{c} as the concentration vector and \mathbf{R} as the kinetic reaction rate. Discretizing Eq. (1) in time
using the backward Euler method,

$$(\mathbf{c}^{k+1} - \mathbf{c}^k)/\Delta t = \mathbf{R}(\mathbf{c}^{k+1}). \quad (2)$$

Solving the equation using the Newton-Raphson method, we denote the residual as

$$\mathbf{f}(\mathbf{c}^{k+1,p}) = (\mathbf{c}^{k+1,p} - \mathbf{c}^k)/\Delta t - \mathbf{R}(\mathbf{c}^{k+1,p}), \quad (3)$$

160 and Jacobian as

$$\mathbf{J} = \frac{\partial \mathbf{f}(\mathbf{c}^{k+1,p})}{\partial \mathbf{c}^{k+1,p}}, \quad (4)$$

the update is

$$\delta \mathbf{c}^{k+1,p} = \mathbf{J}^{-1} \mathbf{f}(\mathbf{c}^{k+1,p}), \quad (5)$$

and the iteration equation is

$$165 \quad \mathbf{c}^{k+1,p+1} = \mathbf{c}^{k+1,p} - \delta \mathbf{c}^{k+1,p}. \quad (6)$$

The iteration continues until either the residual $\mathbf{f}(\mathbf{c}^{k+1,p})$ or the update $\delta \mathbf{c}^{k+1,p}$ is less than a specified tolerance. Specifically,

$$\|\mathbf{f}(\mathbf{c}^{k+1,p})\|_2 < \text{ATOL}, \quad (7)$$

$$170 \quad \frac{\|\mathbf{f}(\mathbf{c}^{k+1,p})\|_2}{\|\mathbf{f}(\mathbf{c}^{k+1,0})\|_2} < \text{RTOL}, \quad (8)$$

or

$$\frac{\|\delta \mathbf{c}^{k+1,p}\|_2}{\|\mathbf{c}^{k+1,p}\|_2} < \text{STOL}. \quad (9)$$

- If none of these tolerances are met in MAXIT iterations or MAXF function evaluations, the iteration is considered to diverge, and PFLOTRAN decreases the time step size for MAX_CUT times.
- 175 The default values in PFLOTRAN are ATOL = 10^{-50} , RTOL = 10^{-8} , STOL = 10^{-8} , MAXIT = 50, MAXF = 10^4 , and MAX_CUT = 16.

Unlike the explicit time stepping in CLM for biogeochemistry where only the reaction rates need to be calculated, the implicit time stepping requires the derivatives. While PFLOTRAN provides the option to calculate the derivatives numerically, analytical derivative calculation is generally preferred (e.g., Xu et al., 2006) because numerical calculation for accurate Jacobian approximation is a notoriously difficult task (Shampine et al., 2005).

Many reactions can be specified, and the rates and derivatives are accumulated in the residual and Jacobian, providing flexibility in specifying various reactions with a user-defined rate formula. As typical rate formulae consist of first order, Monod, and/or inhibition terms, a general rate formula 185 with flexible number of terms and typical moisture, temperature, and pH response functions are coded in PFLOTRAN. Most of the biogeochemical reactions can be specified in the input file, with flexible number of reactions, pools (species), rate terms, and various response functions without source code modification. Code modification is necessary only when different rate formulae, or response functions are introduced. In contrast, the number of pools and reactions are traditionally 190 hard-coded in CLM. Consequently, any change of the pools, reactions, or rate formula may require source code modification. Therefore, this new approach facilitates implementation of increasingly mechanistic reactions and tests of various representations with less code modifications.

3 Approaches for addressing the nonphysical solution negativity

One of the basic challenges for using a numerical geochemical formulation for CLM is that the 195 updated solution in Eq. (6) is not guaranteed to be nonnegative. Both SU (scaling back the update during each iteration) and LT (log transformation) are available in PFLOTRAN to enforce nonnegativity. However, to our knowledge, the limitations and implications of both methods have not been thoroughly examined.

3.1 Scaling back update in iterations

200 SU scales back the update (Eq. 6) with a scaling factor λ (Bethke, 2007; Hammond, 2003) such that

$$\mathbf{c}^{k+1,p+1} = \mathbf{c}^{k+1,p} - \lambda \delta \mathbf{c}^{k+1,p} > 0, \quad (10)$$

where

$$\lambda = \alpha \min [1, \mathbf{c}^{k+1,p}(i)/\delta \mathbf{c}^{k+1,p}(i)] \quad (11)$$

205 for positive $\delta\mathbf{c}^{k+1,p}(i)$ and $i = 1$ to m , with m as the number of species times the number of numerical grid cells. With SU, the concentration of the species that is going negative decreases by $(1 - \alpha)$ times in each iteration instead. This can be shown by solving the zero-order uptake problem $dc/dt = -1$: $c^{k+1} = c^k - \Delta t$ when $\Delta t \geq c^k$; otherwise, $\lambda = \alpha c^k / \Delta t$, and $c^{k+1} = (1 - \alpha)c^k$. With a default $\alpha = 0.99$ in PFLOTRAN, each iteration decreases the concentration by 100 times.

210 With multiple iterations, the concentration can approach machine precision or zero (we use zero for concentration below machine precision in this work).

If $\mathbf{c}^{k+1,p}(i)$ is 0, and $\delta\mathbf{c}^{k+1,p}(i) > 0$, then $\lambda = 0$, and the update is limited to 0. Even if $\lambda = 0$ occurs in only one grid cell for only one species, the update for all of the species and all of the grid cells are constrained to 0, and the residual will not be decreased to satisfy Eqs. (7 or 8) to achieve 215 convergence. If the iteration is deemed converged because the scaled update is 0, and Eq. (9) is met, the calculation will march through this time step without any change in the concentrations, numerically stopping all reactions. This can continue for many time steps until $\delta\mathbf{c}^{k+1,p}(i)$ becomes negative.

Application of a very small scaling factor λ due to decrease of a small concentration for one 220 species may have a similar consequence: the iteration may not decrease the residual $f(\mathbf{c}^{k+1,p})$ or the iteration may not converge because Eqs. (7 or 8) are not met. If the iteration is considered converged because of too small an update $\lambda\delta\mathbf{c}^{k+1,p}$ (Eq. 9 is satisfied), the scaling factor λ (e.g., 10^{-10}) may correctly limit the consumption reaction rates to account for availability limitation, but wrongly limit the production rate in the time step. This can be illustrated by adding to the zero-order uptake 225 problem $dc/dt = -1$ an independent zero order production $de/dt = 1$: when $\Delta t \geq c^k$, $\lambda = \alpha c^k / \Delta t$, $c^{k+1} = (1 - \alpha)c^k$, $e^{k+1} = e^k + \alpha c^k$ rather than $e^{k+1} = e^k + \Delta t$. If $c^k = 0$, SU numerically stops the independent production reaction. To avoid excessive numerical error, PFLOTRAN reports an error and exits when $\lambda < \lambda_{\min}$, with default $\lambda_{\min} = 10^{-10}$, translating the accuracy issue into a stability issue. It is necessary to investigate SU to resolve accuracy and stability issues.

230 3.2 Log transformation

LT is widely used in geochemical codes for describing highly variable concentrations for primary species such as H^+ or O_2 that can vary over many orders of magnitude as pH or redox state changes without the need to use variable switching. It is also used to enforce positivity (Bethke, 2007; Hammond, 2003; Parkhurst and Appelo, 1999). Instead of solving Eq. (3) for \mathbf{c}^{k+1} using Eqs. (4,5,6), LT 235 solves for $(\ln \mathbf{c}^{k+1})$ with

$$\mathbf{J}_{\ln}(i,j) = \frac{\partial \mathbf{f}(i)}{\partial \ln(\mathbf{c}(j))} = \mathbf{c}(j) \frac{\partial \mathbf{f}(i)}{\partial \mathbf{c}(j)} = \mathbf{c}(j) \mathbf{J}(i,j), \quad (12)$$

$$\delta \ln \mathbf{c}^{k+1,p} = \mathbf{J}_{\ln}^{-1} \mathbf{f}(\mathbf{c}^{k+1,p}), \quad (13)$$

and

240 $\mathbf{c}^{k+1,p+1} = \mathbf{c}^{k+1,p} \exp[-\delta \ln(\mathbf{c}^{k+1,p})]. \quad (14)$

One downside is apparent from Eq. (12): suppose $\mathbf{c}^{k+1,p}(j)$ is 10^{-10} ; $\mathbf{J}(i,j)$ (column j of the Jacobian) decreases by 10^{10} times to $\mathbf{J}_{ln}(i,j)$. As a result, the condition number of the Jacobian matrix may increase by orders of magnitude, which may constrain the time step size to be substantially reduced, compared to the case without LT. Secondly, the small eigenvalues may result in large 245 $\delta \ln \mathbf{c}^{k+1,p}$, which may cause overshooting to an unrealistically large concentration (negative update), or essentially 0 concentration (positive update), or even overflow in the exponential function in Eq. (14). To prevent these problems, PFLOTRAN limits the update to

$$\delta \ln(\mathbf{c}^{k+1,p}) = \text{sign}[\delta \ln(\mathbf{c}^{k+1,p})] \min[\text{abs}(\delta \ln \mathbf{c}^{k+1,p}), \delta_{ln,max}] \quad (15)$$

with a default $\delta_{ln,max} = 5$. Thirdly, LT does not prevent zero concentration per se: a number of iterations with positive $\delta \ln \mathbf{c}^{k+1,p}(j)$ update can bring $\mathbf{c}^{k+1,p+1}(j)$ to zero. With zero concentration, \mathbf{J}_{ln} becomes singular, and the numerical solution fails. These can be shown by solving $dc/dt = -1$ in the log transformed form $d \ln c/dt = -1/c$: $c^{k+1} = c^k e^{-\Delta t/c^{k+1}}$. LT converts the linear problem that does not require iteration to a nonlinear problem that needs to be solved iteratively. As c^{k+1} becomes small, Δt has to be decreased accordingly, or the concentration goes to zero and causes 255 division by 0 overflow.

3.3 Downregulation of reaction rate

Even though ensuring nonnegativity, neither SU nor LT can prevent zero concentrations, which can cause numerical problems for both. They require the solution to the mathematical representation itself be positive. To obtain such a representation, it is necessary to downregulate reaction rates to 260 represent the limitation of the availability of each reactant on the reaction rate. In the geochemical modeling literature, this is mostly represented by a rate limiting function as a function of concentration in each reaction for each reactant. CLM downregulates demand (consumption) as a function of rate (demand-based competition) (Tang and Riley, 2015). We further examine downregulation of consumption rate as a function of concentration (DC) and rate (DR) for use with SU or LT to address 265 the nonphysical solution negativity.

3.3.1 Downregulation of consumption rate as a function of concentration

The first-order decay problem (Eqs. A1, A4, A7, A9) is nonnegative. This can be shown by solving the first-order problem $dc/dt = -c$ using either the analytical solution $c^{k+1} = c^k e^{-\Delta t}$ or the backward Euler method $c^{k+1} = c^k / (1 + \Delta t)$; However, it can go to zero. To avoid zero concentration, 270 a residual term is often used to limit the decay to a residual concentration (Tang et al., 2013a). For

example, Eq. (A1) becomes

$$\frac{d[\text{CN}_u]}{dt} = -k_d f_T f_w ([\text{CN}_u] - [\text{CN}_u]_r). \quad (16)$$

When the concentration goes below $[\text{CN}_u]_r$ in an iteration, Eq. (16) implies a hypothetical reverse reaction to bring it back to $[\text{CN}_u]_r$. An alternative is to use a cutoff (Appendix B). Even though it can be smoothed, a cutoff is still a sharp change as shown in the derivative (Fig. 14). In contrast, the residual concentration introduced in Eq. (16) does not change the derivative.

For the litter decomposition reactions (R7, R8, R9) that immobilize nitrogen, and nitrification reaction (R11) associated with decomposition to produce N_2O , the rate formulae (e.g., Eq. A1, 16,) do not account for the limitation of the reaction rate by the availability of nitrogen. They are zero order with respect to nitrogen. Mechanistically, a nitrogen-limiting function needs to be added to those formulae to represent the limitation of decreasing nitrogen concentration on decomposition rate (downregulation), for example,

$$\frac{d[\text{CN}_u]}{dt} = -k_d f_T f_w ([\text{CN}_u] - [\text{CN}_u]_r) f([N]). \quad (17)$$

A widely used downregulation function is the Monod substrate limitation function (Hammond, 2003; Tang et al., 2013a):

$$f([N]) = \frac{[N]}{[N] + k_m}, \quad (18)$$

with half saturation k_m . In the case of $[N] = k_m$, $f([N]) = 0.5$. For $[N] \gg k_m$, Eq. (18) is zero order with respect to $[N]$ or has little impact on the decomposition reactions that immobilize N. For $[N] \ll k_m$, Eq. (18) approximates first order with respect to $[N]$ (Figure 15d). Nevertheless, since $\frac{df([N])}{d[N]} = \frac{k_N}{([N]+k_m)^2}$, the derivative increases to about k_m^{-1} as the concentration decreases to below k_m (Fig. 15). This is similar to the smoothed cutoff in Eq. (B2) (Fig. 14): even though smoothed, it is still a steep transition.

To represent the threshold concentration beyond which certain microorganisms can not use certain electron donors (Fennell and Gossett, 1998), and prevent zero concentration, a residual concentration can be added to Eq. (18) as $f([N]) = \frac{[N]-[N]_r}{[N]-[N]_r+k_m}$. Similar to Eq. (16), this implies a nonphysical backward reaction when $[N] < [N]_r$. If we use the cutoff instead of a residual concentration, a combination of Eqs. (18) and (B2) produces more nonlinearity, as evidenced in Fig. (16), which may result in small time step sizes to march through these transition regions when the concentration gets there.

The plant nitrogen uptake reactions (R13, R14) are zero order, and substrate limiting functions need to be added as well. For plant uptake and immobilization, we have to distribute the demands between NH_4^+ and NO_3^- . If we simulate the NH_4^+ limitation on plant uptake, e.g., with

$$R_a = R_p \frac{[\text{NH}_4^+]}{[\text{NH}_4^+] + k_m}, \quad (19)$$

the NO_3^- plant uptake can be represented by

$$R_n = (R_p - R_a) \frac{[\text{NO}_3^-]}{[\text{NO}_3^-] + k_m} = R_p \frac{k_m}{[\text{NH}_4^+] + k_m} \frac{[\text{NO}_3^-]}{[\text{NO}_3^-] + k_m}. \quad (20)$$

305 This essentially assumes an inhibition of NH_4^+ on NO_3^- uptake, which is consistent with the observation that plant NO_3^- uptake rate remained low until NH_4^+ concentrations dropped below a threshold (Eltrop and Marschner, 1996).

For comparison with CLM, we examine the uptake rate as a function of demands and available concentrations $f_{pi} = R_a + R_n/R_p$. As an example, we consider uptake $R_p = 10^{-9} \text{ M s}^{-1}$ from a solution with various $[\text{NH}_4^+]$ and $[\text{NO}_3^-]$ for a 0.5 h time step. With CLM, $f_{pi} = 1$ when $[\text{NH}_4^+] + [\text{NO}_3^-] \geq R_p\Delta t$; otherwise, it decreases with decreasing $[\text{NH}_4^+] + [\text{NO}_3^-]$ (Fig. 2). The new representation (Eqs. 19, 20) is generally similar, with $f_{pi} = 1$ or 0 when $[\text{NH}_4^+]$ or $[\text{NO}_3^-] \gg k_m$ or $\ll k_m$. For the intermediate concentrations, f_{pi} in the new scheme is less than or equal to that in CLM because NH_4^+ “inhibits” NO_3^- uptake. The difference decreases with decreasing k_m , apparently disappearing at $k_m = 10^{-10}$ (Fig. 3).

310 Various level of preferences of NH_4^+ over NO_3^- uptake were observed for plants (Pfautsch et al., 2009; Warren and Adams, 2007; Nordin et al., 2001; Falkengren-Grerup, 1995; Gherardi et al., 2013). The microbial uptake of inorganic and organic nitrogen species is similar (Fouilland et al., 2007; Kirchman, 1994; Kirchman and Wheeler, 1998; Middelburg and Nieuwenhuize, 2000; Veugelers et al., 2004). CLM implies a strong preference for NH_4^+ over NO_3^- . For example, if NH_4^+ is abundant, NO_3^- will not be taken. The new scheme allows the level of preference to be adjusted by varying k_m .

3.3.2 Downregulation of consumption as a function of rate

DC involves adding residual concentration and half saturations, which has the potential to provide 325 mechanistic treatment of the processes. It also introduces a number of parameters that need to be determined. These parameters may not be well defined and can vary among different microbe and plant species under various conditions across the globe. An alternative is to downregulate consumption as a function of rate (DR) like CLM demand-based competition.

CLM splits the available nitrogen to meet the demands by microbes and plants in proportion 330 to the potential rates (Thornton and Rosenbloom, 2005) (Appendix A5). Suppose that the demand (consumption, sink, including immobilization, plant uptake, nitrification, etc.) rate is R_{dp} ; the supply (source, production, including deposition, mineralization, etc.) rate is R_s , and the concentration is $[N]$ at the beginning of the time step, the demand is downregulated to

$$R_d = \min [R_{dp}, R_s + [N]/\Delta t]. \quad (21)$$

335 Tang and Riley (2015) suggest to include R_s . For a reaction that is limited by multiple substrates (e.g., nitrogen and phosphorus), they use the minimum of the downregulated rates. Like the demand based competition in CLM, this is simple to implement, and prevents negative concentration when the explicit forward difference method is used. It is different for the backward Euler method. Applying to $dc/dt = -1$, $c^{k+1} = c^k - \Delta t$ for $c^k \geq \Delta t$; for $c^k < \Delta t$, $c^{k+1} = c^k/2$: the concentration

340 decreases by half in a time step. This is similar to SU with $\alpha = 0.5$, which is suggested by Bethke (2007). DR essentially downregulates the demands using the first-order rate when nitrogen is limiting. Eq. (21) is similar to Eq. (18), except that Eq. (18) switches from zero to the first-order rate smoothly, while Eq. (21) has a discontinuity.

Implementation of DR in a geochemical code like PFLOTRAN involves separating the supply
345 and consumption rates for each species in each reaction, and checking and conducting downregulation when necessary after contributions from all of the reactions are accumulated. It involves not only the rate terms for the residual but also the derivative terms for the Jacobian. The complexity explodes when more species needs to be downregulated (e.g., NH_4^+ , NO_3^- , and organic N) and there are transformation processes among these species. This is shown in Appendix C, which describes
350 the implementation of downregulation of NH_4^+ and NO_3^- with a nitrification reaction from NH_4^+ to NO_3^- . Basically, it becomes challenging to separate, track, and downregulate consumption and production rates for an indefinite number of species, and calculation of the Jacobian becomes convoluted. This work differs from Tang and Riley (2015) in that we use implicit time stepping while they use explicit time stepping.

355 4 Test problems, results, and discussions

We examine the causes of negative concentration; the advantages and disadvantages of SU, LT, DC,
360 and DR; and the accuracy, stability, and efficiency issue using both simple test problems and coupled CLM-PFLOTRAN site simulations. For simple test problems, we assess the conditions under which large updates occur. With coupled CLM-PFLOTRAN spin-up simulations for arctic, temperate, and tropical sites, we assess the conditions when the nonphysical solution negativity arises by relaxing the downregulation on nitrogen consumption (decreasing half saturation). Spreadsheet and PFLOTRAN input files are provided as supplemental information (SI).

Our implementation of CLM biogeochemistry introduces mainly two parameters: half saturation
365 k_m and residual concentration. A wide range of k_m values were reported for NH_4^+ , NO_3^- , and organic nitrogen for microbes and plants. The median, mean, and standard deviations range from $10 \sim 100$, $50 \sim 500$, and $10 \sim 200 \mu\text{M}$, respectively (Kuzyakov and Xu, 2013). Reported residual (threshold)
370 concentrations are limited and are considered to be 0 (e.g., HØGh-Jensen et al., 1997), likely because of the detection limits of the analytical methods. The detection limits are usually at the μM level, while up to the nM level was reported (Nollet and De Gelder, 2013). In Ecosys, the k_m is 0.40 and 0.35 gN m^{-3} , and the residual concentration is 0.0125 and 0.03 gN m^{-3} (Grant, 2013) for NH_4^+ and NO_3^- for microbes. We start with $k_m = 10^{-6} \text{ M}$ or mol m^{-3} , and residual concentration 10^{-15} M or mol m^{-3} for plants and microbes. To further investigate the nonphysical solution negativity for the current study and for future application for other nutrients (e.g., H_2 and O_2) where the concentrations can be much lower, we examine k_m from 10^{-3} to 10^{-12} in our test problems. The k_m is expected

375 to be different for different plants, microbes, and for NH_4^+ and NO_3^- , and different values can be assigned in the input file in our implementation. we do not differentiate them in this work as we focus on numerical issues.

4.1 Plant nitrogen uptake, nitrification, and denitrification

It was observed that plants can decrease nitrogen concentration to below detection limits in hours
 380 (Kamer et al., 2001). Plant nitrogen uptake is one of the major sinks for nitrogen in TEMs and contributes to decreasing nitrogen concentration numerically to 0 or negative. In CLM, the total plant nitrogen demand is calculated based on photosynthesized carbon allocated for new growth and the C:N stoichiometry for new growth allocation, and the plant nitrogen demand from the soil is equal to the total nitrogen demand minus retranslocated nitrogen stored in the plants (Oleson
 385 et al., 2013). The CLM calculated rate is provided as an input to PFLOTTRAN, which is constant in a 0.5 h time step. Without downregulating plant nitrogen uptake rate, nitrogen concentration is likely to go negative when the net consumption rate overwhelms a low available concentration. As the Monod substrate limiting function is the most widely used downregulation function, we examine the numerical solutions, beginning with the numerical issues during the iteration processes.
 390 Incrementally, we add first order reactions (e.g., nitrification, denitrification, and plant NO_3^- uptake) to look into these issues in increasingly complex systems.

4.1.1 Plant NH_4^+ uptake (Test 1)

We consider the plant NH_4^+ uptake reaction (R13) with the rate R_a

$$\frac{d[\text{NH}_4^+]}{dt} = -R_a \frac{[\text{NH}_4^+]}{[\text{NH}_4^+] + k_m} = -R_{at}. \quad (22)$$

395 A semi-analytical solution (Eq. D2) has one positive and one negative roots. With the positive root, $[\text{NH}_4^+]^{k+1} \rightarrow 0$ when $[\text{NH}_4^+]^k \rightarrow 0$, or $R_a \Delta t \rightarrow \infty$. As $k_m \rightarrow 0$, $[\text{NH}_4^+]^{k+1} \rightarrow 0$ when $[\text{NH}_4^+]^k \leq R_a \Delta t$. If we replace $[\text{NH}_4^+]$ with $[\text{NH}_4^+] - [\text{NH}_4^+]_r$ in Eqs. (22) and (D2), $[\text{NH}_4^+]^{k+1} \rightarrow [\text{NH}_4^+]_r$. The representation of plant NH_4^+ uptake by Eq. (22) ensures $[\text{NH}_4^+]^{k+1} \geq [\text{NH}_4^+]_r$.

Solving Eq. (22) using the backward Euler method and Newton-Raphson method from time step
 400 k to $k + 1$, and $[\text{NH}_4^+]^{k+1,0} = [\text{NH}_4^+]^k$, the update for the first iteration is

$$\xi = \frac{\delta[\text{NH}_4^+]^{k+1,1}}{[\text{NH}_4^+]^{k+1,0}} = \frac{\frac{1}{[\text{NH}_4^+]^k + k_m}}{\frac{1}{R_a \Delta t} + \frac{k_m}{([\text{NH}_4^+]^k + k_m)^2}}. \quad (23)$$

The update is a function of uptake rate (R_a), time step size (Δt), concentration ($[\text{NH}_4^+]^k$), and half saturation (k_m). For $R_a \Delta t$ from 0 to ∞ , ξ increases from 0 to $1 + [\text{NH}_4^+]/k_m$. With $[\text{NH}_4^+]^k = 10^{-6}$ M, and $k_m = 10^{-6}$ M, the update is less than $[\text{NH}_4^+]^k$ when $R_a \Delta t \leq 5 \times 10^{-6}$ M (say using
 405 CLM $\Delta t = 1800$ s, $R_a \leq 2.8 \times 10^{-9}$ mol s $^{-1}$, Fig. 4); when $R_a \Delta t$ increases, the update increases, with an upper limit of 2, suggesting that large time step size and uptake rate alone usually do not

lead to an excessively large update or a small scaling factor. For $[\text{NH}_4^+]^k$ from 0 to ∞ , ξ starts from $R_a\Delta t/(k_m + R_a\Delta t)$, increases to peak at $\sqrt{R_a\Delta t/k_m}/2$ when $[\text{NH}_4^+] = \sqrt{k_m R_a \Delta t} - k_m$, and then decreases to 0. With $k_m = 10^{-6}$ M and $R_a\Delta t = 0.001$ M, ξ starts from ~ 1 , increases to peak at ~ 16 when $[\text{NH}_4^+]^k = 3 \times 10^{-5}$, and decreases to 0 as $[\text{NH}_4^+]^k \rightarrow \infty$. For k_m from ∞ to 0, ξ increases from 0 to $R_a\Delta t/k_m$. With $R_a\Delta t = 10^{-3}$, and $[\text{NH}_4^+]^k = 10^{-6}$, the update increases several orders of magnitude as k_m decreases from 10^{-6} to 10^{-9} , with a limit of $\xi = 10^3$ (Fig.4). While the semi-analytical solution (Eq. D2) indicates that the problem itself is nonnegative, these calculations suggests that overshoot can occur, and large time step, uptake rate, low concentration and half saturation can contribute to large update that can exceed the available concentration by orders of magnitude.

Now we look into the iteration processes: the solution starts with an overshoot to 1.995×10^{-6} and converges to the positive semi-analytical solution 3.1127×10^{-5} (Eq. D2) in eight iterations when $R_a\Delta t = 0.001$ (SI spreadsheet). dR/dC changes from ~ 1 to 10^5 in the first iteration, although the change in the Jacobian is reduced due to a small $R_a\Delta t$. With $R_a\Delta t = 0.002$, the update δ is greater than $[\text{NH}_4^+]_k$; without scaling back the update, the solution converges to the negative solution -1.002×10^{-3} (SI, negative root from Eq. D2). Scaling back the update with $\alpha = 0.9999$, the concentration decreases by $1 - \alpha = 10^{-4}$ times in the first iteration rather than to negative root (SI). The solution converges to the positive root in seven iterations.

If LT is used with $R_a\Delta t = 0.002$, the solution oscillates and results in a δ of -1041.5 in the fifth iteration, which essentially makes the solution 0 (SI). Limiting $\delta \leq \delta_{\ln,\max}$, the solution converges to the semi-analytical solution (Eq.D2) in 7, 7, 6, 5, 6, and 8 iterations with $\delta_{\ln,\max} = 2, 3, 4, 5, 6$, and 7, respectively. For $\delta_{\ln,\max} \geq 8$, the iteration oscillates and converges slowly. If we split $R_a\Delta t = 0.002$ into two steps, the first step takes seven iterations to converge, while the second step takes nine iterations, with $\delta_{\ln,\max} = 5$.

Finally, we consider a PFLOTRAN simulation: a $1 \text{ m} \times 1 \text{ m} \times 1 \text{ m}$ grid cell, with 0.25 water content, initial $4 \mu\text{M}$ NH_4^+ , and a plant NH_4^+ demand $10^{-7} \text{ mol s}^{-1}$ ($\sim 3 \text{ mg d}^{-1}$, reported values for evergreen and deciduous forest range from $0.3 \sim 10 \text{ g m}^{-2} \text{ year}$, Chapin et al., 2011). It takes 10,000 s (2.78 h) to consume NH_4^+ . We use a time step of 0.5 h for a simulation duration of 10 h. Without downregulating the uptake rate, SU decreases $[\text{NH}_4^+]$ to 0 and then keeps it 0 as $\lambda = 0$. With LT, we obtain an accurate solution until it fails when the concentration goes to 0, and the Jacobian becomes singular.

Using SU with DC with a residual concentration $[\text{NH}_4^+]_r$ of 10^{-20} M and a half saturation k_m of 10^{-6} , 10^{-9} , or 10^{-12} , the calculated concentration $[\text{NH}_4^+]$ is kept to be greater than or equal to 10^{-20} (Fig. 5). With a k_m of 10^{-6} , the calculated $[\text{NH}_4^+]$ is above 10^{-20} , and the update does not exceed the concentration in any iteration during the 10 h simulation duration. With a k_m of 10^{-9} and 10^{-12} , the update exceeds the concentration from time 2.5 h to 3.0 h. In the case of $k_m = 10^{-12}$ M for the next time step, the two iterations decrease the concentration by 10,000 times, and the solution is

deemed converged as the default STOL = 10^{-8} is met. This continues for another time step with one iteration (Fig. 5). In these two time steps, the solution does not converge to the exact solution. If we set the residual concentration to 10^{-15} , the false convergence results in a concentration below 10^{-15} (8.71×10^{-16}) in the first of the two time steps. If we set STOL = 10^{-50} to avoid the convergence due to Eq. (9), PFLOTTRAN cuts the time step sizes and quits due to too many cuts.

Using LT can avoid the false convergences in this example, but it requires more iterations, with 56/95, 36/110, and 32/87 for SU/LT for k_m values of 10^{-6} , 10^{-9} , and 10^{-12} M, respectively. These results demonstrate that SU ensures nonnegativity and allows large time step for efficiency at the risk of inaccuracy. LT is accurate but requires more iterations.

4.1.2 Plant NH_4^+ uptake and nitrification (Test 2)

Adding a nitrification reaction (R10) with a first-order rate, Eq. (22) becomes

$$455 \quad \frac{d[\text{NH}_4^+]}{dt} = -R_a \frac{[\text{NH}_4^+]}{[\text{NH}_4^+] + k_m} - k_{nitr} [\text{NH}_4^+] = -R_{at} - R_{nitr}. \quad (24)$$

With $J_{at} = \frac{dR_{at}}{d[\text{NH}_4^+]} = R_a \frac{k_m}{([\text{NH}_4^+] + k_m)^2}$, and $J_{nitr} = \frac{dR_{nitr}}{d[\text{NH}_4^+]} = k_{nitr}$ and using the backward Euler method and Newton-Raphson method, Eq. (5) becomes Eq. (E1). For plant ammonium uptake,

$$\delta[\text{PlantA}]^{k+1,1} = J_{at} \delta[\text{NH}_4^+]^{k+1,1} - R_{at} = -\frac{\frac{1}{\Delta t} + J_{nitr}}{\frac{1}{\Delta t} + J_{at} + J_{nitr}} R_{at} + \frac{J_{at}}{\frac{1}{\Delta t} + J_{at} + J_{nitr}} R_{nitr}, \quad (25)$$

A positive component, $\frac{J_{at}}{1/\Delta t + J_{at} + J_{nitr}} R_{nitr}$, is introduced even though there is no reaction that consumes PlantA. Depending on the relative magnitude of the rates (R_{at} and R_{nitr}) and the derivatives (J_{at} and J_{nitr}), as well as time step size Δt and $[\text{NH}_4^+]^k$, the update can be positive, which can result in too small a scaling factor that causes a nonphysical solution negativity if [PlantA] is very small. This demonstrates that products as well as reactants can be driven negative during the Newton-Raphson iterations.

465 This can also be shown with a simple first-order reaction A \rightarrow B with rate $k[A]$: taking out the equation for B, $-k\delta[A] + \delta[B]/\Delta t = -k[A]$, $\delta[B] = k\Delta t(\delta[A] - [A])$. Suppose $\delta[A] = 2 \times 10^{-6}$, $[A] = 10^{-6}$, and $k\Delta t = 10^{-3}$, then $\delta[B] = 10^{-9}$. If [B] is very small, say 10^{-20} , this results in a $\lambda = 0.99 \times 10^{-11}$. This has an implication for DR: downregulating only some reactants may not be sufficient to guarantee positivity, while downregulating both reactants and products is complicated.

470 4.1.3 Plant uptake, nitrification, and denitrification (Test 3)

Adding plant NO_3^- uptake reaction (R14) with rate $R_{nt} = R_p \frac{[\text{NH}_4^+]}{[\text{NH}_4^+] + k_m} \frac{[\text{NO}_3^-]}{[\text{NO}_3^-] + k_m}$, $J_{nt,n} = \frac{dR_{nt}}{d[\text{NO}_3^-]} = R_p \frac{[\text{NH}_4^+]}{[\text{NH}_4^+] + k_m} \frac{k_m}{([\text{NO}_3^-] + k_m)^2}$, and $J_{nt,a} = \frac{dR_{nt}}{d[\text{NH}_4^+]} = \frac{dR_n}{d[\text{NH}_4^+]} \frac{k_m}{([\text{NH}_4^+] + k_m)^2} \frac{[\text{NO}_3^-]}{[\text{NO}_3^-] + k_m}$, and denitrification reaction (R12) with rate $R_{deni} = k_{deni} [\text{NO}_3^-]$, and $J_{deni} = \frac{dR_{deni}}{d[\text{NO}_3^-]} = k_{deni}$, Eq. (5) becomes Eq. (F1). For NO_3^- ,

$$\delta[\text{NO}_3^-]^{k+1,1} \left(\frac{1}{\Delta t} + J_{nt} + J_{deni} \right) = (J_{nitr} - J_{nt,a})(R_{at} + R_{nitr}) \delta[\text{NH}_4^+]^{k+1,1} - R_{nitr} + R_{nt} + R_{deni}.$$

In addition to plant NH_4^+ uptake and nitrification, $\delta[\text{NO}_3^-]^{k+1,1}$ becomes a function of plant NO_3^- uptake and denitrification, and all of the rate coefficients and derivatives may contribute to a positive update that can result in too small a scaling factor. Even though the problem itself is nonnegative, coupling these reactions together in the Newton-Raphson iteration does introduce the potential to produce relatively significant positive update for NO_3^- and PlantN. It worsens when the interdependent uptake of NH_4^- and NO_3^- (Eqs. 19 and 20) representation brings in the impact of any rates related to NH_4^+ (including deposition, uptake, immobilization, etc.), with potentially large derivative terms for the off-diagonal entries in the Jacobian matrix. It is the nonlinearity introduced in the downregulation and propagated through the reactions that causes the challenge. A smaller half saturation introduces a greater derivative change; therefore, it is more likely to cause the problem. These results also suggest that the likelihood for NO_3^- to go negative is greater than NH_4^+ , and PlantN is greater than PlantA as the latter are influenced by more rates and derivatives.

4.2 N immobilization, mineralization, and nitrification during decomposition (Test 4)

We further examine the implications of a small half saturation on the numerical solutions with a decomposition test problem. We consider a case of decomposing 0.2 M Lit1C + 0.005 M Lit1N to produce SOM1 with an initial 4 μM NH_4^+ using the reactions (R7 and R2) in the CLM-CN reaction network (Fig. 1) at first. Then we add the nitrification reaction (R11) with rate (Eq. A5) to examine the implications of a complex rate formula. We use PFLOTRAN with a fully saturated grid cell of 1 m and porosity of 0.25.

Lit1 decomposes fast at the beginning and then slows down as NH_4^+ is depleted (Fig. 6). The Lit1 decomposition rate is controlled by the mineralization rate from SOM1 decomposition. As the immobilization rate decreases with decreasing Lit1, $[\text{NH}_4^+]$ rebounds. For k_m of 10^{-6} , 10^{-9} , and 10^{-12} M, Lit1 and SOM1 dynamics are similar, but the $[\text{NH}_4^+]$ values are decreased to $\sim 10^{-8}$, 10^{-11} , and 10^{-14} M, respectively. The number of iterations for SU/LT is 76/169, 68/194, and 54/207 for the three k_m values. Obviously, smaller k_m results in lower $[\text{NH}_4^+]$ and more iterations.

The implication of $k_m = 10^{-12}$ becomes obvious when the nitrification reaction (R11) with rate (Eq. A5 with downregulation) is included. In the time step when the mineralization rate surpasses the immobilization, $\delta[\text{N}_2\text{O}] > [\text{N}_2\text{O}]$, and N_2O is lowered to 10^{-23} M in three iterations as STOL is met. In the next time step, $\delta[\text{N}_2\text{O}] = \sim 10^{-13}$, resulting in a scaling factor λ of 10^{-10} that stops the simulation. Checking the calculations in the last iteration, the lower $[\text{NH}_4^+]$ (2×10^{-13} M) introduces a derivative (6.94×10^{11}) into the column of the Jacobian with respect to $[\text{NH}_4^+]$. Because N_2O production is a function of both immobilization and mineralization, a Monod substrate limiting function is added to downregulate the mineralization component (Eq. A5). In the row corresponding to $[\text{N}_2\text{O}]$, the off-diagonal terms are nonzero in the entries with respect to $[\text{NH}_4^+]$, [Lit1C], [Lit1N], and [SOM1]. Even though these values are relatively small because the fraction of net mineralization

rate to produce $[N_2O]$ is small (0.02), inversion of the Jacobian matrix results in substantial values in the row, and the 10^{-13} M update comes mainly from the entries with respect to Lit1C and Lit1N. This further illustrates that extremely small half saturation together with complex rate limiting formula can result in large derivatives at small concentration, which may lead to overshoot that results in a
515 small update scaling factor.

4.3 CLM-PFLOTRAN simulations

We examine the nonphysical solution negativity in the coupled CLM-PFLOTRAN simulations for arctic (US-Brw), temperate (US-WBW), and tropical (BR-Cax) AmeriFlux sites. The CLM-PFLOTRAN simulations are run in the mode that PFLOTRAN only handles subsurface chemistry (decomposition,
520 nitrification, denitrification, N plant uptake). For comparison, 1) depth and O_2 availability impact on decomposition, 2) cryoturbation, 3) SOM transport, and 4) nitrogen leaching are ignored by setting 1) decomp_depth_efolding to 10^6 m, o_scalar to 1, 2) cryoturb_diffusion, 3) som_diffusion, and 4) sf_no3 and sf_sminn to 0. In the base case, $k_m = 10^{-6}$ and residual concentration 10^{-15} . Nitrogen is the focus as it is the rate limiting nutrient in CLM4.5 biogeochemistry. Spin-up simulations are
525 used because they are generally more likely to face a nonphysical solution negativity because the simulations start far away from equilibrium. In these site simulations, PFLOTRAN uses the same 10 layer grid for the 3.8 m one-dimensional column as CLM. The simulation duration is 500, 300, and 300 year for the arctic, temperate, and tropical sites as the system roughly reaches steady state within these simulation durations.

530 **4.3.1 Site descriptions**

The US-Brw site (71.35N, 156.62W) is located near Barrow, Alaska. The mean annual temperature, precipitation, and snowfall are $-12^\circ C$, 11 cm, and 69 cm, respectively (1971 ~ 2000) (Lara et al., 2012). The landscape is poorly drained polygonized tundra. The maximum thaw depth ranges from 30 to 40 cm, and the snow free-period is variable in length but generally begins in early June and
535 lasts until early September (Hinkel and Nelson, 2003). The area is composed of several different representative wet-moist coastal sedge tundra types, including wet sedges, grasses, moss, and assorted lichens. The leaf area index (LAI) is ~ 1.1 .

The US-WBW site (35.96N, 84.29W) is located in the Walker Branch Watershed in Oak Ridge, Tennessee (Hanson and Wullschleger, 2003). The climate is typical of the humid southern Appalachian region. The mean annual precipitation is ~ 139 cm, and the mean median temperature is $14.5^\circ C$. The soil is primarily Ultisols that developed in humid climates in the temperate zone on old or highly weathered material under forest. The temperate deciduous broadleaf forest was regenerated from agriculture land 50 years ago. LAI is ~ 6.2 (Hanson et al., 2004).

The BR-Cax site (-1.72N, -51.46W) is located in the eastern Amazon tropical rainforest. The
545 mean annual rainfall is between 2000 and 2500 mm, with a pronounced dry season between June

and November. The soil is a yellow oxisol (Brazilian classification latossolo amarelo) with a thick stony laterite layer at 3 ~ 4 m depth (da Costa et al., 2010). The vegetation is evergreen broadleaf forest. The LAI is 4 ~ 6 (Powell et al., 2013).

4.3.2 Base case simulation results

- 550 For base case simulations, we use DC and LT because DC is the general geochemical modeling approach and is rigorous, and LT is less likely to cause numerical error than SU. The site climate data from 1998 to 2006, 2002 to 2010, and 2001 to 2006 are used to drive the spin-up simulation for the arctic, temperate, and tropical sites, respectively. This introduces a multi-year cycle in addition to the annual cycle (Figs. 7, 8, 9). Overall, CLM-PFLOTTRAN is close to CLM4.5 in predicting LAI
555 and nitrogen distribution among vegetation, litter, SOM, NH_4^+ and NO_3^- pools for the arctic (Fig. 7), temperate (Fig. 8), and tropical site (Fig. 9). CLM4.5 does reach equilibrium earlier than CLM-PFLOTTRAN as the latter represents slightly more inhibition on plant nitrogen uptake, depending on the half saturation. The impact is most obvious for LitN, NH_4^+ , and NO_3^- in the tropical site (Fig. 9). Because the nitrogen demand competition scheme implemented in CLM-PFLOTTRAN is
560 different from that in CLM4.5, we do not expect identical results as Fang et al. (2013).

The arctic site shows a distinct summer growing season (Fig. 7): LAI and VEGN jump up at the beginning, then level off, and drop down at the end of the growing season when LITN jumps up due to litter fall. NH_4^+ and NO_3^- drop to very low level at the beginning of growing season and accumulate at the other times. In addition to a longer growing season than the arctic site, the temperate site
565 shows more litter fall by the end of the growing season as it is a temperate deciduous forest, which introduces immobilization demand that further lowers NH_4^+ and NO_3^- concentrations (Fig. 8e inset). The seasonality is much less apparent in the tropical site than in the arctic and temperate sites. LAI, VEGN, LITN, and SOMN accumulate with less seasonal variations to reach an equilibrium. It takes much longer simulation duration to spin up the simulation for the arctic site than the temperate and
570 tropical sites because the arctic site receives less solar radiation.

Except for the tropical site where the higher k_m of 10^{-3} mol m $^{-3}$ results in lower immobilization, higher accumulation of LITN, and higher $[\text{NH}_4^+]$ and $[\text{NO}_3^-]$ during the spinup (Fig. 9), the range of k_m values (10^{-3} , 10^{-6} , and 10^{-9} mol m $^{-3}$) generally has limited impact on the overall calculations except that the nitrogen concentrations drop lower with lower k_m values (e.g., inset in Figs. 7e,f, 8e),
575 which is similar to Fig. (6). The lack of sensitivity is because these very low concentrations do not make up a mass of nitrogen that is significant enough to influence the carbon and nitrogen cycle. However, as a small k_m means weak downregulation and steep transition between zero order and first order, it has implications on accuracy, stability, and efficiency of the numerical solutions.

4.3.3 Efficiency

580 We use OIC (ORNL Institutional Cluster phase5 queue esd13q) for the comparison of computing time for the three sites using DC or DR with SU or LT. The results suggest that SU takes about 30~80% more time than CLM, and LT costs about two to four times that of CLM (Table 1). For LT with DC, the computing time increases with decreasing k_m . This is expected because a smaller k_m may require smaller time step size to march through steeper transition. The comparison between DC
585 and DR is mixed. This is because DC has more complicated substrate limiting representation that requires smaller time step sizes to march through steep transitions, while DR has simpler substrate limiting representation, but a discontinuity that may require more iterations. We expect that CLM-PFLOTTRAN will require more computational cost as more complex biogeochemical processes are represented, transport is added, and thermal hydrology (including freeze and thaw) are incorporated.
590 Future code optimization for performance on next-generation supercomputers is expected to mitigate the computational cost issue.

4.3.4 Accuracy

CLM checks carbon and nitrogen mass balance for every time step, and report $\geq 10^{-8} \text{ gm}^{-2}$ errors. Among our CLM-PFLOTTRAN simulations, no mass balance error is recorded when LT is used. In
595 contrast, when SU is used, mass balance errors are recorded for all of the simulations. We use the tropical site to demonstrate the accuracy issue by SU. With $k_m = 10^{-3} \text{ mol m}^{-3}$, the LAI, VEGN, LITN, SOMN, NH_4^+ , and NO_3^- are similar to that calculated using LT (Fig. 9 vs. 10). With decreasing k_m , the results diverge by SU but not by LT. With k_m of $10^{-6} \text{ mol m}^{-3}$, much less nitrogen is predicted to accumulate with SU than with LT. Obviously, numerical error is introduced, and the
600 error increases with decreasing k_m

Looking into NH_4^+ and NO_3^- concentration in the first year, SU calculation is similar to LT when NH_4^+ and NO_3^- are abundant at early times (Fig. 11). As nitrogen becomes limiting at late times, NH_4^+ and NO_3^- accumulate lower by SU than by LT. Further looking into the daily cycles (Fig. 11 inset), plant uptake decreases NH_4^+ and NO_3^- concentrations to very low level in the morning when
605 photosynthesis starts; then the low concentration limits plant uptake, and nitrogen accumulates and rebounds until the next day. From day 226 to 230, SU calculations are only slightly lower than LT calculations during the rebound. From day 250 to 254, the SU calculated rebounds are about two orders of magnitude lower than LT calculations. From day 295 to 300, it worsens both in frequency and magnitude. Overall, NH_4^+ and NO_3^- accumulation appear to be numerically “inhibited” when the
610 system becomes increasingly nitrogen limiting and when SU is used. As a result, the mass balance error shots up during the nitrogen limiting periods (Fig. 12).

Further checking nitrogen dynamics between day 250 and 254, these “inhibited” intervals in Fig. (11) coincide with the concentration valleys in the diagnostic species N_2Od , which is used to track

N₂O production (reaction R11) due to decomposition (Eq. A5, Fig. 13). At the beginning of each
615 CLM 0.5 h time step, we reset [N₂Od] to 10⁻¹⁰ in PFLOTRAN so that it can be used to calculate
the reaction rate for CLM (instead of saving concentration in previous time step) As we do not have
any consumption reaction for N₂Od, the decreases are purely numerical. The contribution of the
derivatives of the rate (Eq. A5) results in a large enough positive update to N₂Od, which results in
a small scaling factor, and the solution is considered converged due to less than STOL update. This
620 small scaling factor and false convergence result in the “inhibition” of NH₄⁺ and NO₃⁻ production,
and manifest as much less nitrogen accumulation in Fig. (10). Decreasing STOL and/or λ_{min} does
not resolve the issue (Fig. (11)), nor does removing these diagnostic species, and/or decreasing α
(say from 0.99 to 0.5). As long as there are low concentrations, e.g., NH₄⁺, NO₃⁻, it is possible to
have a sufficiently large update to result in a small scaling factor to "inhibit" production.

625 **4.3.5 Stability and robustness**

PFLOTRAN may stop CLM-PFLOTRAN run due to 1) too small a scaling factor for SU to avoid
excessive numerical error, 2) too small a time step size or too many time step cuts for either SU or
LT, and 3) singular Jacobian matrix for LT.

With SU and DC, the simulations run to conclusion for $k_m = 10^{-3}, 10^{-6}$, and $10^{-9} \text{ mol m}^{-3}$ but
630 abort for $k_m = 10^{-12}$ for all of the three sites. The reasons for the unfinished runs are too small a
scaling factor for NO₃⁻. As indicated in Test 3, it is not surprising to see positive updates to NO₃⁻ as
they depend on source and sink terms not only for NO₃⁻ but also for NH₄⁺ because of the nitrogen
demand distribution scheme (Eq. 20) for plant uptake and immobilization. With smaller k_m , the
derivative of the substrate limiting function becomes bigger when the concentration is lower, and
635 the off-diagonal entries in the Jacobian matrix become more dominant, increasing the probability
of producing positive update. When a greater than [NO₃⁻] update is produced, [NO₃⁻] is expected
to decrease by 1- α times (0.01 by default) for each subsequent iterations, making it lower than the
residual concentration (10^{-15}). If the λ is not greater than 10^{-10} , and the solution may be considered
converged because STOL is satisfied, [NO₃⁻] can be very low, much below residual concentration
640 (10^{-26}). With such low concentrations, a positive update can make the scaling factor to be less than
 10^{-10} . The arctic site simulation aborts at a much later time than the temperate and tropical site
because the uptake and immobilization rate is much smaller in the former site.

A lower STOL can be used to avoid or mitigate the false convergence. With STOL = 10^{-12} , the
simulations abort for $k_m = 10^{-3}, 10^{-6}$, and $10^{-9} \text{ mol m}^{-3}$ for the three sites due to too big an update
645 for the diagnostic variables (PlantN, DeniN, or N₂Od). Instead of aborting the run when the scaling
factor is less than the minimum allowable value λ_{min} , an alternative is to consider the iteration
diverges and return to the time stepping subroutine to cut time step size to increase accuracy. This
mitigates the stability and accuracy challenge at the expense of computing time. Our results show
the default $\lambda_{min} = 10^{-10}$ can result in excessive error even when STOL is extremely small. It is not

650 clear how large a λ_{min} value can insure accuracy without introducing unnecessary computing time.
For this reason, we caution against use of SU.

Similar to SU, the simulations using LT run to conclusion for $k_m = 10^{-3}, 10^{-6}$, and $10^{-9} \text{ mol m}^{-3}$ but abort for $k_m = 10^{-12}$ for the three sites. The runs abort at 21.6, 2.46, and 1.1 year for the arctic, temperate, and tropical site because the number of time step cuts for one time step exceeds
655 MAX_CUT = 16 (default). Increasing MAX_CUT to 50, the simulations run to conclusion with longer computing time (Table 1). Looking into the reason for time step cuts, it usually involves iteration with NO_3^- concentration oscillation between the first order and zero order (Fig. 15b, for example, between 1.69×10^{-11} and 1.14×10^{-13} in the sixth layer in 0.5 year for BR-Cax site). While requiring a large MAX_CUT, all of the simulations conclude with less than a 20% increase of
660 computing time for k_m of 10^{-12} over 10^{-9} . These results suggest that log transformation together with downregulation with reasonable half saturation and residual concentration can provide accurate, stable, and feasibly efficient solutions for CLM-PFLOTRAN biogeochemistry simulations.

5 Summary and Conclusions

We implement CLM biogeochemistry (as an example, CLM-CN decomposition, nitrification, denitrification, and plant nitrogen uptake reactions) in CLM-PFLOTRAN. As CLM uses explicit time stepping and demand-based competition, the concentration is always nonnegative. PFLOTRAN uses implicit time stepping and Newton-Raphson method; the concentration can become negative, which is not physical, may cause numerical instability, and introduce numerical error.

Both scaling back update in each iteration and log transformation can enforce nonnegativity. SU
670 allows large time step sizes to achieve efficiency, but may introduce excessive numerical error if the scaling factor is small and the iterations are considered converged due to small updates. Log transformation involves multiplication of the Jacobian matrix by the concentration vector, decreasing the condition number by orders of magnitude as the concentration can be low. As a result, LT often decreases the time step size and increases computational cost. Neither SU nor LT prevents too small
675 or 0 concentrations. When the concentration becomes too small and essentially 0, SU may stop all reactions by a small positivity update caused by consumption, numerical overshoot, or truncation error and LT can fail due to too stiff or singular Jacobian matrix. Both SU and LT require that the limitation of reactant availability on reaction rates to be represented such that solution to the mathematical representation does not introduce negative or 0 concentrations.

680 The first-order rate accounts for limitation of the reactant availability on the reaction rate, and the solution is nonnegative. Adding a residual concentration makes it positive. For the zero-order rate (the rate is not a function of a reactant), the Monod substrate limiting function provides a smooth transition from a zero order when the reactant is abundant to a first order when the reactant becomes

limiting. The downregulation relaxes with decreasing half saturation and residual concentration and
685 disappears when both are 0.

Our CLM-PFLOTRAN spin-up simulation at an arctic, temperate, and tropical site indicates that accurate and stable solution can be achieved with log transformation with two to three times the computing time of CLM4.5 for a range of half saturation values from 10^{-3} to 10^{-9} and a residual concentration of 10^{-15} for nitrogen. With half saturation of 10^{-12} , the number of maximum time
690 cuts has to be increased from the default 16 to accommodate for small time step sizes to resolve small concentration changes in the transition of zero and first-order rate. The computing time increase from the half saturation of 10^{-12} to 10^{-9} is less than 20%. As physical half saturation ranges from 10^{-5} to 10^{-6} M for nitrogen, and the detection limits are often above 10^{-9} M, log transformation together with downregulation by reasonable half saturation and residual concentration is
695 expected to provide an accurate, stable, and efficient solution for CLM-PFLOTRAN implementation for current CLM biogeochemistry. As more substrate limiting processes, such as labile C, P, O₂, and H₂, are implemented with lower half saturation and residual concentration, and more complicated rate relationship, the maximum allowable time step cut may need to be increased, and the computing time is expected to increase.

700 6 Code availability

PFLOTRAN is an open-source software. It is distributed under the terms of the GNU Lesser General Public License as published by the Free Software Foundation either version 2.1 of the License, or any later version. It is available at <https://bitbucket.org/pfotran>.

7 Author contribution

705 G. B., B. A., R. M., J. K., and F. H. developed the CLM-PFLOTRAN framework that this work built upon. F.Y., G.T., G. B., and X.X. added biogeochemistry to the CLM-PFLOTRAN interface. F. Y. proposed the nitrification and denitrification reactions and rate formulae. G. T., F. Y., and X. X. implemented the CLM biogeochemistry in PFLOTRAN under guidance of G. H., P. L., S. P., and P.T. G.T. prepared the manuscript with contributions from all co-authors. G. T., F. Y., G. B., and G.H.
710 contributed equally to the work.

Acknowledgements. Thanks to Nathaniel O. Collier at ORNL for many discussion that contributed significantly to this work. Thanks to Kathie Tallant and Cathy Jones at ORNL for editing service. This research was funded by the U.S. Department of Energy, Office of Sciences, Biological and Environmental Research, Terrestrial Ecosystem Sciences and Subsurface Biogeochemical Research Program, and is a product of the Next-Generation
715 Ecosystem Experiments in the Arctic (NGEE-Arctic) project. ORNL is managed by UT-Battelle, LLC, for the U.S. Department of Energy under contract DE-AC05-00OR22725.

References

- Antonelli, L., Briani, M., D'Ambra, P., and Fraioli, V.: Positivity Issues in Adaptive Solutions of Detailed Chemical Schemes for Engine Simulations, in: Communications to SIMAI Congress, vol. 3, pp. 303.1–303.12, 2009.
- Bethke, C. M.: Geochemical and biogeochemical reaction modeling, Cambridge University Press, 2007.
- Bonan, G. B., Hartman, M. D., Parton, W. J., and Wieder, W. R.: Evaluating litter decomposition in earth system models with long-term litterbag experiments: an example using the Community Land Model version 4 (CLM4), *Global Change Biology*, pp. 957–974, doi:10.1111/gcb.12031, <http://dx.doi.org/10.1111/gcb.12031>, 2012.
- Boyer, E. W., Alexander, R. B., Parton, W. J., Li, C., Butterbach-Bahl, K., Donner, S. D., Skaggs, R. W., and Grosso, S. J. D.: Modeling Denitrification In Terrestrial And Aquatic Ecosystems At Regional Scales, *Ecological Applications*, 16, 2123–2142, doi:10.1890/1051-0761(2006)016, [http://dx.doi.org/10.1890/1051-0761\(2006\)016](http://dx.doi.org/10.1890/1051-0761(2006)016), 2006.
- Broekhuizen, N., Rickard, G. J., Bruggeman, J., and Meister, A.: An improved and generalized second order, unconditionally positive, mass conserving integration scheme for biochemical systems, *Applied Numerical Mathematics*, 58, 319–340, doi:10.1016/j.apnum.2006.12.002, <http://www.sciencedirect.com/science/article/pii/S0168927406002224>, 2008.
- Bruggeman, J., Burchard, H., Kooi, B. W., and Sommeijer, B.: A second-order, unconditionally positive, mass-conserving integration scheme for biochemical systems, *Applied Numerical Mathematics*, 57, 36–58, doi:10.1016/j.apnum.2005.12.001, <http://www.sciencedirect.com/science/article/pii/S0168927405002242>, 2007.
- Burchard, H., Deleersnijder, E., and Meister, A.: A high-order conservative Patankar-type discretisation for stiff systems of production–destruction equations, *Applied Numerical Mathematics*, 47, 1–30, doi:10.1016/S0168-9274(03)00101-6, <http://www.sciencedirect.com/science/article/pii/S0168927403001016>, 2003.
- Burchard, H., Deleersnijder, E., and Meister, A.: Application of modified Patankar schemes to stiff biogeochemical models for the water column, *Ocean Dynamics*, 55, 326–337, doi:10.1007/s10236-005-0001-x, <http://link.springer.com/article/10.1007%2Fs10236-005-0001-x>, 2005.
- Chapin, F., Matson, P., and Vitousek, P.: Principles of Terrestrial Ecosystem Ecology, Springer, 2011.
- Chen, D. and Plemmons, R. J.: Nonnegativity constraints in numerical analysis, in: *Symposium on the Birth of Numerical Analysis*, pp. 109–140, 2009.
- da Costa, A. C. L., Galbraith, D., Almeida, S., Portela, B. T. T., da Costa, M., de Athaydes Silva Junior, J., Braga, A. P., de Gonçalves, P. H. L., de Oliveira, A. A. R., Fisher, R., Phillips, O. L., Metcalfe, D. B., Levy, P., and Meir, P.: Effect of 7 yr of experimental drought on vegetation dynamics and biomass storage of an eastern Amazonian rainforest, *New Phytologist*, 187, 579–591, doi:10.1111/j.1469-8137.2010.03309.x, <http://dx.doi.org/10.1111/j.1469-8137.2010.03309.x>, 2010.
- Dickinson, R. E., Berry, J. A., Bonan, G. B., Collatz, G. J., Field, C. B., Fung, I. Y., Goulden, M., Hoffmann, W. A., Jackson, R. B., Myneni, R., Sellers, P. J., and Shaikh, M.: Nitrogen Controls on Climate Model Evapotranspiration, *Journal of Climate*, 15, 278–295, doi:10.1175/1520-0442(2002)015, [http://dx.doi.org/10.1175/1520-0442\(2002\)015](http://dx.doi.org/10.1175/1520-0442(2002)015), 2002.

- Eltrop, L. and Marschner, H.: Growth and mineral nutrition of non-mycorrhizal and mycorrhizal Norway spruce (*Picea abies*) seedlings grown in semi-hydroponic sand culture, *New Phytologist*, 133, 469–478, doi:10.1111/j.1469-8137.1996.tb01914.x, <http://dx.doi.org/10.1111/j.1469-8137.1996.tb01914.x>, 1996.
- 760 Falkengren-Grerup, U.: Interspecies differences in the preference of ammonium and nitrate in vascular plants, *Oecologia*, 102, 305–311, doi:10.1007/BF00329797, <http://dx.doi.org/10.1007/BF00329797>, 1995.
- Fang, Y., Huang, M., Liu, C., Li, H., and Leung, L. R.: A generic biogeochemical module for Earth system models: Next Generation BioGeoChemical Module (NGBGC), version 1.0, *Geosci. Model Dev.*, 6, 1977–1988, doi:10.5194/gmd-6-1977-2013, <http://www.geosci-model-dev.net/6/1977/2013>, gMD, 2013.
- 765 Fennell, D. E. and Gossett, J. M.: Modeling the Production of and Competition for Hydrogen in a Dechlorinating Culture, *Environmental Science & Technology*, 32, 2450–2460, doi:10.1021/es980136l, <http://pubs.acs.org/doi/pdfplus/10.1021/es980136l>, 1998.
- Foulland, E., Gosselin, M., Rivkin, R. B., Vasseur, C., and Mostajir, B.: Nitrogen uptake by heterotrophic bacteria and phytoplankton in Arctic surface waters, *Journal of Plankton Research*, 29, 369–376, <http://plankt.oxfordjournals.org/content/29/4/369.abstract>, 2007.
- 770 Gherardi, L. A., Sala, O. E., and Yahdjian, L.: Preference for different inorganic nitrogen forms among plant functional types and species of the Patagonian steppe, *Oecologia*, 173, 1075–1081, doi:10.1007/s00442-013-2687-7, <http://dx.doi.org/10.1007/s00442-013-2687-7>, 2013.
- Grant, R. F.: Modelling changes in nitrogen cycling to sustain increases in forest productivity under elevated atmospheric CO₂ and contrasting site conditions, *Biogeosciences*, 10, 7703–7721, doi:10.5194/bg-10-7703-2013, <http://www.biogeosciences.net/10/7703/2013>, bG, 2013.
- 775 Gu, C. and Riley, W. J.: Combined effects of short term rainfall patterns and soil texture on soil nitrogen cycling — A modeling analysis, *Journal of Contaminant Hydrology*, 112, 141–154, doi:10.1016/j.jconhyd.2009.12.003, <http://www.sciencedirect.com/science/article/pii/S0169772209001648>, 2010.
- 780 Hammond, G. E.: Innovative Methods for Solving Multicomponent Biogeochemical groundwater Transport on Supercomputers, Thesis, University of Illinois at Urbana-Champaign, 2003.
- Hanson, P. and Wullschleger, S.: North American Temperate Deciduous Forest Responses to Changing Precipitation Regimes, Springer Verlag, 2003.
- 785 Hanson, P. J., Amthor, J. S., Wullschleger, S. D., Wilson, K. B., Grant, R. F., Hartley, A., Hui, D., Hunt, J. E. R., Johnson, D. W., Kimball, J. S., King, A. W., Luo, Y., McNulty, S. G., Sun, G., Thornton, P. E., Wang, S., Williams, M., Baldocchi, D. D., and Cushman, R. M.: Oak forest carbon and water simulations: model intercomparisons and evaluations against independent data, *Ecological Monographs*, 74, 443–489, doi:10.1890/03-4049, <http://www.esajournals.org/doi/pdf/10.1890/03-4049>, 2004.
- 790 Hinkel, K. M. and Nelson, F. E.: Spatial and temporal patterns of active layer thickness at Circumpolar Active Layer Monitoring (CALM) sites in northern Alaska, 1995–2000, *Journal of Geophysical Research: Atmospheres*, 108, 8168, doi:10.1029/2001JD000927, <http://dx.doi.org/10.1029/2001JD000927>, 2003.
- Hungate, R.: The rumen microbial ecosystem, *Annual Review of Ecology and Systematics*, pp. 39–66, 1975.
- 795 HØGh-Jensen, H., Wollenweber, B., and Schjoerring, J. K.: Kinetics of nitrate and ammonium absorption and accompanying H⁺ fluxes in roots of *Lolium perenne* L. and N₂-fixing *Trifolium repens* L, *Plant, Cell & Envi-*

- ronment, 20, 1184–1192, doi:10.1046/j.1365-3040.1997.d01-145.x, <http://dx.doi.org/10.1046/j.1365-3040.1997.d01-145.x>, 1997.
- IPCC: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, doi:10.1017/CBO9781107415324, www.climatechange2013.org, 2013.
- Jarrell, K. F.: Extreme Oxygen Sensitivity in Methanogenic Archaeabacteria, *BioScience*, 35, 298–302, <http://bioscience.oxfordjournals.org/content/35/5/298.abstract>, 1985.
- Kamer, K., Kennison, R. L., and Fong, P.: Rates of inorganic nitrogen uptake by the estuarine green macroalgae *Enteromorpha intestinalis* and *Ulva expansa*, vol. 2003, pp. 130–141, 2001.
- Kirchman, D. L.: The uptake of inorganic nutrients by heterotrophic bacteria, *Microbial Ecology*, 28, 255–271, doi:10.1007/BF00166816, <http://dx.doi.org/10.1007/BF00166816>, 1994.
- Kirchman, D. L. and Wheeler, P. A.: Uptake of ammonium and nitrate by heterotrophic bacteria and phytoplankton in the sub-Arctic Pacific, *Deep Sea Research Part I: Oceanographic Research Papers*, 45, 347–365, doi:10.1016/S0967-0637(97)00075-7, <http://www.sciencedirect.com/science/article/pii/S0967063797000757>, 1998.
- Kuzyakov, Y. and Xu, X.: Competition between roots and microorganisms for nitrogen: mechanisms and ecological relevance, *New Phytologist*, 198, 656–669, doi:10.1111/nph.12235, <http://dx.doi.org/10.1111/nph.12235>, 2013.
- Lara, M. J., Villarreal, S., Johnson, D. R., Hollister, R. D., Webber, P. J., and Tweedie, C. E.: Estimated change in tundra ecosystem function near Barrow, Alaska between 1972 and 2010, *Environmental Research Letters*, 7, 015 507, 2012.
- Lichtner, P. C., Hammond, G. E., Lu, C., Karra, S., Bisht, G., Andre, B., Mills, R. T., and Jitu, K.: PFLOTRAN User Manual: A Massively Parallel Reactive Flow and Transport Model for Describing Surface and Subsurface Processes, Report, 2015.
- Maggi, F., Gu, C., Riley, W. J., Hornberger, G. M., Venterea, R. T., Xu, T., Spycher, N., Steefel, C., Miller, N. L., and Oldenburg, C. M.: A mechanistic treatment of the dominant soil nitrogen cycling processes: Model development, testing, and application, *Journal of Geophysical Research: Biogeosciences*, 113, G02 016, doi:10.1029/2007JG000578, <http://dx.doi.org/10.1029/2007JG000578>, 2008.
- Manzoni, S. and Porporato, A.: Soil carbon and nitrogen mineralization: Theory and models across scales, *Soil Biology and Biochemistry*, 41, 1355–1379, doi:10.1016/j.soilbio.2009.02.031, <http://www.sciencedirect.com/science/article/pii/S0038071709000765>, 2009.
- Middelburg, J. J. and Nieuwenhuize, J.: Nitrogen uptake by heterotrophic bacteria and phytoplankton in the nitrate-rich Thames estuary, *Marine Ecology Progress Series*, 203, 13–21, <http://www.int-res.com/abstracts/meps/v203/p13-21/>, 2000.
- Nollet, L. M. L. and De Gelder, L. S. P.: *Handbook of Water Analysis* (3rd Edition), CRC Press, 2013.
- Nordin, A., Högberg, P., and Näsholm, T.: Soil nitrogen form and plant nitrogen uptake along a boreal forest productivity gradient, *Oecologia*, 129, 125–132, doi:10.1007/s004420100698, <http://dx.doi.org/10.1007/s004420100698>, 2001.

- 835 Oleson, K., Lawrence, D., Bonan, G., Levis, S., Swenson, S., Thornton, P., Bozbayik, A., Fisher, R., Heald, C.,
 Kluzek, E., Lamarque, J.-F., Lawrence, P., Lipscomb, W., Muszala, S., and Sacks, W.: Technical description
 of version 4.5 of the Community Land Model (CLM), Ncar/tn-503+str, ncar technical note, NCAR,
 doi:10.5065/D6RR1W7M, http://www.cesm.ucar.edu/models/cesm1.2/clm/CLM45_Tech_Note.pdf, 2013.
- Parkhurst, D. L. and Appelo, C.: User's guide to PHREEQC (Version 2): A computer program for speciation,
 840 batch-reaction, one-dimensional transport, and inverse geochemical calculations, Water-Resources Investigations 99-4259, USGS, 1999.
- Parton, W. J., Mosier, A. R., Ojima, D. S., Valentine, D. W., Schimel, D. S., Weier, K., and Kulmala, A. E.:
 Generalized model for N₂ and N₂O production from nitrification and denitrification, Global Biogeochemical
 Cycles, 10, 401–412, doi:10.1029/96GB01455, <http://dx.doi.org/10.1029/96GB01455>, 1996.
- 845 Parton, W. J., Holland, E. A., Del Gross, S. J., Hartman, M. D., Martin, R. E., Mosier, A. R., Ojima, D. S., and
 Schimel, D. S.: Generalized model for NO_x and N₂O emissions from soils, Journal of Geophysical Research:
 Atmospheres, 106, 17 403–17 419, doi:10.1029/2001JD900101, <http://dx.doi.org/10.1029/2001JD900101>,
 2001.
- Pfautsch, S., Rennenberg, H., Bell, T. L., and Adams, M. A.: Nitrogen uptake by Eucalyptus regnans and
 850 Acacia spp. – preferences, resource overlap and energetic costs, Tree Physiology, 29, 389–399, <http://treephys.oxfordjournals.org/content/29/3/389.abstract>, 2009.
- Pierre, M. and Schmitt, D.: Blowup in Reaction-Diffusion Systems with Dissipation of Mass, SIAM Review, 42, 93–106, doi:10.1137/S0036144599359735, <http://dx.doi.org/10.1137/S0036144599359735><http://epubs.siam.org/doi/abs/10.1137/S0036144599359735>, 2000.
- 855 Powell, T. L., Galbraith, D. R., Christoffersen, B. O., Harper, A., Imbuzeiro, H. M. A., Rowland, L., Almeida, S.,
 Brando, P. M., da Costa, A. C. L., Costa, M. H., Levine, N. M., Malhi, Y., Saleska, S. R., Sotta, E., Williams,
 M., Meir, P., and Moorcroft, P. R.: Confronting model predictions of carbon fluxes with measurements of
 Amazon forests subjected to experimental drought, New Phytologist, 200, 350–365, doi:10.1111/nph.12390,
<http://dx.doi.org/10.1111/nph.12390>, 2013.
- 860 Riley, W. J., Maggi, F., Kleber, M., Torn, M. S., Tang, J. Y., Dwivedi, D., and Guerry, N.: Long residence times
 of rapidly decomposable soil organic matter: application of a multi-phase, multi-component, and vertically
 resolved model (BAMS1) to soil carbon dynamics, Geosci. Model Dev., 7, 1335–1355, doi:10.5194/gmd-7-
 1335-2014, <http://www.geosci-model-dev.net/7/1335/2014/>, gMD, 2014.
- Schmidt, M. W. I., Torn, M. S., Abiven, S., Dittmar, T., Guggenberger, G., Janssens, I. A., Kleber, M., Kogel-
 865 Knabner, I., Lehmann, J., Manning, D. A. C., Nannipieri, P., Rasse, D. P., Weiner, S., and Trumbore, S. E.:
 Persistence of soil organic matter as an ecosystem property, Nature, 478, 49–56, <http://dx.doi.org/10.1038/nature10386>, 10.1038/nature10386, 2011.
- Sellers, P. J., Dickinson, R. E., Randall, D. A., Betts, A. K., Hall, F. G., Berry, J. A., Collatz, G. J., Denning,
 A. S., Mooney, H. A., Nobre, C. A., Sato, N., Field, C. B., and Henderson-Sellers, A.: Modeling the Ex-
 870 changes of Energy, Water, and Carbon Between Continents and the Atmosphere, Science, 275, 502–509,
 doi:10.1126/science.275.5299.502, <http://www.sciencemag.org/content/275/5299/502.abstract>, 1997.
- Seneviratne, S. I., Corti, T., Davin, E. L., Hirschi, M., Jaeger, E. B., Lehner, I., Orlowsky, B., and Teuling,
 A. J.: Investigating soil moisture–climate interactions in a changing climate: A review, Earth-Science Re-

- views, 99, 125–161, doi:10.1016/j.earscirev.2010.02.004, <http://www.sciencedirect.com/science/article/pii/S0012825210000139>, 2010.
- Shampine, L. F., Thompson, S., Kierzenka, J. A., and Byrne, G. D.: Non-negative solutions of ODEs, *Applied Mathematics and Computation*, 170, 556–569, doi:<http://dx.doi.org/10.1016/j.amc.2004.12.011>, <http://www.sciencedirect.com/science/article/pii/S0096300304009683>, 2005.
- Steefel, C. I., Appelo, C. A. J., Arora, B., Jacques, D., Kalbacher, T., Kolditz, O., Lagneau, V., Lichtner, P. C.,
880 Mayer, K. U., Meeussen, J. C. L., Molins, S., Moulton, D., Shao, H., Šimůnek, J., Spycher, N., Yabusaki,
S. B., and Yeh, G. T.: Reactive transport codes for subsurface environmental simulation, *Computational Geosciences*, pp. 1–34, doi:10.1007/s10596-014-9443-x, <http://dx.doi.org/10.1007/s10596-014-9443-x>, 2014.
- Tang, G., Watson, D. B., Wu, W.-M., Schadt, C. W., Parker, J. C., and Brooks, S. C.: U(VI) Bioreduction with
Emulsified Vegetable Oil as the Electron Donor – Model Application to a Field Test, *Environmental Science
885 & Technology*, 47, 3218–3225, doi:10.1021/es304643h, <http://dx.doi.org/10.1021/es304643h>, 2013a.
- Tang, J., Riley, W., Koven, C., and Subin, Z.: CLM4-BeTR, a generic biogeochemical transport and reaction
module for CLM4: model development, evaluation, and application, *Geoscientific Model Development*, 6,
127–140, 2013b.
- Tang, J. Y. and Riley, W. J.: Technical Note: A generic law-of-the-minimum flux limiter for simulating sub-
890 strate limitation in biogeochemical models, *Biogeosciences Discuss.*, 12, 13 399–13 425, doi:10.5194/bgd-
12-13399-2015, <http://www.biogeosciences-discuss.net/12/13399/2015/>, bGD, 2015.
- Thornton, P. E. and Rosenbloom, N. A.: Ecosystem model spin-up: Estimating steady state con-
ditions in a coupled terrestrial carbon and nitrogen cycle model, *Ecological Modelling*, 189,
25–48, doi:10.1016/j.ecolmodel.2005.04.008, <http://www.sciencedirect.com/science/article/pii/S0304380005001948>, 2005.
- Veuger, B., Middelburg, J. J., Boschker, H. T. S., Nieuwenhuize, J., van Rijswijk, P., Rochelle-Newall, E. J.,
and Navarro, N.: Microbial uptake of dissolved organic and inorganic nitrogen in Randers Fjord, Estuar-
ine, Coastal and Shelf Science, 61, 507–515, doi:10.1016/j.ecss.2004.06.014, <http://www.sciencedirect.com/science/article/pii/S0272771404001593>, 2004.
- 900 Wang, G., Post, W. M., and Mayes, M. A.: Development of microbial-enzyme-mediated decompo-
sition model parameters through steady-state and dynamic analyses, *Ecological Applications*, 23,
255–272, doi:10.1890/12-0681.1, <http://dx.doi.org/10.1890/12-0681.1><http://www.esajournals.org/doi/abs/10.1890/12-0681.1>, 2012.
- Warren, C. R. and Adams, P. R.: Uptake of nitrate, ammonium and glycine by plants of Tasmanian wet eu-
905 calypt forests, *Tree Physiology*, 27, 413–419, <http://treephys.oxfordjournals.org/content/27/3/413.abstract>,
10.1093/treephys/27.3.413, 2007.
- Xu, T., Sonnenthal, E., Spycher, N., and Pruess, K.: TOUGHREACT—A simulation program
for non-isothermal multiphase reactive geochemical transport in variably saturated geologic me-
dia: Applications to geothermal injectivity and CO₂ geological sequestration, *Computers & Geo-
910 sciences*, 32, 145–165, doi:10.1016/j.cageo.2005.06.014, <http://www.sciencedirect.com/science/article/pii/S0098300405001500>, 2006.

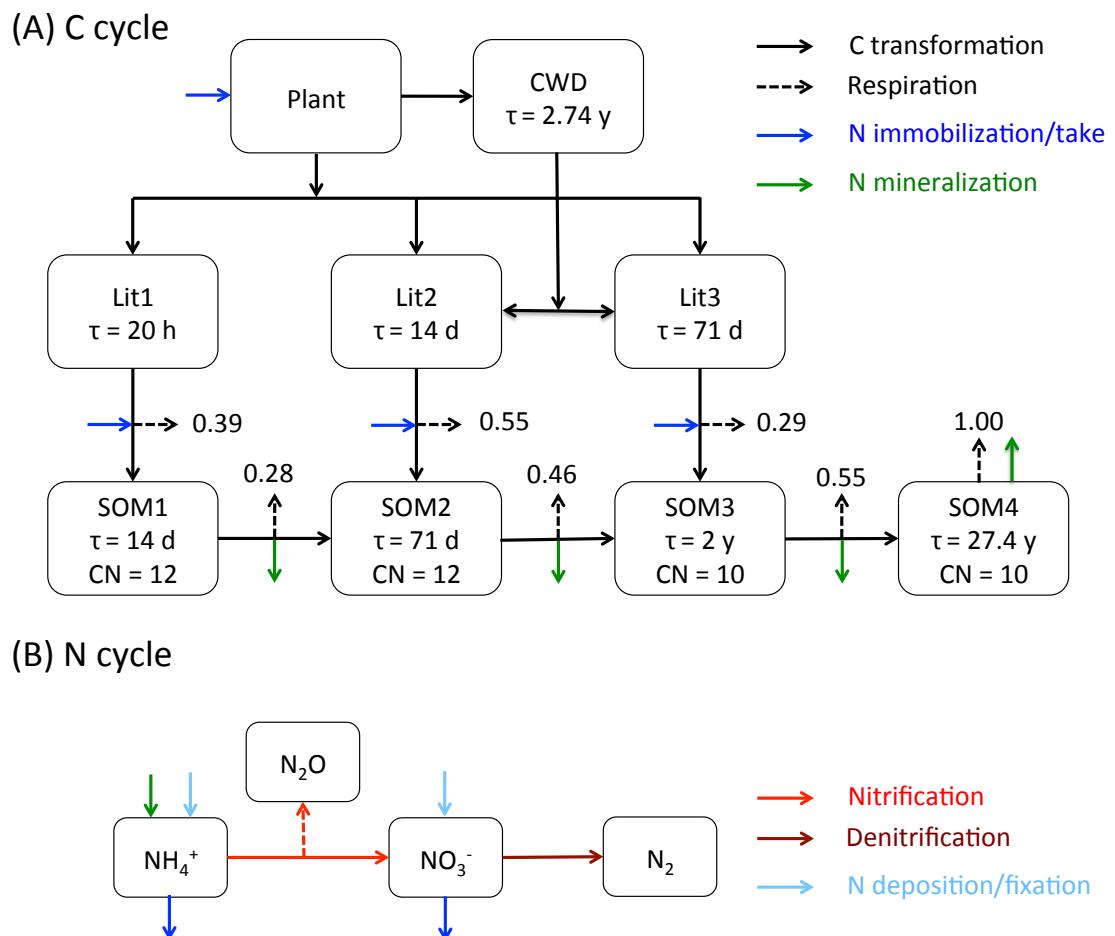


Figure 1. The reaction network for the carbon (A) and nitrogen (B) cycles implemented in this work. The carbon cycle is modified from Thornton and Rosenbloom (2005) and Bonan et al. (2012)

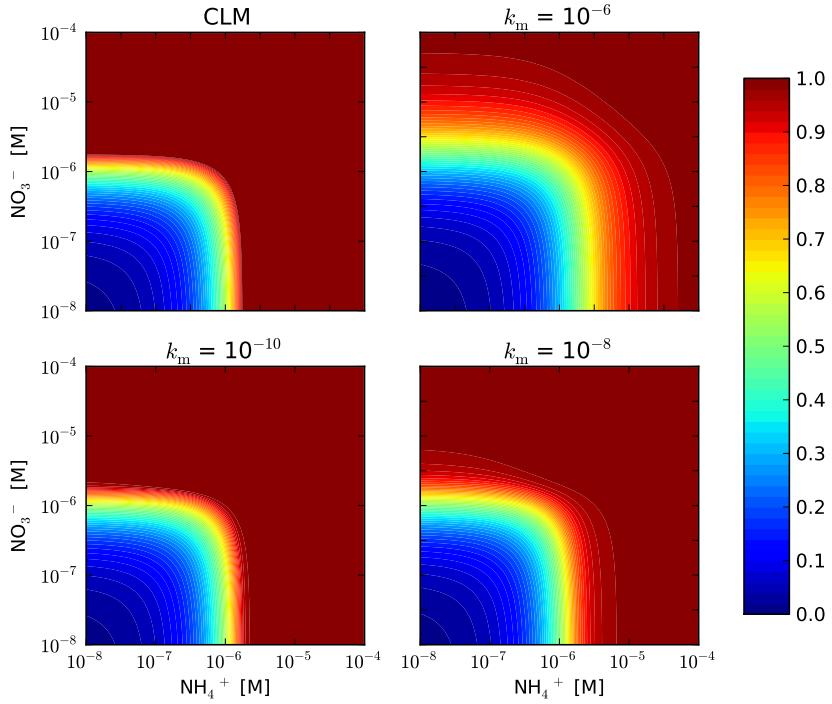


Figure 2. The ratio of uptake and demand (f_{pi}) as a function of concentrations with CLM and representation by Eqs. (19) and (20) in a 0.5 h time step with an uptake rate of 10^{-9} M s^{-1} . f_{pi} for the new representation is less than or equal to that for CLM. The difference decreases with decreasing half saturation k_m .

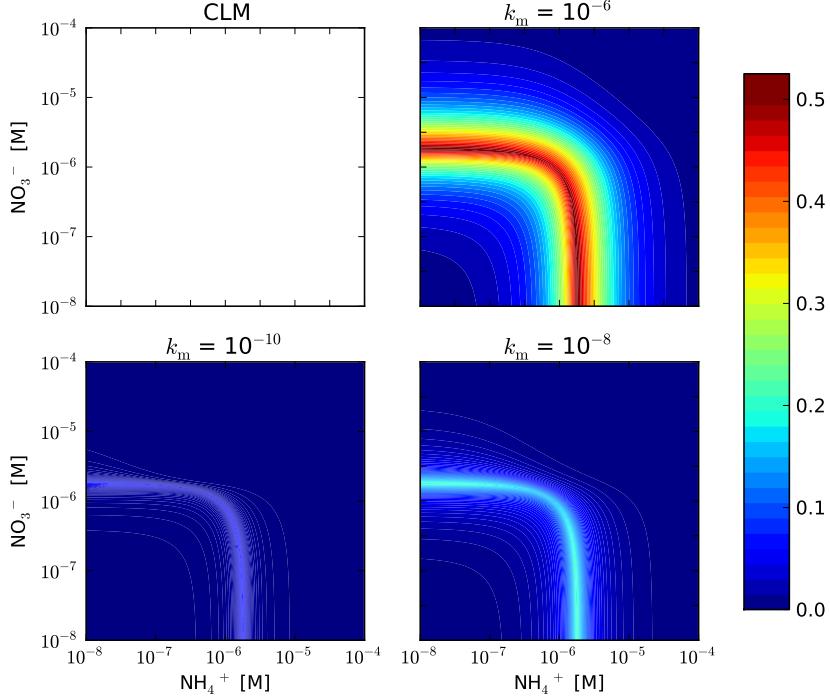


Figure 3. The difference plots for Fig. (2).

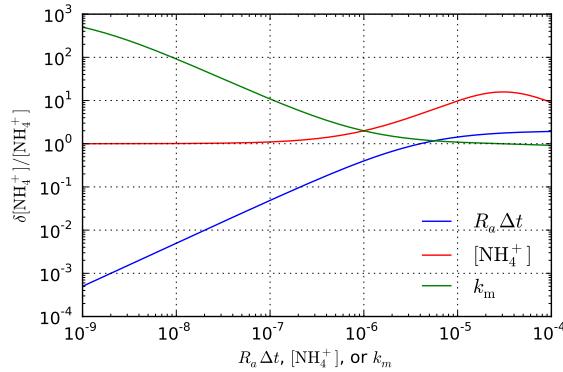


Figure 4. The calculated ratio of update over concentration for the first iteration as a function of time step size ($R_a \Delta t$), initial concentration (NH_4^+), and half saturation (k_m) for solving $d[\text{NH}_4^+]/dt = -R_a[\text{NH}_4^+]/(\text{NH}_4^+ + k_m)$ using the backward Euler method and Newton-Raphson method. Without scaling back the update (δ), the concentration in the first iteration step ($[\text{NH}_4^+]^{k+1,1}$) can go negative if $R_a \Delta t > 10^{-5}$, $[\text{NH}_4^+]^k > 10^{-6}$, or $k_m < 10^{-6}$. Unless specified in the figure, the base case parameters are $R_a \Delta t = 10^{-3}$, $\text{NH}_4^+ = 10^{-6}$, and $k_m = 10^{-6}$.

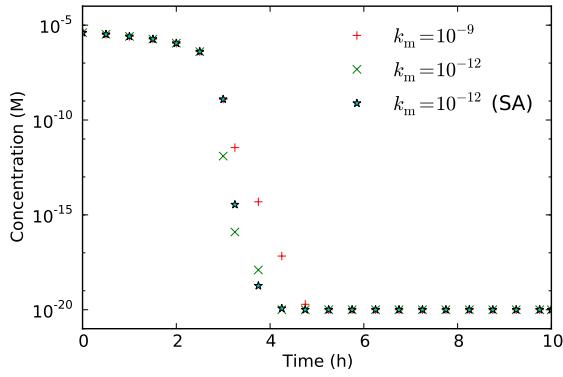


Figure 5. Influence of half saturation (k_m) on the simulated NH_4^+ concentration decrease due to plant uptake. Too small a half saturation may result in false convergence due to too small an update (scaling factor) during the iteration. The semi-analytical (SA) solution (Eq. D2) is used for comparison.

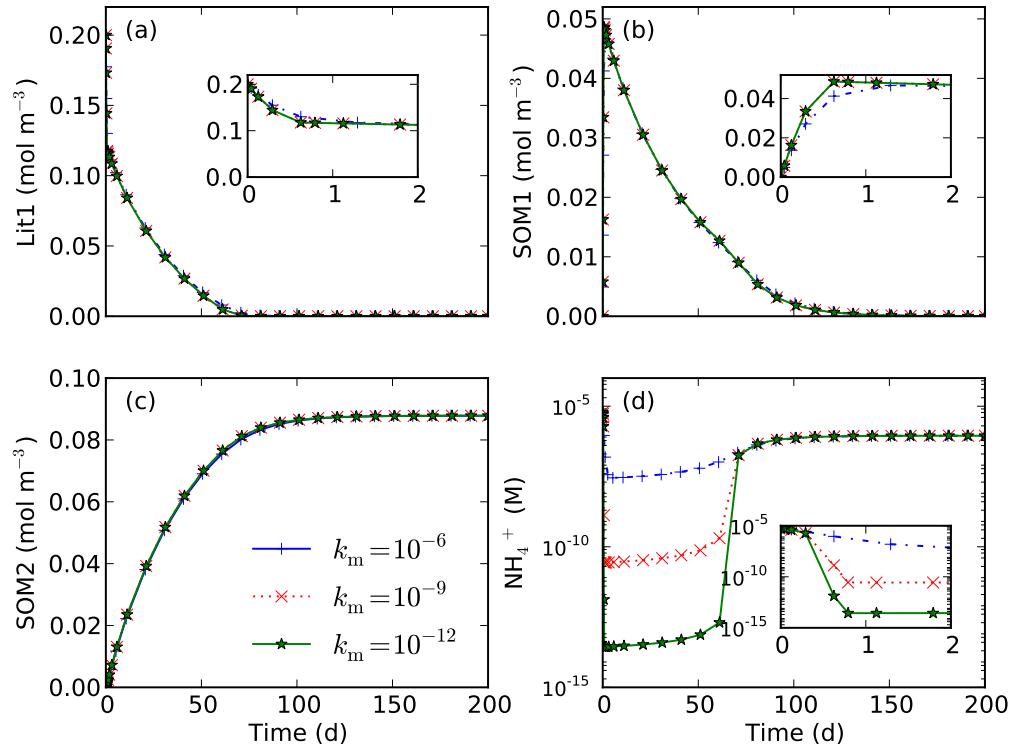


Figure 6. Influence of half saturation k_m on decomposition that involves both nitrogen immobilization and mineralization. Smaller half saturation can result in lower nitrogen concentration (d) but does not substantially impact the calculated concentrations other than NH_4^+ (a,b,c).

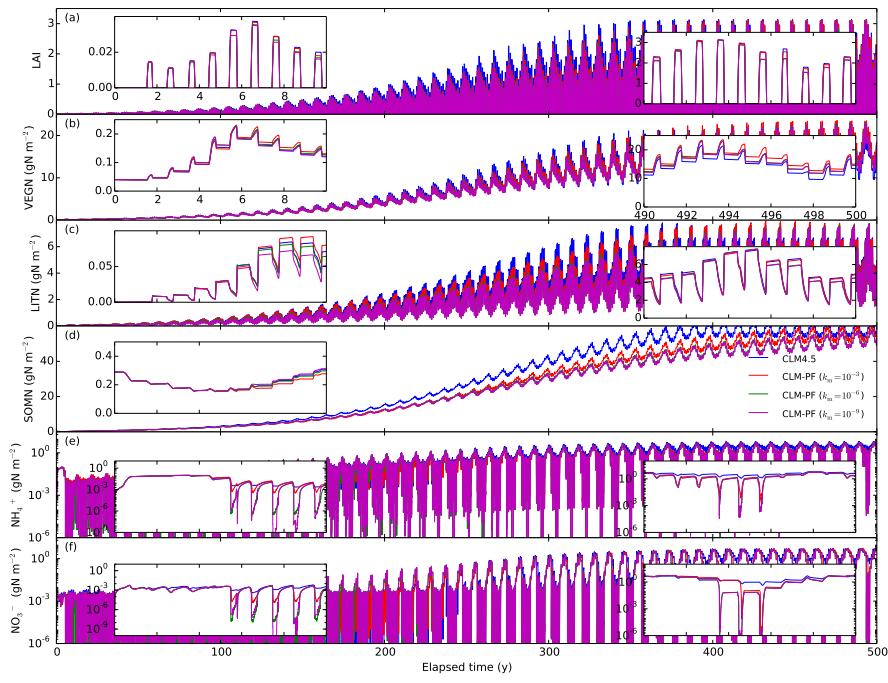


Figure 7. Calculated LAI and nitrogen distribution among vegetation, litter, SOM, NH_4^+ , and NO_3^- pools in spin-up simulations for the US-Brw site. Log transformation is used to enforce nonnegativity for CLM-PFLTRAN.

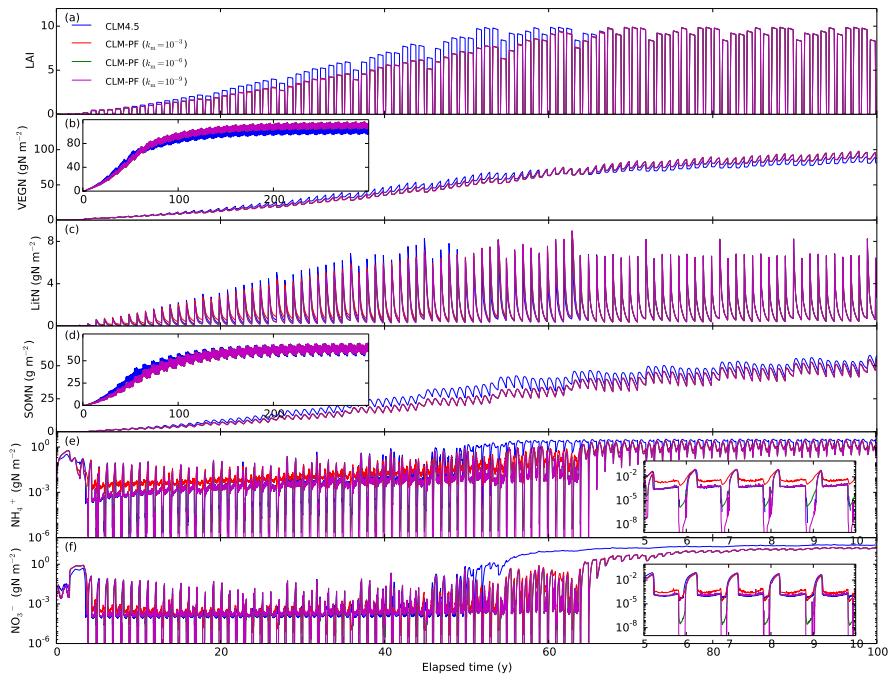


Figure 8. Calculated LAI and nitrogen distribution among vegetation, litter, SOM, NH_4^+ , and NO_3^- pools in spin-up simulations for US-WBW site. Log transformation is used to enforce nonnegativity for CLM-PFLTRAN calculations.

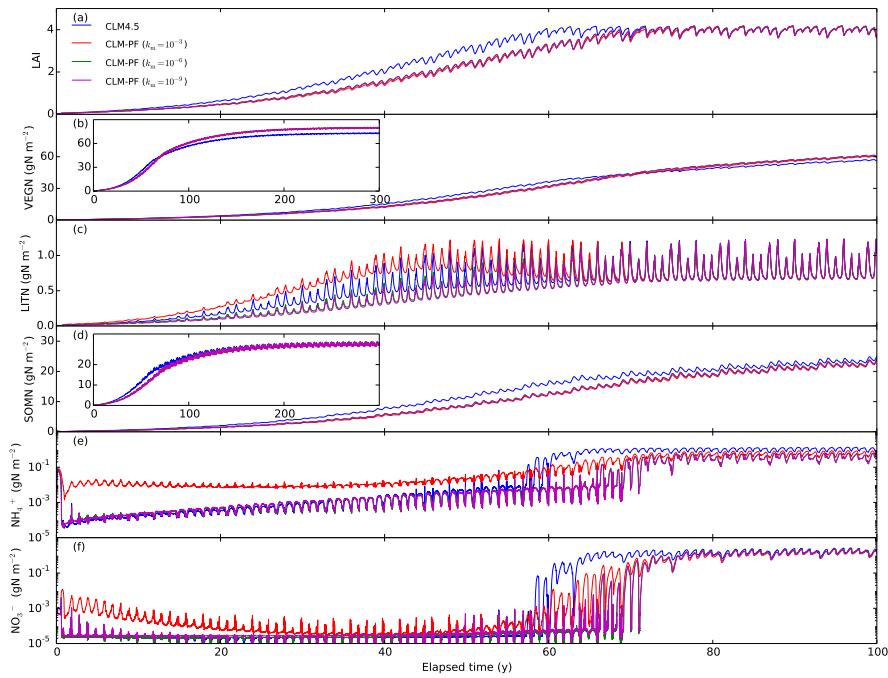


Figure 9. Calculated LAI and nitrogen distribution among vegetation, litter, SOM, NH_4^+ , and NO_3^- pools in spin-up simulations for BR-Cax site. Log transformation is used to enforce nonnegativity for CLM-PFLOTRAN calculations.

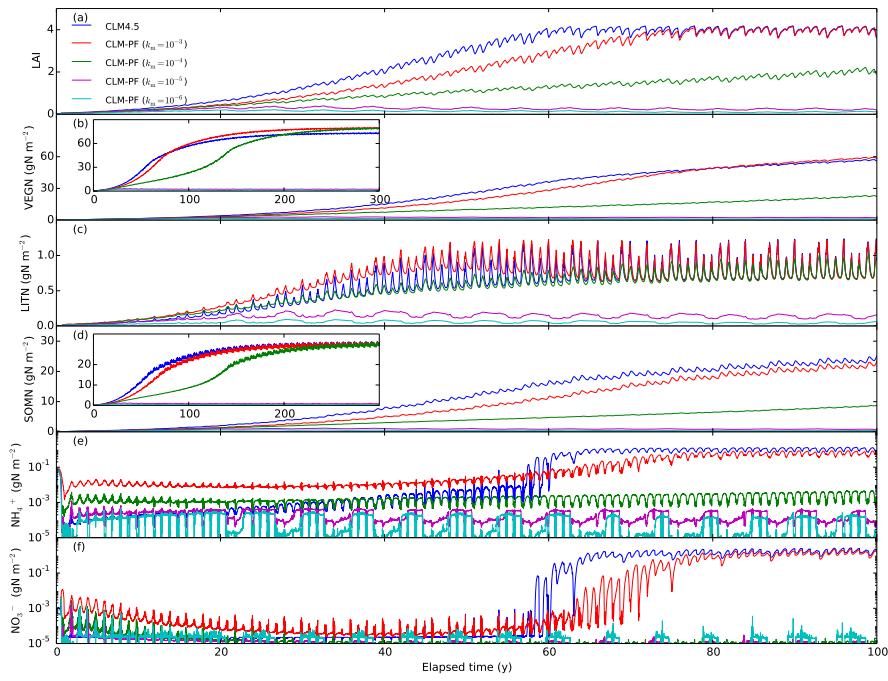


Figure 10. Calculated LAI and nitrogen distribution among vegetation, litter, SOM, NH_4^+ , and NO_3^- pools in spin-up simulations for BR-Cax site. Scaling back update in each iteration is used enforce nonnegativity for CLM-PFLOTTRAN.

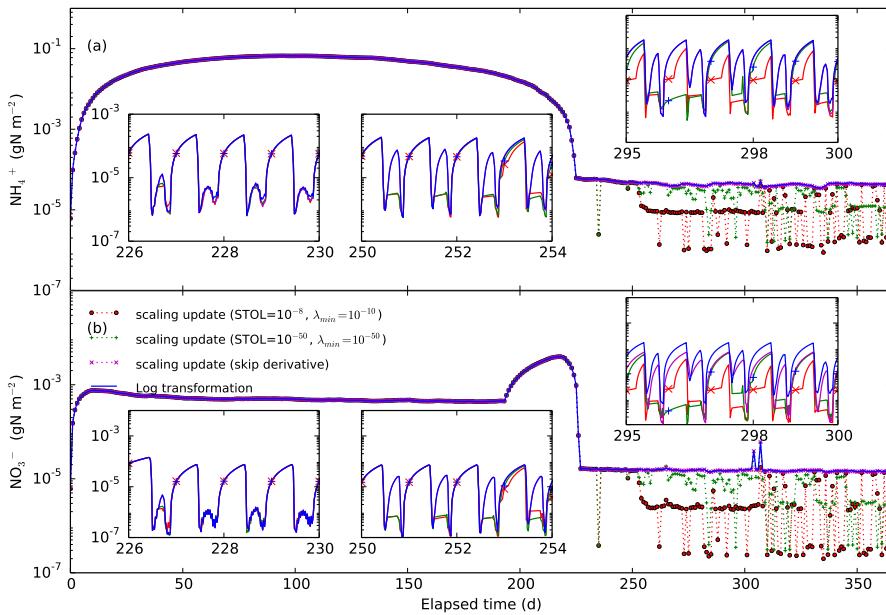


Figure 11. Calculated NH_4^+ and NO_3^- concentration in the first year using SU (scaling update) vs. LT (log transformation) in the spin-up simulation for BR-Cax site with $k_m = 10^{-6} \text{ mol m}^{-3}$. “skip derivative” refers to not including the derivatives for the reaction (R11) with rate (Eq. A5) in the entries for N_2Od in the Jacobian matrix.

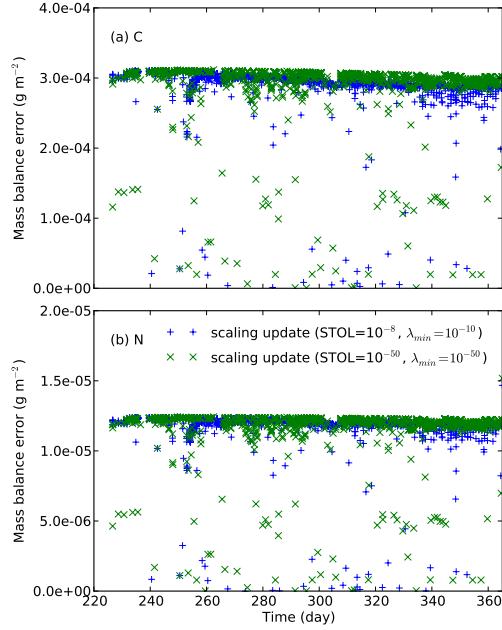


Figure 12. Mass balance error for (a) C and (b) N in the first year where SU with $k_m = 10^{-6}$ mol m $^{-3}$ is used for the tropical site.

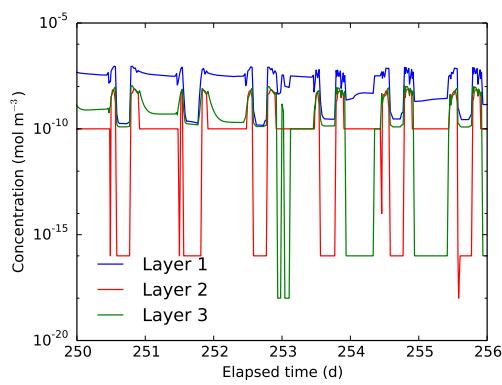


Figure 13. Diagnostic N₂O concentration (N₂Od from nitrification associated with net nitrogen mineralization Reaction R11 and rate Eq. A5) in the spin-up simulation for the BR-Cax site with $k_m = 10^{-6}$ mol m $^{-3}$. The concentration is reset to 10^{-10} at the beginning of each CLM half hour time step. Scaling back update in each iteration is used to enforce nonnegativity.

Table 1. Wall time (hour) for spin-up simulation at the arctic, temperate, and tropical sites on OIC

| Site | CLM | SU (10^{-3}) | SU (10^{-6}) | SU (10^{-9}) | LT (10^{-3}) | LT (10^{-6}) | LT (10^{-9}) | LT (10^{-12}) |
|--------|------|------------------|------------------|------------------|------------------|------------------|------------------|-------------------|
| DC | | | | | | | | |
| US-Brw | 18.1 | 24.8 | 25.5 | 29.2 | 38.5 | 40.8 | 47.0 | 49.5 |
| US-Pit | 11.7 | 17.1 | 14.8 | 14.9 | 30.0 | 37.5 | 40.2 | 43.1 |
| BR-Cax | 10.9 | 16.2 | 18.6 | 18.0 | 40.2 | 40.5 | 45.7 | 52.0 |
| DR | | | | | | | | |
| US-Brw | 18.1 | 21.5 | 21.5 | | 35.8 | 35.9 | | |
| US-Pit | 11.7 | 14.1 | 14.1 | | 26.9 | 26.7 | | |
| BR-Cax | 10.9 | 16.4 | 16.4 | | 42.2 | 41.6 | | |

SU = scaling update, LT = log transformation, 10^{-3} , 10^{-6} , and 10^{-9} are k_m , DC and DR = downregulating consumption as a function of concentration and rate, OIC = ORNL Institutional Cluster (Phase5), US-Brw simulation duration = 500 year and US-Pit and BR-Cax simulation duration = 300 year. For the last column LT (10^{-12}), MAX_CUT is increased from default 16 to 50.

Appendix A: CLM biogeochemical reactions and rates

A1 CLM-CN decomposition

The CLM-CN decomposition cascade consists of three litter pools with variable CN ratios, four soil
915 organic matter (SOM) pools with constant CN ratios, and seven reactions (Fig. 1A). The reaction
can be described by



with CN_u and CN_d as the upstream and downstream pool (molecular formula, for 1 mol upstream
and downstream pool, there is u and d mol N), N as either NH_4^+ or NO_3^- , f as the respiration fraction,
920 and $n = u - (1 - f)d$. The rate is

$$\frac{d[\text{CN}_u]}{dt} = -k_d f_T f_w [\text{CN}_u], \quad (\text{A1})$$

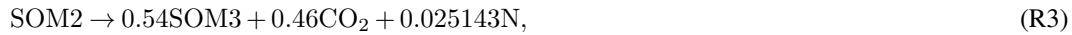
with k_d as the rate coefficient and f_T and f_w as the temperature and moisture response functions,

$$f_T = Q_{10}^{(T-25)/10}, \quad (\text{A2})$$

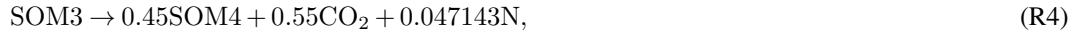
and

$$925 \quad f_w = \ln(\psi) Q_{10}^{(T-25)/10}, \quad (\text{A3})$$

With a constant CN ratio, the decomposition reactions for the four SOM pools are



930



and



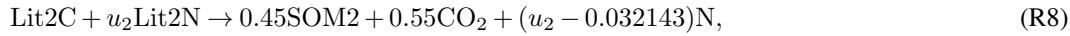
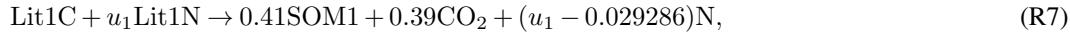
The exact stoichiometric coefficients are calculated in the code using values for respiration factor,
935 CN ratio, and molecular weight specified in the input file.

CLM4.5 has an option to separate N into NH_4^+ and NO_3^- . The N mineralization product is NH_4^+ .

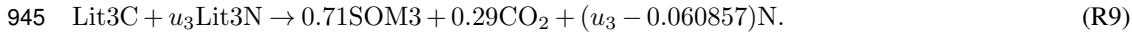
As the CN ratio is variable for the three litter pools, litter N pools need to be tracked such that
reaction (R1) becomes



940 with $u = [\text{LitN}]/[\text{LitC}]$. The three litter decomposition reactions are



and



As the CN ratio of the litter pools is generally high, u_1 , u_2 , and u_3 are usually small, and n in these reactions (e.g., $n_1 = u_1 - 0.029286$ for Lit1) is normally negative. Namely, these reactions consume (immobilize) N, which can be NH_4^+ , NO_3^- , or both.

A2 Nitrification

950 The nitrification reaction to produce NO_3^- is



with \dots for additional reactants and products to balance the reaction. The rate is (Dickinson et al., 2002)

$$\frac{d[\text{NH}_4^+]}{dt} = -\frac{d[\text{NO}_3^-]}{dt} = -k_n f_T f_w [\text{NH}_4^+]. \quad (\text{A4})$$

955 The nitrification reaction to produce N_2O is



with one component related to decomposition as

$$\frac{d[\text{NH}_4^+]}{dt} = -2 \frac{d[\text{N}_2\text{O}]}{dt} = -f_{nm} f_T f_w f_{pH} \max(R_{nm}, 0) \quad (\text{A5})$$

with f_{nm} as a fraction (Parton et al., 1996) and R_{nm} as the net N mineralization rate,

$$960 R_{nm} = \sum_i n_i R_i, \quad (\text{A6})$$

where R_i denotes the rate of reaction (R2, R3, R4, R5, R7, R8, R9). The second component is (Parton et al., 1996)

$$\frac{d[\text{NH}_4^+]}{dt} = -2 \frac{d[\text{N}_2\text{O}]}{dt} = -k_{n2o} f_T f_w f_{pH} (1 - e^{-0.0105[\text{NH}_4^+]}). \quad (\text{A7})$$

Ignoring the high-order terms and moving the unit conversion factor into k_{n2o} , it can be simplified

965 as a first-order rate as

$$\frac{d[\text{NH}_4^+]}{dt} = -2 \frac{d[\text{N}_2\text{O}]}{dt} = -k_{n2o} f_T f_w f_{pH} [\text{NH}_4^+]. \quad (\text{A8})$$

A3 Denitrification

The denitrification reaction is



970 with rate (Dickinson et al., 2002)

$$\frac{d[\text{NO}_3^-]}{dt} = -2\frac{d[\text{N}_2]}{dt} = -k_{\text{deni}} f_T f_w f_{\text{pH}} [\text{NO}_3^-]. \quad (\text{A9})$$

A4 Plant nitrogen uptake

The plant nitrogen uptake reaction can be written as



975 and



The rate is specified by CLM (plant nitrogen demand) and assumed to be constant in each half-hour time step.

A5 Demand-based competition and distributing nitrogen demand between NH_4^+ and NO_3^-

980 Denote $R_{d,p}$, $R_{d,i}$, $R_{d,nitr}$, $R_{d,deni}$ as the potential plant, immobilization, nitrification, and denitrification demand (rate); $R_{a,tot} = R_{d,p} + R_{d,i} + R_{d,nitr}$ as the total NH_4^+ demand; and $R_{n,tot}$ as the total NO_3^- demand. CLM uses a demand-based competition approach to split the available sources in proportion to the demand rates to meet the demands (Oleson et al., 2013; Thornton and Rosenbloom, 2005). Specifically, for each time step, if $R_{a,tot}\Delta t \leq [\text{NH}_4^+]$, the uptakes are equal
985 to potential demands, and $R_{n,tot} = 0$; otherwise, the uptakes for NH_4^+ are $[\text{NH}_4^+]R_{d,p}/R_{a,tot}\Delta t$, $[\text{NH}_4^+]R_{d,i}/R_{a,tot}\Delta t$, and $[\text{NH}_4^+]R_{d,nitr}/R_{a,tot}\Delta t$ for plants, immobilization, and nitrification, respectively; $R_{n,tot} = R_{a,tot} - [\text{NH}_4^+]/\Delta t + R_{d,deni}$. If $R_{n,tot}\Delta t < [\text{NO}_3^-]$, all of the remaining demand $R_{n,tot}$ is met with available NO_3^- . Otherwise, available NO_3^- is split to meet the remaining plant, immobilization, and denitrification demands in proportion to their rates.

990 Appendix B: Downregulation with cutoff

An alternative to using a residual concentration (Eq. 16) is to introduce a cutoff, e.g., Eq. (A1) becomes

$$\frac{d[\text{CN}_u]}{dt} = -k_d f_T f_w [\text{CN}_u] f([\text{CN}_u]), \quad (\text{B1})$$

with $f([CN_u]) = 1$ when $[CN_u] \geq [CN_u]_r$, and 0 otherwise. It is simple but does not prevent $[CN_u]$ from getting below $[CN_u]_r$ or 0 in theory. In addition, it introduces a discontinuity. A polynomial function can be used to smooth the cutoff

$$f([CN_u]) = 1 - \left[1 - \left(\frac{[CN_u] - [CN_u]_r}{[CN_u]_1 - [CN_u]_r} \right)^2 \right]^2. \quad (B2)$$

This function varies from 0 at $[CN_u]_r$ to 1 at $[CN_u]_1$, with zero derivatives at both points. While not implying a nonphysical reverse reaction as Eq. (16),

$$\frac{df([CN_u])}{d[CN_u]} = 4 \frac{[CN_u] - [CN_u]_r}{([CN_u]_1 - [CN_u]_r)^2} \left[1 - \left(\frac{[CN_u] - [CN_u]_r}{[CN_u]_1 - [CN_u]_r} \right)^2 \right]. \quad (B3)$$

Depending on $[CN_u]_1 - [CN_u]_r$, this cutoff does introduce a large derivative change during the transition: $df([CN_u])/d[CN_u]$ varies from 0 to 10^8 for $[CN_u]_r=10^{-10}$ and $[CN_u]_1=10^{-8}$ (Fig. 14). The maximum increases at about the same order of magnitude as the decrease of $[CN_u]_r$ and $[CN_u]_1$. Even though smoothed, the cutoff is still a sharp change. The smaller the cutoff concentrations, the sharper the transitions. Marching through a steep transition in time usually involves small time steps.

Appendix C: Downregulation of consumption as a function of rates

The contribution of n reactions to the rate component in the residual and Jacobian for species i are

$$\mathbf{R}(i) = \sum_{j=1}^n \nu_{ij} R_j, \quad (C1)$$

and

$$\mathbf{J}(i, k) = \frac{\partial \mathbf{R}(i)}{\partial \mathbf{C}(k)} = \sum_{j=1}^n \frac{\partial (\nu_{ij} R_j)}{\partial \mathbf{C}(k)}, \quad (C2)$$

with \mathbf{C} as the concentration vector, \mathbf{R} as the rate vector, \mathbf{J} as the component of Jacobian matrix (m by m) that is associated with \mathbf{R} , ν_{ij} as stoichiometric coefficient of species i in reaction j ($\sum_{i=1}^m \nu_{ij} \mathbf{C}(i) = 0$), and R_j as the rate of reaction j , a function of the activities (concentration) of the reactants, products, and environmental variables (moisture, temperature, pH, redox, etc.).

C1 Downregulation of consumption for one species

Depending on whether a species a (e.g., NH_4^+) is a reactant ($\nu_{aj} > 0$) or a product ($\nu_{aj} < 0$), we divide the n reactions into a demand (subscript da for demand a) and a supply (subscript sa for supply a) group:

$$R_{da} = \sum_{j=1}^{n_{da}} \nu_{da,aj} R_{da,j}, \quad (C3)$$

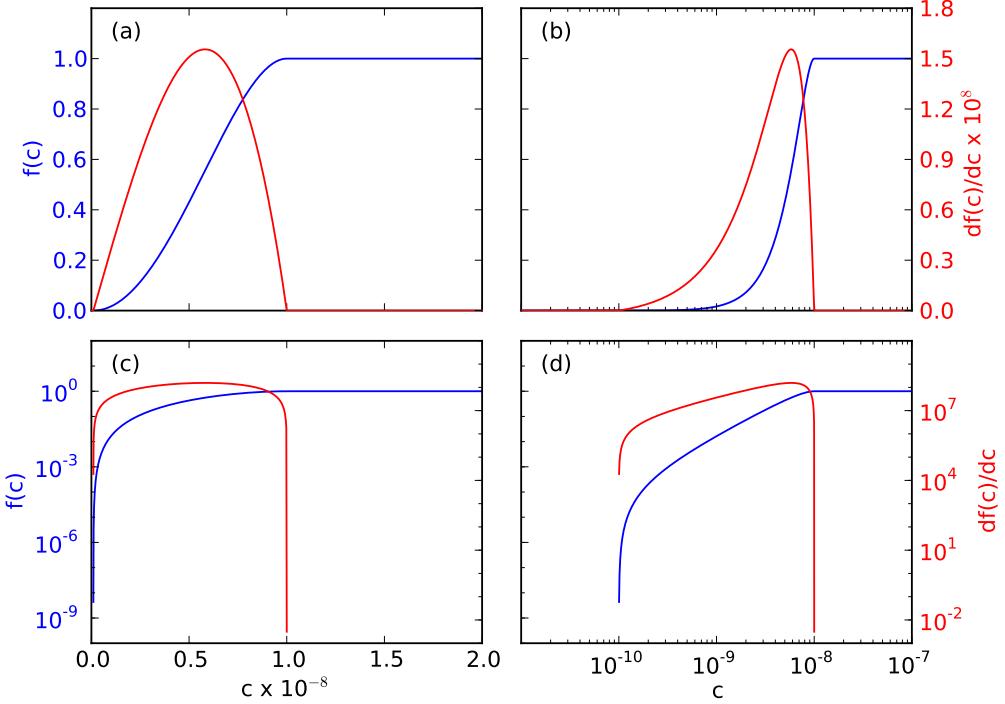


Figure 14. Smoothed cutoff function (Eq. B2) and derivative (Eq. B3). The y axis is in log in (c) and (d), while the x axis is in log in (b) and (d). Even though smoothed, the cutoff is a steep transition. The derivative varies by orders of magnitude in a small concentration range and requires small time step size to march through.

and

$$R_{sa} = \sum_{j=1}^{n_{sa}} \nu_{sa,aj} R_{sa,j}, \quad (C4)$$

with n_{da} as the number of reactions that consume species a , $\nu_{da,aj}$ as the stoichiometric coefficient of species i in a consuming reaction j , n_{sa} as the number of reactions that produce species a , $\nu_{sa,ij}$ as

1025 the stoichiometric coefficient of species i in a producing reaction j , R_{da} as a consumption (demand) rate (mol s^{-1} , negative), and R_{sa} as a production (supply) rate (mol s^{-1} , positive). Therefore,

$$\mathbf{R}(i) = \sum_{j=1}^{n_{sa}} \nu_{sa,aj} R_{sa,j} + \sum_{j=1}^{n_{da}} \nu_{da,aj} R_{da,j} = \mathbf{R}_{sa}(i) + \mathbf{R}_{da}(i), \quad (C5)$$

and

$$\mathbf{J}(i, k) = \frac{\partial \mathbf{R}(i)}{\partial \mathbf{C}(k)} = \sum_{j=1}^n \frac{\partial(\nu_{ij} R_j)}{\partial \mathbf{C}(k)} = \sum_{j=1}^{n_{sa}} \frac{\partial(\nu_{sa,ij} R_{sa,j})}{\partial \mathbf{C}(k)} + \sum_{j=1}^{n_{da}} \frac{\partial(\nu_{da,ij} R_{da,j})}{\partial \mathbf{C}(k)} = \mathbf{J}_{sa} + \mathbf{J}_{da}. \quad (C6)$$

1030 Define a downregulation factor

$$d_a = \min \left(1, -\frac{R_{sa} \Delta t + [\mathbf{C}(a) - \epsilon] V}{R_{da} \Delta t} \right) = \min \left(1, -\frac{s_a}{D_a} \right), \quad (C7)$$

with V as the bulk volume or volume of liquid water of the grid cell for species with concentration unit mol m^{-3} or M. After downregulation,

$$\mathbf{R}_a = \mathbf{R}_{sa} + d_a \mathbf{R}_{da}, \quad (\text{C8})$$

1035

$$\mathbf{J}_a(i, k) = \frac{\partial \mathbf{R}_a(i)}{\partial \mathbf{C}(k)} = \mathbf{J}_{sa}(i, k) + d_a \mathbf{J}_{da}(i, k) + \mathbf{R}_{da}(i) \frac{\partial d_a}{\partial \mathbf{C}(k)}, \quad (\text{C9})$$

$$\frac{\partial d_a}{\partial \mathbf{C}(k)} = - \left(\frac{\partial s_a}{\partial \mathbf{C}(k)} D_a - s_a \frac{\partial D_a}{\partial \mathbf{C}(k)} \right) D_a^{-2}, \quad (\text{C10})$$

1040

$$\frac{\partial s_a}{\partial \mathbf{C}(a)} = \sum_{j=1}^{n_{sa}} \frac{\partial (\nu_{sa,aj} R_{sa,j})}{\partial \mathbf{C}(a)} \Delta t + V, \quad (\text{C11})$$

$$\frac{\partial s_a}{\partial \mathbf{C}(k)} = \sum_{j=1}^{n_{sa}} \frac{\partial (\nu_{sa,kj} R_{sa,j})}{\partial \mathbf{C}(k)} \Delta t, \quad (\text{C12})$$

$$\frac{\partial D_a}{\partial \mathbf{C}(k)} = \sum_{j=1}^{n_{da}} \frac{\partial (\nu_{da,kj} R_{da,j})}{\partial \mathbf{C}(k)} \Delta t, \quad (\text{C13})$$

1045 Implementation in PFLOTTRAN involves 1) adding variables \mathbf{R}_{sa} , \mathbf{R}_{da} , and \mathbf{J}_a , 2) accumulating the values in each reaction rate formula, and 3) conducting downregulation and adding the contribution to the global residual vector and Jacobian matrix.

C2 Downregulation of consumption for a second species

In addition to species a (e.g., NH_4^+), we want to downregulate another species l , for example, NO_3^- .

1050 The treatment is the same except for the reactions that consume species a and produce species l . Suppose we have a nitrification reaction (R10) with rate R_{al} , and $R'_{al} = dR_{al}/d[\text{NH}_4^+]$. The rate and derivative are added in demand rate and derivative for NH_4^+ (\mathbf{R}_{sa} , \mathbf{R}_{da} , \mathbf{J}_{sa} , \mathbf{J}_{da}), not in the sink rate and derivative for NO_3^- (\mathbf{R}_{sl} , \mathbf{R}_{dl} , \mathbf{J}_{sl} , \mathbf{J}_{dl}). Define a downregulation factor

$$d_l = \min \left(1, - \frac{R_{sl} \Delta t + (\mathbf{C}(l) - \epsilon) V_w + R_{al} d_a \Delta t}{R_{dl} \Delta t} \right) = \min(1, - \frac{s_l}{D_l}), \quad (\text{C14})$$

1055

$$\mathbf{R}_l = \mathbf{R}_{sl} + d_l \mathbf{R}_{dl}, \quad (\text{C15})$$

$$\mathbf{J}_l(i, k) = \frac{\partial \mathbf{R}_l(i)}{\partial \mathbf{C}(k)} = \mathbf{J}_{sl}(i, k) + d_l \mathbf{J}_{dl}(i, k) + \mathbf{R}_{dl}(i) \frac{\partial d_l}{\partial \mathbf{C}(k)}, \quad (\text{C16})$$

1060 $\frac{\partial d_l}{\partial \mathbf{C}(k)} = - \left(\frac{\partial s_l}{\partial \mathbf{C}(k)} D_l - s_l \frac{\partial D_l}{\partial \mathbf{C}(k)} \right) D_l^{-2},$ (C17)

$$\frac{\partial s_l}{\partial \mathbf{C}(l)} = \sum_{j=1}^{n_{sl}} \frac{\partial (\nu_{sl,lj} R_{sl,j})}{\partial \mathbf{C}(l)} \Delta t + V_w + \underline{\frac{\partial (R_{al} d_a)}{\partial \mathbf{C}(l)}},$$
 (C18)

$$\frac{\partial s_l}{\partial \mathbf{C}(k)} = \sum_{j=1}^{n_{sl}} \frac{\partial (\nu_{sl,kj} R_{sl,j})}{\partial \mathbf{C}(k)} \Delta t + \underline{\frac{\partial (R_{al} d_a)}{\partial \mathbf{C}(K)}},$$
 (C19)

1065 and

$$\frac{\partial D_a}{\partial \mathbf{C}(k)} = \sum_{j=1}^{n_{da}} \frac{\partial (\nu_{da,kj} R_{da,j})}{\partial \mathbf{C}(k)} \Delta t.$$
 (C20)

In addition to variables \mathbf{R}_{sl} , \mathbf{R}_{dl} , and \mathbf{J}_l , downregulating the second species requires two additional variables, R_{al} , and R'_{al} . When accumulating the values in each reaction rate formula, the production rate from the reaction that consumes the first species has to be carefully treated as it is downregulated for the first species. Conducting downregulation involves additional terms, as underlined in the equations. Compared with the downregulation for the first species, downregulating the second species that can be produced from the first species with a simple $A \rightarrow B$ reaction becomes much more complicated. If we add another reaction that consumes the second species to produce the first species, this approach has to be modified.

1075 In general, many if not all species need to be downregulated. Extending this approach involves adding many variables (vectors and matrices) to track the consumption and production rates and their derivatives. The consumption and production relationship among many species can be complicated, making generalization of this approach challenging if not infeasible.

Appendix D: A semi-analytical solution for Monod equation

1080 As an analytical solution is not available, discretizing Eq. (22) using the backward Euler method, it becomes

$$[\text{NH}_4^+]^{k+1} + (k_m + R_a \Delta t - [\text{NH}_4^+]^k) [\text{NH}_4^+]^{k+1} - k_m [\text{NH}_4^+]^k = 0.$$
 (D1)

The two roots are

$$[\text{NH}_4^+]^{k+1} = 0.5 \left[[\text{NH}_4^+]^k - k_m - R_a \Delta t \pm \sqrt{([\text{NH}_4^+]^k - k_m - R_a \Delta t)^2 + 4k_m [\text{NH}_4^+]^k} \right].$$
 (D2)

1085 Appendix E: Matrix equation for example Test 2

$$\begin{bmatrix} \frac{1}{\Delta t} + J_{at} + J_{nitr} & 0 & 0 \\ -J_{at} & \frac{1}{\Delta t} & 0 \\ -J_{nitr} & 0 & \frac{1}{\Delta t} \end{bmatrix} \begin{pmatrix} \delta[\text{NH}_4^+]^{k+1,1} \\ \delta[\text{PlantA}]^{k+1,1} \\ \delta[\text{NO}_3^-]^{k+1,1} \end{pmatrix} = \begin{pmatrix} R_{at} + R_{nitr} \\ -R_{at} \\ -R_{nitr} \end{pmatrix}.$$
 (E1)

Appendix F: Matrix equation for example Test 3

$$\begin{bmatrix} \frac{1}{\Delta t} + J_{at} + J_{nitr} & 0 & 0 & 0 & 0 \\ -J_{at} & \frac{1}{\Delta t} & 0 & 0 & 0 \\ -J_{nitr} + J_{nt,a} & 0 & \frac{1}{\Delta t} + J_{nt} + J_{deni} & 0 & 0 \\ -J_{nt,a} & 0 & -J_{nt,n} & 1/\Delta t & 0 \\ 0 & 0 & -0.5J_{deni} & 0 & 1/\Delta t \end{bmatrix} \begin{pmatrix} \delta[\text{NH}_4^+]^{k+1,1} \\ \delta[\text{PlantA}]^{k+1,1} \\ \delta[\text{NO}_3^-]^{k+1,1} \\ \delta[\text{PlantN}]^{k+1,1} \\ \delta[\text{N}_2]^{k+1,1} \end{pmatrix} = \begin{pmatrix} R_{at} + R_{nitr} \\ -R_{at} \\ -R_{nitr} + R_{nt} + R_{deni} \\ -R_{nt} \\ -0.5R_{deni} \end{pmatrix}. \quad (\text{F1})$$

Appendix G: Supplemental figures

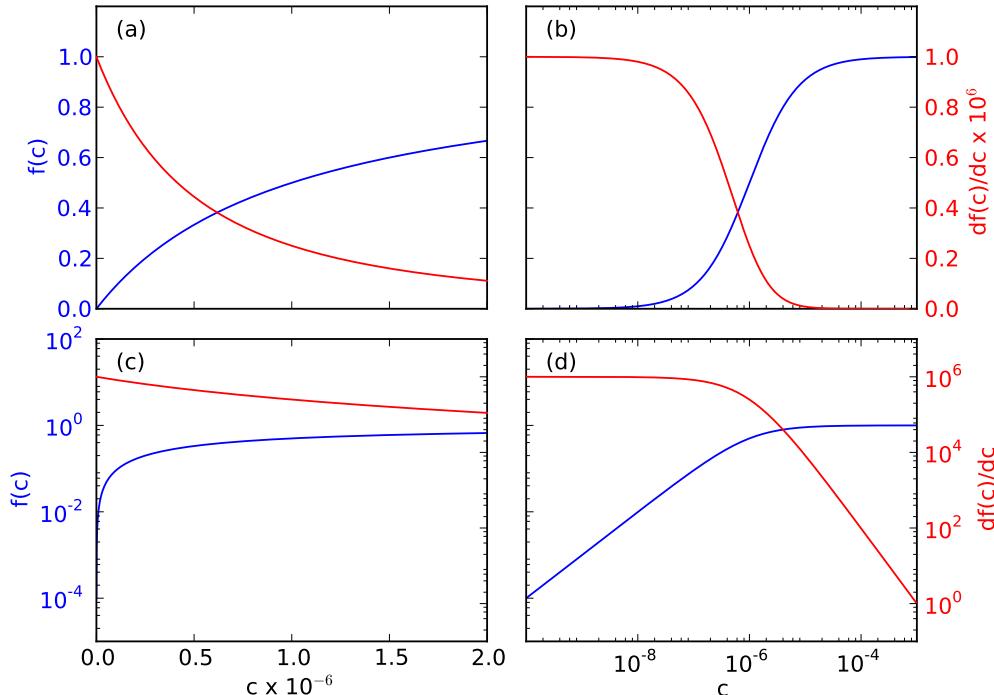


Figure 15. Monod substrate limiting function (Eq. 18) and derivative. The y axis is in log in (c) and (d), while the x axis is in log in (b) and (d). As the function switches from zero-order to first-order, the derivative jumps six (\$k_m^{-1}\$) orders of magnitudes, requiring small time step size to step through.

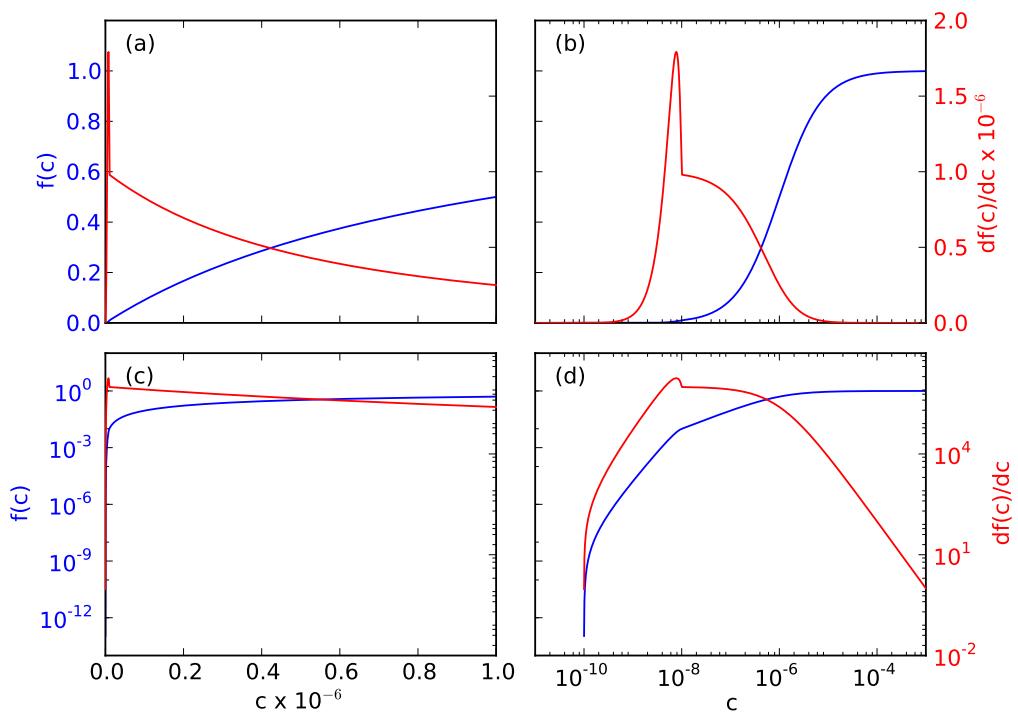


Figure 16. A combination of the Monod substrate limiting function (Eq. 18, Fig. 15) and the smoothed cutoff function (Eq. B2, Fig. 14) introduces steep transitions that require small time step sizes to march through.