

Manuscript prepared for Geosci. Model Dev.
with version 2014/09/16 7.15 Copernicus papers of the L^AT_EX class copernicus.cls.
Date: 11 October 2015

Using Reactive Transport Codes to Provide Mechanistic Biogeochemistry Representations in Global Land Surface Models

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
5 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. We explore a flexible and extensible approach to biogeochemistry in land surface models to facilitate testing of alternative models and incorporation of new understanding: coupling to a configurable subsurface reactive transport code. A reaction network with the CLM-CN decomposition, nitrification, denitrification, and plant uptake is used as an example. We implement the reactions in the open source PFLOTRAN code, couple to CLM, and test at arctic, temperate, and tropical sites. To make the reaction network designed for use in explicit time-stepping in CLM compatible with implicit time-stepping, the Monod substrate rate-limiting function with a residual concentration is used to represent the limitation of nitrogen availability on plant uptake and immobilization. To achieve accurate, efficient, and robust numerical solutions, care needs to be taken to using scaling, clipping, or log transformation to avoid negative concentrations during the Newton iterations. With a tight relative update tolerance (STOL, e.g., 10^{-12}) to avoid false convergence, accurate solution can be achieved with about 50% more computing time than CLM in point mode site simulations using either the scaling or the clipping methods. Log transformation takes 60% to 100% more computing time than CLM. The computing time increases slightly for clipping and scaling, substantially for log transformation for half saturation decrease from 10^{-3} to $10^{-9} \text{ mol m}^{-3}$, which normally results in decreasing nitrogen concentrations. Frequent occurrence of very low concentrations (e.g. below nM) can increase the computing time for clipping by about 20%, and double for log transformation. Caution needs to be taken for the scaling method as a small scaling factor caused by a positive update to a small concentration may diminish the update and result in false convergence even with very tight STOL. As methane and nitrous oxide production and consumption involve very low half saturation and threshold concentrations, this work provides insights for addressing nonnegativity issue and facilitates mechanistic biogeochemistry representation in earth system models to reduce climate prediction uncertainty.

1 Introduction

Land surface (terrestrial ecosystem) models (LSM) 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 (Sellers et al., 1997). Evolving from the first generation 35 “bucket”, second generation “biophysical”, and third generation “physiological” models (Seneviratne et al., 2010), current LSMs, e.g., the Community Land Model (CLM), implement comprehensive thermal, hydrological, and biogeochemical processes (Oleson et al., 2013). The important role of soil biogeochemistry is suggested by the confirmation that the increase of carbon dioxide (CO₂), methane (CH₄), and nitrous oxide (N₂O) in the atmosphere since the preindustrial time is 40 the main driving cause of climate change, and interdependent water, carbon and nitrogen cycles in terrestrial ecosystems are very sensitive to climate changes (IPCC, 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 45 for improving climate prediction (e.g., Riley et al., 2014).

As LSMs usually hardcode the reaction network (pools/species, reactions, rate formulae), substantial effort is often required to modify the source code for testing alternative biogeochemical models, and incorporating new process understanding. Fang et al. (2013) demonstrated use of a reaction-based approach to facilitate implementation of CLM-CN and CENTURY models, and incorporation 50 of phosphorus cycle. Tang et al. (2013) solved the advection diffusion equation in CLM using operator splitting. In contrast, TOUGH-REACT, a reactive transport modeling (RTM) code, was used to develop multi-phase mechanistic carbon and nitrogen models with many speciation and microbial reactions (Maggi et al., 2008; Gu and Riley, 2010; Riley et al., 2014), but has not been coupled to a 55 LSM. PHREEQC was coupled with DayCent to describe soil and stream water equilibrium chemistry (Hartman et al., 2007). Coupling a RTM code with CLM will facilitate testing of increasingly mechanistic biogeochemical models in LSMs.

An essential aspect of LSMs is to simulate competition for nutrients (e.g., mineral nitrogen, phosphorus, etc.) among plants and microbes. In CLM, plant and immobilization nitrogen demands are calculated independent of soil mineral nitrogen. The limitation of nitrogen availability on plant uptake and immobilization is simulated by a demand based competition: demands are downregulated 60 by soil nitrogen concentration (Oleson et al., 2013; Thornton and Rosenbloom, 2005). This avoids negative concentrations and does not introduce numerical errors (Tang and Riley, 2015) as CLM uses explicit time stepping.

RTM model often accounts for limitation of reactant availability on reaction rates for each individual reaction for mechanistic representations and flexibility in adding reactions. RTM codes 65 generally use implicit time stepping and the Newton-Raphson method. Negative concentration can be introduced during iterations, which is not physical, and can cause numerical instability and errors (Shampine et al., 2005). This is expected to worsen when we implement microbial reactions for

methane and nitrous oxide consumption and production as the threshold and half saturation are at or below nM (10^{-9} M) level (Conrad, 1996). The redox potential Eh needs to be decreased to -0.35
70 V (oxygen concentration $< 10^{-22}$ M Hungate, 1975) for methanogens to grow (Jarrell, 1985).

Three 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). The other two either scale back the update vector (Bethke, 2007; Hammond, 2003) or clip the concentrations for the species that are going negative (Yeh et al., 2004; White and McGrail,
75 2005; Sonnenthal et al., 2014) in each iteration. Except that log transformation is more computationally demanding (Hammond, 2003), how the various algorithms for enforcing nonnegativity affect computational accuracy and efficiency is rarely discussed.

As LSMs need to run under varies conditions at the globe scale for simulation duration of centuries, it is necessary to resolve accuracy and efficiency issues to use RTM codes for LSMs. The
80 objective of this work is to explore some of the implementation issues associated with using RTM codes in LSMs, with the ultimate goal being accurate, efficient, robust, and configurable representations of subsurface biogeochemical reactions in CLM. To this end, we develop an alternative implementation of an existing CLM biogeochemical reaction network using PFLOTRAN, coupled that model to CLM, test the implementation at arctic, temperate, and tropical sites, and examine the
85 implication of using scaling, clipping, and log transformation for avoiding negative concentrations. Although we focus on a carbon-nitrogen decomposition cascade with nitrification, denitrification, and plant nitrogen uptake implemented in PFLOTRAN and CLM, the critical numerical issue of avoiding negative concentrations has broader relevance as biogeochemical representations in LSMs become more mechanistic.

90 2 Methods

LSMs generally include biogeochemical reactions for carbon and nitrogen cycles, in particular, the organic matter decomposition, nitrification, denitrification, plant nitrogen uptake, and methane production and oxidation. The kinetics are usually described by a first-order rate modified by response functions for environmental variables (temperature, moisture, pH, etc.) (Bonan et al., 2012; Boyer
95 et al., 2006; Schmidt et al., 2011). In this work, we use the CLM-CN 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 (Fig. 1) as an example. The reactions and rate formulae are detailed in Appendix A.

2.1 CLM-PFLOTRAN biogeochemistry coupling

100 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

subsurface. This work focuses on the biogeochemistry as CLM solves the energy and water flow equations and handles the solute transport (mixing, advection, diffusion, and leaching). In each CLM time step, CLM provides production rates for Lit1C, Lit1N, Lit2C, Lit2N, Lit3C, Lit3N for litter fall; CWDC, and CWDN for coarse woody debris production, ammonium and nitrate for nitrogen deposition and fixation; and plant nitrogen demand; 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 reactions, and provides the final concentrations back to CLM.

The reactions and rates 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, and other environmental variables (see `reaction_clm.F90` in `pflotran-dev` source code for details). PFLOTRAN accumulates these rates 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 (Appendix B).

Unlike the explicit time stepping in CLM where only reaction rates need to be calculated, implicit time stepping requires derivatives. While PFLOTRAN provides an option to calculate derivatives numerically, analytical derivative calculation is generally preferred (e.g., Xu et al., 2006) because numerical calculation for accurate Jacobian is a notoriously difficult task (Shampine et al., 2005). PETSc, the parallel framework that PFLOTRAN is built upon, suggests that finite difference approximation is intended for situations where one is unable to compute Jacobian analytically and for checking hand-coded analytical Jacobian calculation (Balay et al., 2015).

Many reactions can be specified in an input file, providing flexibility in adding various reactions with user-defined rate formulae. As typical rate formulae consist of first order, Monod, and inhibition terms, a general rate formula 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 an input file, with flexible number of reactions, 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 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.

2.2 Mechanistic representation of rate-limiting processes

To use RTMs in LSMs, we need to make reaction networks designed for use in explicit time-stepping LSMs compatible with implicit time-stepping RTMs. The limitation of reactant availability on reaction rate is well represented by the first-order rate (Eqs. A1, A2, A5, A7): the rate decreases to

zero as the concentration decreases to zero. A residual concentration is often added to represent a threshold below which a reaction stops. For example the decomposition rate, Eq. (A1) becomes

$$140 \quad \frac{d[\text{CN}_u]}{dt} = -k_d f_T f_w ([\text{CN}_u] - [\text{CN}_u]_r). \quad (1)$$

Where CN_u is the upstream pool with 1:u as the CN ratio in mole; [] denotes concentration; k_d , f_T , and f_w is the rate coefficient, temperature and moisture response function. When CN_u goes below $[\text{CN}_u]_r$ in an iteration, Eq. (1) implies a hypothetical reverse reaction to bring it back to $[\text{CN}_u]_r$. The residual concentration can be set to zero to nullify the effect.

145 For the litter decomposition reactions (R7, R8, R9) that immobilize nitrogen (N), the nitrification reaction (R11) associated with decomposition to produce nitrous oxide, and the plant nitrogen uptake reactions (R13,R14), the rate formulae do not account for the limitation of the reaction rate by nitrogen availability. Mechanistically, a nitrogen limiting function needs to be added. For example, using the widely used Monod substrate limitation function (Hammond, 2003), Eq. (1) becomes

$$150 \quad \frac{d[\text{CN}_u]}{dt} = -k_d f_T f_w ([\text{CN}_u] - [\text{CN}_u]_r) \frac{[N] - [N]_r}{[N] - [N]_r + k_m}, \quad (2)$$

with half saturation k_m and a mineral nitrogen residual concentration $[N]_r$. In the case of $[N] - [N]_r = k_m$, the rate is decreased by half. For $[N] \gg k_m$, Eq. (2) approximates zero order with respect to $[N]$. For $[N] \ll k_m$, Eq. (2) approximates first order with respect to $[N]$. The derivative of the Monod term, $k_N([N] + k_m)^{-2}$, increases to about k_m^{-1} as the concentration decreases to below k_m . This represents 155 a steep transition when k_m is small. The half saturation is expected to be greater than the residual concentration. When both are zero, the rate is not limited by the substrate availability.

To separate mineral nitrogen into ammonium (NH_4^+) and nitrate (NO_3^-), we need to distribute the demands between ammonium and nitrate for plant uptake and immobilization. If we simulate the ammonium limitation on plant uptake with

$$160 \quad R_a = R_p \frac{[\text{NH}_4^+]}{[\text{NH}_4^+] + k_m}, \quad (3)$$

the plant nitrate uptake can be represented by

$$R_n = (R_p - R_a) \frac{[\text{NO}_3^-]}{[\text{NO}_3^-] + k_m} = R_p \frac{k_m}{[\text{NH}_4^+] + k_m} \frac{[\text{NO}_3^-]}{[\text{NO}_3^-] + k_m}. \quad (4)$$

Where R_p , R_a , and R_n are the plant uptake rate for nitrogen, ammonium, and nitrate. This essentially assumes an inhibition of ammonium on nitrate uptake, which is consistent with the observation that 165 plant nitrate uptake rate remained low until ammonium concentrations dropped below a threshold (Eltrop and Marschner, 1996). However, the preference differs for different plants (Pfautsch et al., 2009; Warren and Adams, 2007; Nordin et al., 2001; Falkengren-Grerup, 1995; Gherardi et al., 2013), which require different representations in future developments.

CLM uses a demand based competition approach (Appendix A5) to represent the limitation of 170 available nitrogen on plant uptake and immobilization. It is similar to the Monod function except that

it introduces a discontinuity during the transition between the zero and first order rate. Implementation of the demand based competition in a RTM involves separating the supply and consumption rates for each species in each reaction, and conducting downregulation when necessary after contributions from all of the reactions are accumulated. It involves not only the rate terms for the residual
175 but also the derivative terms for the Jacobian. The complexity increases quickly when more species need to be downregulated (e.g., ammonium, nitrate, and organic nitrogen) and there are transformation processes among these species. 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. In contrast, use of Monod function with a residual concentration for
180 individual reactions is easier to implement, and allows more flexibility in adding reactions..

2.3 Scaling, clipping and log transformation for avoiding negative concentration

The concentration update for iteration p from time step k to $k + 1$, $\delta\mathbf{c}^{k+1,p}$, in a Newton-Raphson iteration can be greater than the concentration $\mathbf{c}^{k+1,p}$ in some entries (Eq. B6), which can result in nonphysical negative concentrations. One approach to avoid negative concentration is to scale back
185 the update with a scaling factor λ (Bethke, 2007; Hammond, 2003) such that

$$\mathbf{c}^{k+1,p+1} = \mathbf{c}^{k+1,p} - \lambda\delta\mathbf{c}^{k+1,p} > 0, \quad (5)$$

where

$$\lambda = \min_{i=0,m} [1, \alpha\mathbf{c}^{k+1,p}(i)/\delta\mathbf{c}^{k+1,p}(i)] \quad (6)$$

for positive $\delta\mathbf{c}^{k+1,p}(i)$ with m as the number of species times the number of numerical grid cells.
190 RTM codes STOMP, HYDROGEOCHEM 5.0, and TOUGH-REACT using clipping, i.e., for any $\delta\mathbf{c}^{k+1,p}(i) \geq \mathbf{c}^{k+1,p}(i)$, $\delta\mathbf{c}^{k+1,p}(i) = \mathbf{c}^{k+1,p}(i) - \epsilon$ with ϵ as a small number (e.g., 10^{-20}).

Log transformation also ensures positive solution (Bethke, 2007; Hammond, 2003; Parkhurst and Appelo, 1999). It 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
195 state changes without the need to use variable switching. Instead of solving Eq. (B3) for \mathbf{c}^{k+1} using Eqs. (B4,B5,B6), this method solves for $(\ln \mathbf{c}^{k+1})$ (Hammond, 2003) with

$$\mathbf{J}_{\ln}(i,j) = \frac{\partial \mathbf{f}(i)}{\partial \ln(\mathbf{c}(j))} = \mathbf{c}(j) \frac{\partial \mathbf{f}(i)}{\partial \mathbf{c}(j)} = \mathbf{c}(j) \mathbf{J}(i,j), \quad (7)$$

$$\delta \ln \mathbf{c}^{k+1,p} = \mathbf{J}_{\ln}^{-1} \mathbf{f}(\mathbf{c}^{k+1,p}), \quad (8)$$

200 and

$$\mathbf{c}^{k+1,p+1} = \mathbf{c}^{k+1,p} \exp[-\delta \ln(\mathbf{c}^{k+1,p})]. \quad (9)$$

3 Tests, results, and discussions

The Newton-Raphson method and scaling, clipping, and log transformation are widely used and extensively tested for RTM, but not for LSM simulations. CLM describes biogeochemical dynamics
205 within daily cycles for simulation durations of hundreds of years; the nitrogen concentration can be very low (mM to nM) while the carbon concentrations can be very high (e.g., thousands mol m⁻³ carbon in organic layer); the concentrations and dynamics can vary dramatically in different locations around the globe. It is not surprising that the complex biogeochemical dynamics in a wide range of temporal and spatial scales in CLM poses numerical challenges for the RTM methods. Our
210 simulations reveal some numerical issues (numerical errors, divergence, and small time step sizes) that were not widely reported. We identify the issues from coupled simulations, and reproduce them in simple test problems. We examine remedies in the simple test problems, and test them in the coupled simulations. For convenience of presentation in this paper, we examine the reaction network and illustrate the numerical issues and remedies using simple test problems, and then test them in
215 the coupled simulations.

For Test 1, we start with plant ammonium uptake to examine numerical solution for Monod function, and then add nitrification and denitrification incrementally to assess the implication of adding reactions. For Test 2, We check the implementation of mineralization and immobilization in the decomposition reactions. Thirdly, we compare the nitrogen demand partition into ammonium and
220 nitrate between CLM and PFLOTRAN. With coupled CLM-PFLOTRAN spinup simulations for arctic, temperate, and tropical sites, we assess the application of scaling, clipping and log transformation to achieve accurate, efficient, and robust simulations. Spreadsheet and PFLOTRAN input files are provided as supplemental information (SI).

Our implementation of CLM biogeochemistry introduces mainly two parameters: half saturation
225 k_m and residual concentration. A wide range of k_m values were reported for ammonium, nitrate, and organic nitrogen for microbes and plants. The median, mean, and standard deviations range from 10~100, 50~500, and 10~200 μM , respectively (Kuzyakov and Xu, 2013). Reported residual concentrations are limited and are considered to be zero (e.g., HØGh-Jensen et al., 1997), likely because of the detection limits of the analytical methods. The detection limits are usually at μM level,
230 while up to 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 ammonium and nitrate for microbes. We start with $k_m = 10^{-6}$ 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., H₂ and O₂) where the
235 concentrations can be much lower, we examine k_m from 10^{-3} to 10^{-9} in our test problems. The k_m is expected to be different for different plants, microbes, and for ammonium and nitrate. We do not differentiate them in this work as we focus on numerical issues.

3.1 Plant nitrogen uptake, nitrification, and denitrification

It was observed that plants can decrease nitrogen concentration to below detection limit in hours
240 (Kamer et al., 2001). In CLM, the total plant nitrogen demand is calculated based on photosynthetized 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 et al., 2013). The demand is provided as an input to PFLO-TRAN. We use the Monod function to represent the limitation of nitrogen availability on uptake.
245 We examine the numerical solutions for the Monod equation. Incrementally, we add first order reactions (e.g., nitrification, denitrification, and plant nitrate uptake) to look into the numerical issues in increasingly complex reaction networks.

3.1.1 Plant ammonium uptake (Test 1)

We consider the plant ammonium uptake reaction (R13) with a rate R_a

$$250 \frac{d[\text{NH}_4^+]}{dt} = -R_a \frac{[\text{NH}_4^+]}{[\text{NH}_4^+] + k_m} = -R_{at}. \quad (10)$$

Discretizing it in time using the backward Euler method for a time step size Δt , a solution is

$$[\text{NH}_4^+]^{k+1} = 0.5 \left[[\text{NH}_4^+]^k - k_m - R_a \Delta t \pm \sqrt{([\text{NH}_4^+]^k - k_m - R_a \Delta t)^2 + 4k_m [\text{NH}_4^+]^k} \right]. \quad (11)$$

Ignoring the negative root, $[\text{NH}_4^+]^{k+1} \geq 0$. Adding a residual concentration by replacing $[\text{NH}_4^+]$ with $[\text{NH}_4^+] - [\text{NH}_4^+]_r$, $[\text{NH}_4^+] \geq [\text{NH}_4^+]_r$. Namely, the representation of plant ammonium uptake
255 with the Monod function mathematically ensures $[\text{NH}_4^+]^{k+1} \geq [\text{NH}_4^+]_r$.

We use spreadsheet to examine the Newton-Raphson iteration process for solving Eq. (10) and the application of clipping, scaling, and log transformation (SI test1.xlsx). When an overshoot gets the concentration closer to the negative than the positive root (Eq. 11), the iterations converge to the nonphysical negative semi-analytical solution (sheet case3). This can be avoided by using clipping,
260 scaling, or log transformation (sheet case4, case8, case9).

While clipping avoids convergence to the negative solution, the fact that the ammonium consumption is clipped but the PlantA production is not (sheet case5) requires additional iterations to eliminate the mass balance error. This is different from Tang and Riley (2015) in that the explicit method does not have subsequent iterations to correct mass balance error. However, if a nonreactive
265 species is added with a concentration of 1000 mol m^{-3} (e.g., SOM4 in organic layer), the relative update $r\text{tol} = \|\delta c^{k+1,p}\|_2 / \|c^{k+1,p}\|_2$ decreases to 1.9×10^{-9} in the iteration (sheet case5a). With $\text{STOL} = 10^{-8}$, the iteration would be deemed converged with the mass balance error. Satisfying the relative update tolerance criteria does not guarantee that the residual equations are satisfied (Lichtner et al., 2015). For this reason, we need to a tight STOL to avoid this false convergence so that additional
270 iterations can be taken to resolve the mass balance error.

In contrast to clipping, scaling applies the same scaling factor to dampen both ammonium consumption and PlantA production following the stoichiometry of the reaction to maintain mass balance (sheet case6). However, if we add a production reaction that is independent of plant ammonium uptake, say nitrate deposition, then the nitrate increases due to deposition will be damped by the
275 same scaling factor, introducing numerical error in the iteration (sheet case7). Like clipping, the error can be eliminated in the subsequent iterations, and a loose STOL may lead to false convergence.

Small to zero concentration for ammonium and PlantA has no impact on the iterations for the clipping or scaling methods in this test. In contrast, a small initial PlantA concentration can cause challenge for log transformation even though PlantA is only a product. When it is zero, the Jacobian
280 matrix is singular because zero is multiplied to the column corresponding to PlantA (Eq. 7). An initial PlantA concentration of 10^{-9} can result in overflow of the exponential function (sheet casea, as a 64-bit real number, which corresponds to double precision, is precise to 15 significant digits and has a range of e^{-709} to e^{709} , Lemmon and Schafer,2005). Clipping the update (say to be between -5 and 5) is needed to prevent numerical overflow or excessively large update. Like the cases with clip-
285 ping without log transformation, mass balance error is introduced, and can be resolved in subsequent iterations (sheet caseb).

This simple test for the Monod function indicates that 1) Newton-Raphson iterations may converge to a negative concentration; 2) scaling, clipping, and log transformation can be used to avoid convergence to negative concentration; 3) small or zero concentration makes the Jacobian matrix stiff
290 or singular when log transformation is used, and clipping is needed to guard against overflow of the exponential function; 4) mass balance error is introduced in the iterations when clipping is applied as it dampens only the consumption, but not the corresponding production; 5) independent reactions are numerically inhibited in the iterations when scaling is applied; 6) the errors due to clipping or scaling can be eliminated with subsequent iterations; and 7) loose update tolerance convergence
295 criteria may cause false convergence.

3.1.2 Plant ammonium uptake and nitrification (Test 2)

Adding a nitrification reaction (R10) with a first-order rate, Eq. (10) becomes

$$\frac{d[\text{NH}_4^+]}{dt} = -R_a \frac{[\text{NH}_4^+]}{[\text{NH}_4^+] + k_m} - k_{nitr} [\text{NH}_4^+] = -R_{at} - R_{nitr}. \quad (12)$$

A semi-analytical solution similar to Eq. (11) can be derived. With $J_{at} = dR_{at}/d[\text{NH}_4^+] = R_a k_m ([\text{NH}_4^+] + k_m)^{-2}$, and $J_{nitr} = dR_{nitr}/d[\text{NH}_4^+] = k_{nitr}$, the matrix equation Eq. (B5), becomes
300

$$\begin{bmatrix} 1/\Delta t + J_{at} + J_{nitr} & 0 & 0 \\ -J_{at} & 1/\Delta t & 0 \\ -J_{nitr} & 0 & 1/\Delta t \end{bmatrix} \begin{pmatrix} \delta[\text{NH}_4^+]^{k+1,1} \\ \delta[\text{PlantA}]^{k+1,1} \\ \delta[\text{NO}_3^-]^{k+1,1} \end{pmatrix} = \begin{pmatrix} R_{at} + R_{nitr} \\ -R_{at} \\ -R_{nitr} \end{pmatrix}, \quad (13)$$

for the first iteration. As $R_{at} + R_{nitr} \geq 0$, the ammonium update is positive even when ammonium concentration is not very low. The off-diagonal terms for PlantA and nitrate in the Jacobian matrix

bring the positive ammonium update into the updates for PlantA and nitrate even though there is no
 305 reaction that consumes them. Specifically,

$$\frac{\delta[\text{PlantA}]^{k+1,1}}{\Delta t} = -\frac{\frac{1}{\Delta t} + J_{nitr}}{\frac{1}{\Delta t} + J_{at} + J_{nitr}} R_{at} + \frac{J_{at}}{\frac{1}{\Delta t} + J_{at} + J_{nitr}} R_{nitr}, \quad (14)$$

$$\frac{\delta[\text{NO}_3^-]^{k+1,1}}{\Delta t} = \frac{J_{nitr}}{\frac{1}{\Delta t} + J_{at} + J_{nitr}} R_{at} - \frac{\frac{1}{\Delta t} + J_{at}}{\frac{1}{\Delta t} + J_{at} + J_{nitr}} R_{nitr}. \quad (15)$$

Depending on the rates (R_{at} , R_{nitr}), derivatives (J_{at} , J_{nitr}), and time step size Δt , the update can
 310 be positive for PlantA and nitrate, producing a zero order "numerical consumption" for PlantA or
 nitrate, in which the limitation of PlantA or nitrate availability is not explicitly represented. This
 has important implications for the scaling method.

The scaling factor (λ) is not only a function of the update, but also of the concentration (Eq. 6). If a
 positive update is produced for a zero concentration, the scaling factor is zero, decreasing the scaled
 315 update to zero. The iteration converges without any change to the concentrations, numerically stop
 all of the reactions in the time step unless STOL is negative. We add the denitrification reaction with
 $R_{nitr} = 10^{-6} \text{ s}^{-1}$ to SI test1.xlsx case6 to create SI test2.xlsx to demonstrate that a small enough
 initial concentration relative to the positive update may numerically inhibit all of the reactions as
 well. An update of $6.6 \times 10^{-6} \text{ M}$ is produced for nitrate (sheet scale1). When the initial nitrate
 320 concentration $[\text{NO}_3^-]_0$ is not too small, say 10^{-6} M , the solution converges to the semi-analytical
 solution in six iterations. With decreasing $[\text{NO}_3^-]_0$ to 10^{-9} M , the relative update stol is $9.2 \times$
 10^{-10} . If $\text{STOL} = 10^{-9}$, the solution is deemed converged as Eq. (B9) is met, but not to the semi-
 analytical solution. The ammonium uptake and nitrification reactions are numerically "inhibited"
 because the small scaling factor and a high concentration of a nonreactive species decreases the
 325 update to below STOL to reach false convergence. If we tighten STOL to 10^{-30} , the iterations
 continues, with decreasing nitrate concentration, λ , and stol by two orders of magnitude ($1-\alpha$ as
 default $\alpha = 0.99$) in each iteration, until stol reaches 10^{-30} (sheet scale2). Unless the STOL is
 sufficient small, or MAXIT is small (Appendix B), false convergence will occur before the time
 stepping routine continues with time step cut for the scaling method. The impact of "numerical
 330 consumption" on clipping and log transformation is much less dramatic than the scaling method
 as the latter applies the same scaling factor to the whole update vector following stoichiometric
 relationships of the reactions to maintain mass balance, and the limiting concentration decreases by
 $(1 - \alpha)$ times in each iteration, with the possibility of resulting in less than STOL relative update in
 MAXIT iterations.

335 In summary, this test problem demonstrates that 1) positive update can be produced even for
 products during Newton-Raphson iteration; 2) when a positive update is produced for a very low
 concentration, a very small scaling factor may numerically inhibits all of the reactions even with
 very tight STOL.

3.1.3 Plant uptake, nitrification, and denitrification (Test 3)

340 The matrix and update equations with added plant nitrate uptake and denitrification are available
in Appendix C. In addition to nitrate and PlantA, PlantN and the denitrification product nitrogen
gas may have positive updates. In addition to the off-diagonal terms due to the derivative of plant
uptake with respect to ammonium concentration, the derivative of plant uptake with respect to ni-
trate concentration is added in the Jacobian matrix for PlantN (Eq. C1). As a result, positive update
345 for both ammonium and nitrate will contribute to positive PlantN update through the two nonzero
off-diagonal terms. Therefore, the likelihood for a positive update to PlantN is greater than PlantA
as the former are influenced by more rates and derivatives. We add plant nitrate uptake, and deni-
trification into SI test2.xlsx and assess the implications of increased reactions and complexity in SI
test3.xlsx. In addition to nitrate, this introduces a positive update for nitrogen gas in the first iteration
350 (sheet scale1). As the iterations resolve the balance between nitrite production from nitrification,
and consumption due to plant uptake and denitrification, update to PlantN becomes positive, and
eventually leads to false convergence. The time step size needs to be decreased from 1800 to 15 s
to resolve the false convergence (sheet scale2). In contrast, the added reactions have less impact on
clipping and log transformation.

355 **3.2 Nitrogen immobilization and mineralization during decomposition (Test 4)**

We examine another part of the reaction network: decomposition, nitrogen immobilization, and min-
eralization (Fig. 1). We consider a case of decomposing $0.2 \text{ M Lit1C} + 0.005 \text{ M Lit1N}$ to produce
SOM1 with initial $4 \mu\text{M}$ ammonium using the reactions (R2 and R7) in the CLM-CN reaction net-
work (Fig. 1). We use PFLOTRAN with a saturated grid cell with porosity of 0.25. At the beginning,
360 Lit1 decreases and SOM1 increases sharply because the rate coefficient for Lit1 is 16 times that
for SOM1 ((Figs. 3a,b)). As ammonium concentration decreases by orders of magnitude because of
the faster immobilization than mineralization rate (Fig. 3c,d), Lit1 decomposition rate slows down
to the level such that the immobilization rate is less than the mineralization rate. Namely, SOM1
decomposition controls Lit1 decomposition through limitation of mineralization on immobilization.
365 As the immobilization rate decreases with decreasing Lit1, ammonium concentration rebounds after
Lit1 is depleted. For k_m of 10^{-6} , 10^{-9} , and 10^{-12} M , Lit1 and SOM1 dynamics are similar ex-
cept slight difference in the early transit periods, but the ammonium values are decreased to $\sim 10^{-8}$,
 10^{-11} , and 10^{-14} M , respectively. Smaller k_m results in lower ammonium concentration, which has
implications for the clipping, scaling and log transformation methods.

370 **3.3 Nitrogen demand partitioning between ammonium and nitrate**

For comparison with CLM, we examine the uptake rate as a function of demands and available con-
centrations $f_{pi} = (R_a + R_n)/R_p$ as implemented in Eqs. (3,4). 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. 375 2). The new representation (Eqs. 3, 4) is generally similar, with $f_{pi} = 1$ or 0 when $[\text{NH}_4^+]$ or $[\text{NO}_3^-]$ \gg 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 $k_m = 10^{-10}$. Various level of preferences of ammonium over nitrate 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; Veuger et al., 2004). CLM implies a strong preference for ammonium over nitrate. For example, if ammonium is abundantly sufficient, nitrate will not be taken. The new scheme allows the level of preference to be adjusted by varying k_m .

385 3.4 CLM-PFLOTRAN simulations

We test the implementation by running 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, nitrification, denitrification, plant nitrogen uptake). For comparison with CLM, 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_smminn` to 0 (Oleson et al., 2013). Spin-up simulations are used because they are numerically more challenging as 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 1000, 600, and 600 year for the arctic, temperate, and tropical site, respectively. In the base case, $k_m = 10^{-6} \text{ mol m}^{-3}$ and residual concentration $10^{-15} \text{ mol m}^{-3}$. To assess the sensitivity of various preference levels for ammonium and nitrate uptake, and downregulation levels, we examine $k_m = 10^{-3}$ to $10^{-9} \text{ mol m}^{-3}$. We evaluate how scaling, clipping, and log transformation for avoiding negative concentrations influence accuracy and efficiency.

400 3.4.1 Site descriptions

The US-Brw site (71.35N,156.62W) is located near Barrow, Alaska. The mean annual temperature, precipitation, and snowfall are -12°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 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 (AmeriFlux data).

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 °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 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).

3.4.2 CLM-PFLOTRAN site simulation results

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 (US-Brw), temperate (US-WBW), and tropical (BR-Cax) sites, respectively. This introduces a multi-year cycle in addition to the annual cycle (Figs. 4, 5, 6). Overall, CLM-PFLOTRAN is close to CLM4.5 in predicting LAI and nitrogen distribution among vegetation, litter, SOM, ammonium and nitrate pools for the arctic (Fig. 4), temperate (Fig. 5), and tropical (Fig. 6) sites . CLM4.5 does reach equilibrium earlier than CLM-PFLOTRAN. The maximum differences occur during the transient (200-400 year for the arctic, and 50-70 year for the temperate and tropical sites) for SOMN, ammonium, and nitrate. This is expected as the nitrogen demand competition scheme implemented in CLM-PFLOTRAN is different from that in CLM4.5 (Fig. 2), the former solves the reaction network simultaneously while latter sequentially (resolve the plant uptake and decomposition at first, and then nitrification, and denitrification) , and the carbon nitrogen cycle is very sensitive to the nitrogen competition representation.

The arctic site shows a distinct summer growing season (Fig. 4): 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. Ammonium and nitrate concentration 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 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 ammonium and nitrate concentrations (Fig. 5e 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 equilibrium.

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 ammonium and nitrate concentrations during the spinup (Fig. 6), 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.,
445 inset in Figs. 4e,f, 5e). 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, and efficiency of the numerical solutions.

3.4.3 Accuracy and efficiency

450 Numerical errors introduced in clipping, scaling or log transformation are captured in CLM when it checks carbon and nitrogen mass balance for every time step for each column, and report $\geq 10^{-8}$ g m⁻² errors. When log transformation is used, mass balance errors are not recorded for the arctic, temperate, and tropical sites with k_m values 10^{-3} , 10^{-6} , and 10^{-9} mol m⁻³. The computing time for CLM-PFLOTRAN is about 60% to 100% more than that of CLM (Table 1). This is not unreasonably
455 high as the implicit method involves matrix inversion, and log transformation converts the linear problem into nonlinear problem. The computational cost increases substantially with decreasing half saturation, which is expected as a smaller half saturation requires smaller time step sizes to march through steeper transition between the zero and first order rate in Monod function. Overall, log transformation is accurate, robust, and reasonably efficient.

460 Mass balance errors are reported for k_m values of 10^{-6} , and 10^{-9} but not for 10^{-3} mol m⁻³ when clipping is applied. With $k_m = 10^{-3}$ mol m⁻³, the plant uptake and immobilization are inhibited at relatively high concentration so that nitrogen concentrations are high. With decreasing from $k_m = 10^{-6}$ to 10^{-9} mol m⁻³, nitrogen concentrations are lowered to much lower level (Figs. 4, 5, 6, similar to Fig. 3c), increasing the likelihood of overshoot. Mass balance errors are recorded when the
465 relative update is below STOL to preventing further iterations to resolve the mass balance errors introduced by clipping. The frequency of mass balance errors decreases with increasing k_m , and decreasing STOL. Tightening STOL from 10^{-8} to 10^{-12} , the reported greater than 10^{-8} g m⁻³ mass balance errors are eliminated. The computing time is about 50% more than CLM, which is more efficient than log transformation (Table 1), particularly for $k_m = 10^{-9}$ mol m⁻³. Tightening STOL only
470 slightly increases the computing time. Because clipping often occurs at very low concentrations, the reported mass balance errors are usually small (mostly $\sim 10^{-8}$ gN m⁻³ and some $\sim 10^{-7}$ gN m⁻³), and do not have substantial impact on the final simulation results.

The results for scaling is similar to clipping: mass balance errors are recorded for k_m values of 10^{-6} and 10^{-9} but not for 10^{-3} mol m⁻³; tightening STOL to 10^{-12} resolves the error; it takes
475 about 50% more computing time than CLM. To examine the influence of low concentrations on the accuracy and efficiency of the scaling method, we conduct numerical experiments in which we reset the nitrous oxide concentration produced from decomposition (reaction R11, rate Eq. A3) to 10^{-12} , 10^{-10} , or 10^{-8} mol m⁻³ in each CLM half hour time step for the tropical site for the first year. This can be used to calculate the nitrous oxide production rate from decomposition and feed back

480 to CLM without saving the concentration for the previous time step. Overall, nitrogen is abundant
in the first half year, and then becomes limiting in the last five months (Fig. 6e,f, inset). We look
into the daily ammonium cycles as an example during the nitrogen limiting period (day 250 to 260,
Fig. 7a). Every day the ammonium concentration increases with time due to deposition, but drops
when the plant nitrogen demand shots up. With reset concentration of 10^{-8} mol m $^{-3}$, the minimum
485 nitrous oxide concentration for the ten layers is 10^{-8} mol m $^{-3}$, and ammonium concentrations show
two peaks followed by two drops due to the two plant uptake peaks every day. Decreasing the reset
concentration to 10^{-10} mol m $^{-3}$, the minimum concentration drops to 10^{-12} , 10^{-14} , and 10^{-16}
mol m $^{-3}$, corresponding to 1, 2, and 3 iterations with overshoot for nitrous oxide. These result
in numerical "inhibition" of ammonium rebound everyday. It worsens with further decreasing the
490 reset concentration to 10^{-12} mol m $^{-3}$. This introduces mass balance errors as reported in CLM
because the false convergence numerically inhibits all of the reactions including nitrogen deposition,
litter input from CLM to PFLOTRAN. Unlike clipping, these false convergences introduce excessive
numerical errors.

The frequent positive update to nitrous oxide is produced because the rate for the nitrification
495 reaction (R11) is parameterized as a fraction of net mineralization rate to reflect the relationship
between labile carbon content and nitrous oxides production (Parton et al., 1996). A Monod function
is added to describe the limitation of ammonium concentration on nitrification. Calculation of
the net mineralization rate involves all of the decomposition reactions, and the litter decomposition
500 reactions bring in ammonium and nitrate limitation, and ammonium inhibition on nitrate immobi-
lization. As a result, the off-diagonal terms for nitrous oxide in the Jacobian matrix corresponding
to Lit1C, Lit1N, Lit2C, Lit2N, Lit3C, Lit3N, SOM1, SOM2, SOM3, SOM4, ammonium, and
nitrate are nonzero. Positive updates to all of these species are expected to contribute to positive
update to nitrous oxide. While the empirical parameterization of nitrification rate as a function of net
mineralization rate is conceptually convenient, it increases the complexity of the reaction network
505 and numerical challenges due to the less sparse Jacobian matrix. While we use nitrous oxide as an
example here, similar results can be obtained for PlantN, PlantA, and nitrogen gas produced from
denitrification, etc. Theoretically, the numerical "inhibition" of all reaction can be caused by positive
update to very low concentrations for any species.

The numerical errors can be decreased and eliminated by decreasing STOL. Similar to the Test 2
510 (SI test2.xlsx sheet scale2), a small STOL can result in small stol, then a very small STOL is needed.
For the case of reset concentration 10^{-10} for the one year tropical site simulation, the numerical
"inhibition" decreases with decreasing STOL and vanishes for the observed time window when
STOL = 10^{-50} (Fig. 8), indicating the need for very small, zero or even negative STOL to avoid
false convergence. We test STOL = 10^{-200} and find that the temperate site simulations involve
515 extensive time step cutting with default maximum number of iterations of 16 (MAXIT, Appendix
B). The impact of resetting nitrous oxide concentration on clipping and log transformation is less

dramatic. Nevertheless, the computing time increases about 10% for clipping, and double for the log transformation.

4 Summary and Conclusions

520 Global land surface models have traditionally represented subsurface biogeochemical processes using preconfigured reaction networks. This hardcoded approach makes it necessary to revise source code to test alternative models or to incorporate improved process understanding. We couple PFLOTRAN with CLM to facilitate testing of alternative models and incorporation of new understanding. We implement CLM-CN decomposition cascade, nitrification, denitrification, and plant nitrogen
525 uptake reactions in CLM-PFLOTTRAN. We illustrate that with implicit time stepping and Newton-Raphson method, which are typically used in reactive transport codes, the concentration can become negative during the iterations even for species that have no consumption, which must be prevented by intervening in the Newton-Raphson iteration procedure.

Clipping, scaling and log transformation can all prevent negative concentration. However, our
530 results reveal implications when the relative update tolerance (STOL) is used as one of the convergence criteria. While use of STOL improves efficiency in many situations, satisfying STOL does not guarantee satisfying the residual equation, and therefore may introduce false convergence. Clipping reduces the consumption but not the production in some reactions, which requires subsequent iterations to solve the residual equations. A tight STOL is needed to avoid false convergence and
535 prevent mass balance errors. While the scaling method dampens the whole update vector following the stoichiometry of the reactions to keep mass balance, a small scaling factor caused by a positive update to a small concentration may diminish the update and result in false convergence, numerically inhibiting all reactions. For accuracy and efficiency, a very tight STOL is needed when the concentration can be very low. Log transformation is accurate and robust, but requires more computing
540 time. The computational cost increases with decreasing concentrations, most substantially for log transformation.

Our CLM-PFLOTTRAN spin-up simulations at arctic, temperate, and tropical sites produce results similar to CLM4.5, and indicate that accurate and robust solution can be achieved with clipping, scaling or log transformation. The computing time is 50% to 100% more than CLM4.5 for a range
545 of half saturation values from 10^{-3} to 10^{-9} and a residual concentration of 10^{-15} for nitrogen. As physical half saturation ranges from 10^{-5} to 10^{-6} M for nitrogen, and the detection limits are often above 10^{-9} M, our results demonstrate accurate, efficient, and robust solution for current CLM biogeochemistry using PFLOTTRAN. We thus demonstrate the feasibility of using an open-source, general-purpose reactive transport code with CLM, thus enabling significantly more complicated
550 and more mechanistic biogeochemical reaction systems.

An alternative to our approach of coupling LSMs with reactive transport codes is to code the solution to the advection, diffusion, and reaction equations directly in the LSM. This has been done using explicit time stepping and operator splitting to simulate the transport and transformation of carbon, nitrogen, and other species in CLM Tang et al. (2013). An advantage of our approach of
555 using a community RTM to solve the advection-dispersion-reaction system is that the significant advances that the RTM community has made in the past several decades can be leveraged to better represent the geochemical processes (e.g., pH, pE) in a systematic, flexible, and numerically reliable way. Given that a wide range of conditions may be encountered in any one global LSM simulation, it is particularly important to have robust solution methods such as fully implicit coupling of the
560 advection-dispersion-reaction equations. As a next step, we hope CLM-PFLOTRAN will facilitate investigation of the role of the redox sensitive microbial reactions for methane production and consumption, and nitrification and denitrification reactions in ecological responses to climate change.

5 Code availability

PFLOTRAN is an open-source software. It is distributed under the terms of the GNU Lesser General
565 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>. CLM-PFLOTRAN is under development and will be available subject to guidelines of NGEE-Arctic and ACME projects.

6 Author contribution

G. B., B. A., R. M., J. K., and F. H. developed the CLM-PFLOTRAN framework that this work
570 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. contributed equally to the work.

575 *Acknowledgements.* Thanks to Nathaniel O. Collier at ORNL for many discussions 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 Ecosystem Experiments in the Arctic (NGEE-Arctic) project. ORNL is managed by UT-Battelle,
580 LLC, for the U.S. Department of Energy under contract DE-AC05-00OR22725.

References

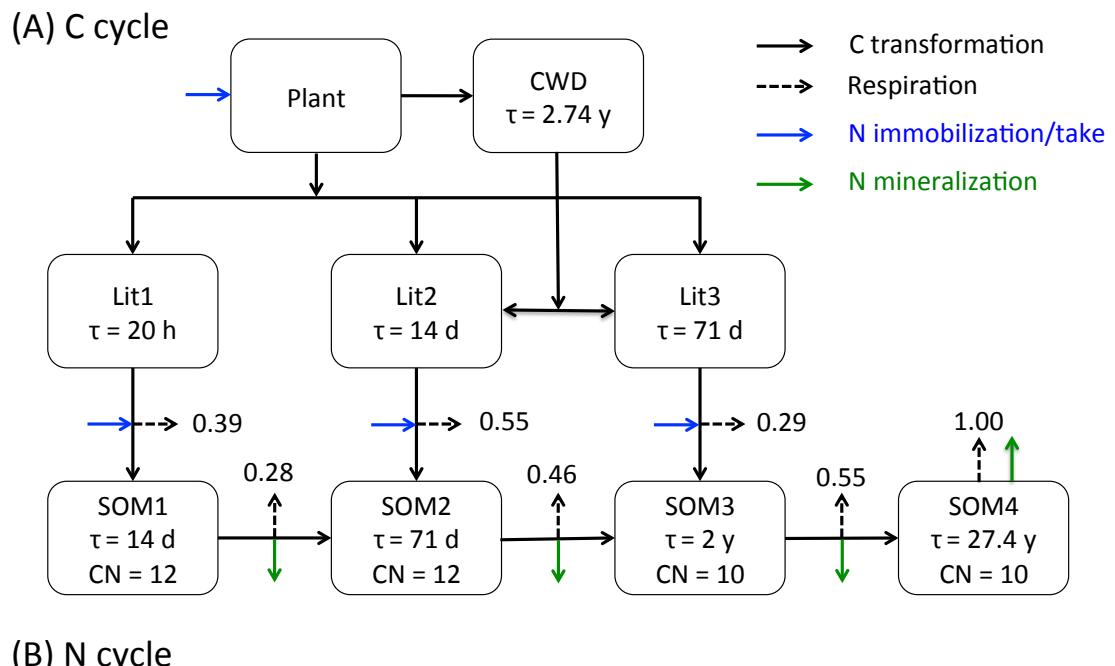
- Balay, S., Abhyankar, S., Adams, M., Brown, J., Brune, P., Buschelman, K., Dalcin, L., Eijkhout, V., Gropp, W., Karpeyev, D., Kaushik, D., Knepley, M., Curfman McInnes, L., Rupp, K., Smith, B., Zampini, S., and Zhang, H.: PETSc Users Manual Revision 3.6, Report ANL-95/11, Argonne National Laboratory, 2015.
- 585 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.
- 590 Boyer, E. W., Alexander, R. B., Parton, W. J., Li, C., Butterbach-Bahl, K., Donner, S. D., Skaggs, R. W., and Gross, 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.
- Conrad, R.: Soil microorganisms as controllers of atmospheric trace gases (H₂, CO, CH₄, OCS, N₂O, and NO),
595 Microbiological Reviews, 60, 609–640, <http://mmbrr.asm.org/content/60/4/609.full.pdf>, 8987358[pmid] Microbiol Rev, 1996.
- 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
600 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.
- Falkengren-Grerup, U.: Interspecies differences in the preference of ammonium and nitrate in vascular plants,
610 *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.
- Foulland, E., Gosselin, M., Rivkin, R. B., Vasseur, C., and Mostajir, B.: Nitrogen uptake by heterotrophic
615 bacteria and phytoplankton in Arctic surface waters, *Journal of Plankton Research*, 29, 369–376, <http://plankt.oxfordjournals.org/content/29/4/369.abstract>, 2007.
- 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.

- 620 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.
- 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, 625 doi:10.1016/j.jconhyd.2009.12.003, <http://www.sciencedirect.com/science/article/pii/S0169772209001648>, 2010.
- 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.
- 630 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, 635 doi:10.1890/03-4049, <http://www.esajournals.org/doi/pdf/10.1890/03-4049>, 2004.
- Hartman, M. D., Baron, J. S., and Ojima, D. S.: Application of a coupled ecosystem-chemical equilibrium model, DayCent-Chem, to stream and soil chemistry in a Rocky Mountain watershed, *Ecological Modelling*, 200, 493–510, doi:10.1016/j.ecolmodel.2006.09.001, <http://www.sciencedirect.com/science/article/pii/S0304380006003942>, 2007.
- 640 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.
- HØGh-Jensen, H., Wollenweber, B., and Schjoerring, J. K.: Kinetics of nitrate and ammonium absorption and 645 accompanying H⁺ fluxes in roots of *Lolium perenne* L. and N₂-fixing *Trifolium repens* L, *Plant, Cell & Environment*, 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, 650 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 655 *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 Pa-*

- 660 pers, 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.
- 665 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.
- Lemmon, D. R. and Schafer, J. L.: Developing Statistical Software in Fortran 95, *Statistics and Computing*, Springer, 2005.
- 670 Lichtner, P. C., Hammond, G. E., Lu, C., Karra, S., Bisht, G., Andre, B., Mills, R. T., and Jitu, K.: PFLO-TRAN 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 675 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.
- 680 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., Höglberg, 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.
- 685 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, 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.: 695 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.
- 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>,
700 2001.
- Pfautsch, S., Rennenberg, H., Bell, T. L., and Adams, M. A.: Nitrogen uptake by *Eucalyptus regnans* and
Acacia spp. – preferences, resource overlap and energetic costs, *Tree Physiology*, 29, 389–399, <http://treephys.oxfordjournals.org/content/29/3/389.abstract>, 2009.
- Powell, T. L., Galbraith, D. R., Christoffersen, B. O., Harper, A., Imbuzeiro, H. M. A., Rowland, L., Almeida, S.,
705 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.
- Riley, W. J., Maggi, F., Kleber, M., Torn, M. S., Tang, J. Y., Dwivedi, D., and Guerry, N.: Long residence times
710 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-
Knabner, I., Lehmann, J., Manning, D. A. C., Nannipieri, P., Rasse, D. P., Weiner, S., and Trumbore, S. E.:
715 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-
changes of Energy, Water, and Carbon Between Continents and the Atmosphere, *Science*, 275, 502–509,
720 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.
- 725 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.
- Sonnenthal, E. L., Spycher, N., Xu, T., Zheng, L., Miller, N. L., and Pruess, K.: TOUGHREACT V3. 0-OMP
Reference Manual: A Parallel Simulation Program for Non-Isothermal Multiphase Geochemical Reactive
730 Transport, LBNL Report in preparation, Report, 2014.
- 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, 2013.
- Tang, J. Y. and Riley, W. J.: Technical Note: A generic law-of-the-minimum flux limiter for simulating sub-
735 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, Estuarine, Coastal and Shelf Science, 61, 507–515, doi:10.1016/j.ecss.2004.06.014, <http://www.sciencedirect.com/science/article/pii/S0272771404001593>, 2004.
- 745 Warren, C. R. and Adams, P. R.: Uptake of nitrate, ammonium and glycine by plants of Tasmanian wet eucalypt forests, Tree Physiology, 27, 413–419, <http://treephys.oxfordjournals.org/content/27/3/413.abstract>, 10.1093/treephys/27.3.413, 2007.
- White, M. D. and McGrail, B. P.: Stomp (subsurface transport over multiple phase) version 1.0 addendum: Eckechem equilibrium-conservationkinetic equation chemistry and reactive transport, Report, 2005.
- 750 Xu, T., Sonnenthal, E., Spycher, N., and Pruess, K.: TOUGHREACT—A simulation program for non-isothermal multiphase reactive geochemical transport in variably saturated geologic media: Applications to geothermal injectivity and CO₂ geological sequestration, Computers & Geosciences, 32, 145–165, doi:10.1016/j.cageo.2005.06.014, <http://www.sciencedirect.com/science/article/pii/S0098300405001500>, 2006.
- 755 Yeh, G. T., Sun, J., Jardine, P. M., Burgos, W. D., Fang, Y., Li, M. H., and Siegel, M. D.: HYDROGEOCHEM 5.0: A Coupled Model of Fluid Flow, Thermal Transport, and HYDROGEO-CHEMical Transport through Saturated-Unsaturated Media: Version 5.0, Report, 2004.



(B) N cycle

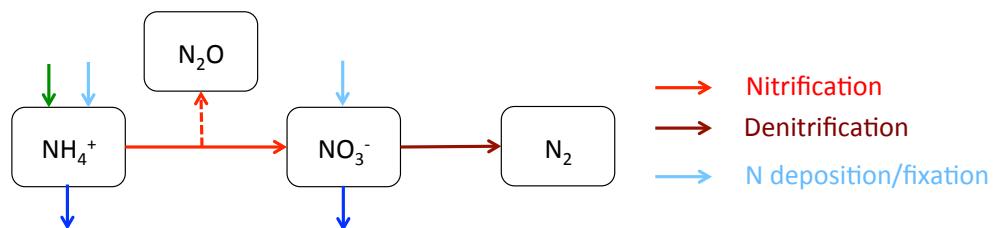


Figure 1. The reaction network for the carbon (A) and nitrogen (B) cycles implemented in this work. The carbon cycle is modified from Thornton and Rosenbloom (2005) and Bonan et al. (2012). τ is the turnover time, and CN is the CN ratio in gC over gN.

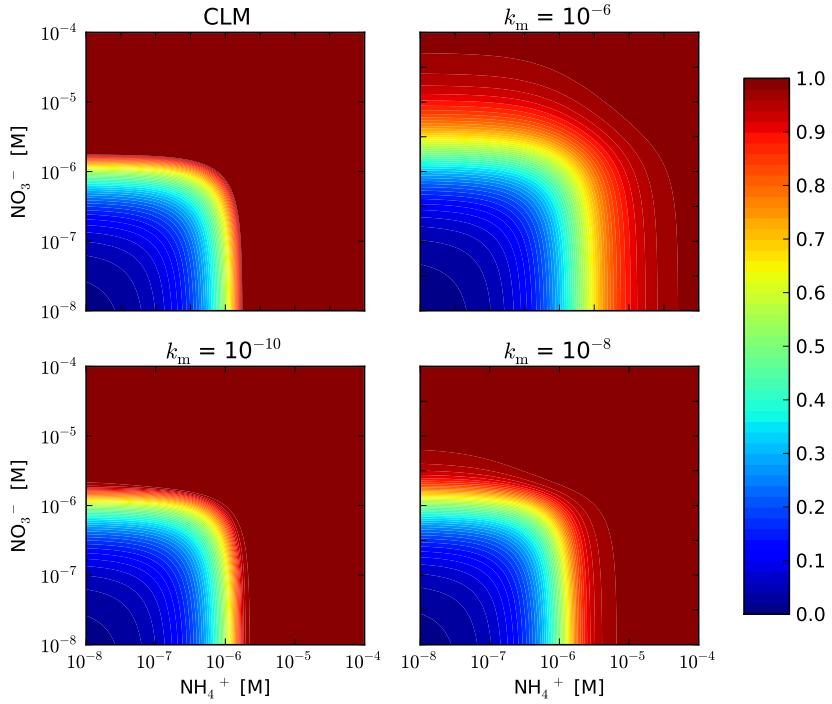


Figure 2. The ratio of uptake and demand (f_{pi}) as a function of concentrations with CLM and representation by Eqs. (3) and (4) in a 0.5 h time step with an uptake rate of 10^{-9} M s^{-1} . f_{pi} for the new representation is less than or equal to that for CLM. The difference decreases with decreasing half saturation k_m .

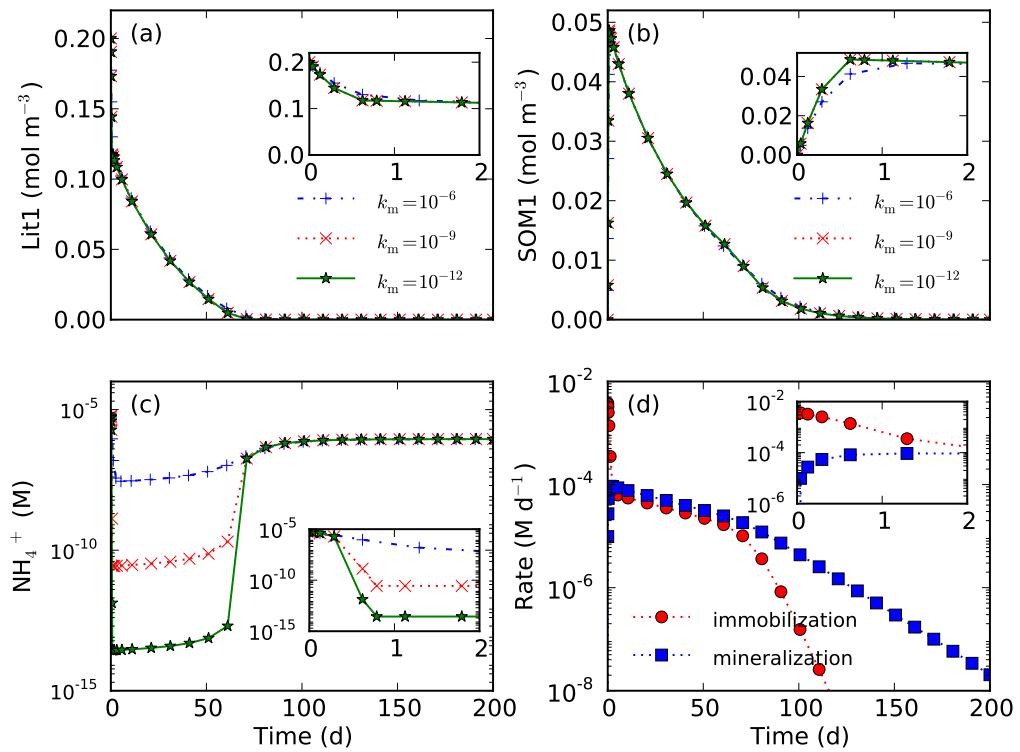


Figure 3. Influence of half saturation k_m on decomposition that involves both nitrogen immobilization and mineralization. Smaller half saturation can result in lower nitrogen concentration (c) but does not substantially impact the calculated concentrations other than ammonium (a,b).

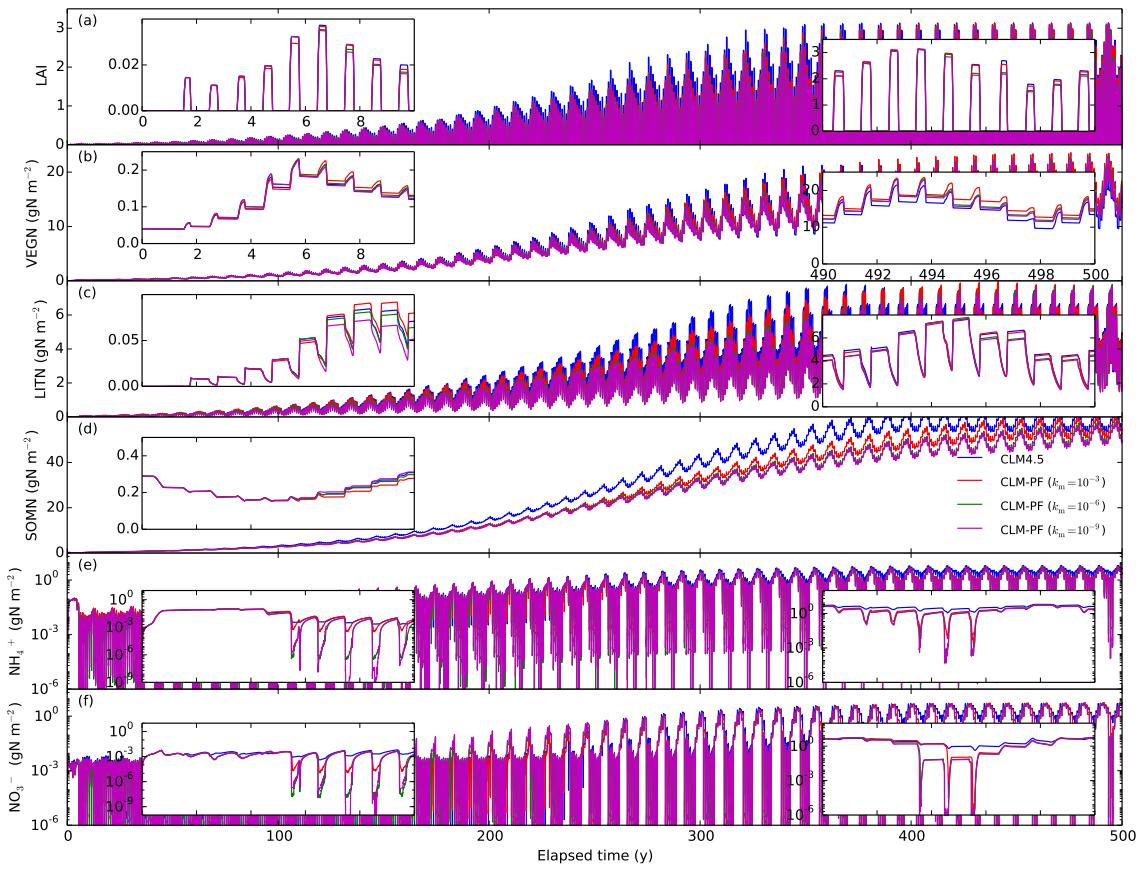


Figure 4. Calculated LAI and nitrogen distribution among vegetation, litter, SOM, NH_4^+ , and NO_3^- pools in spin-up simulations for the US-Brw site.

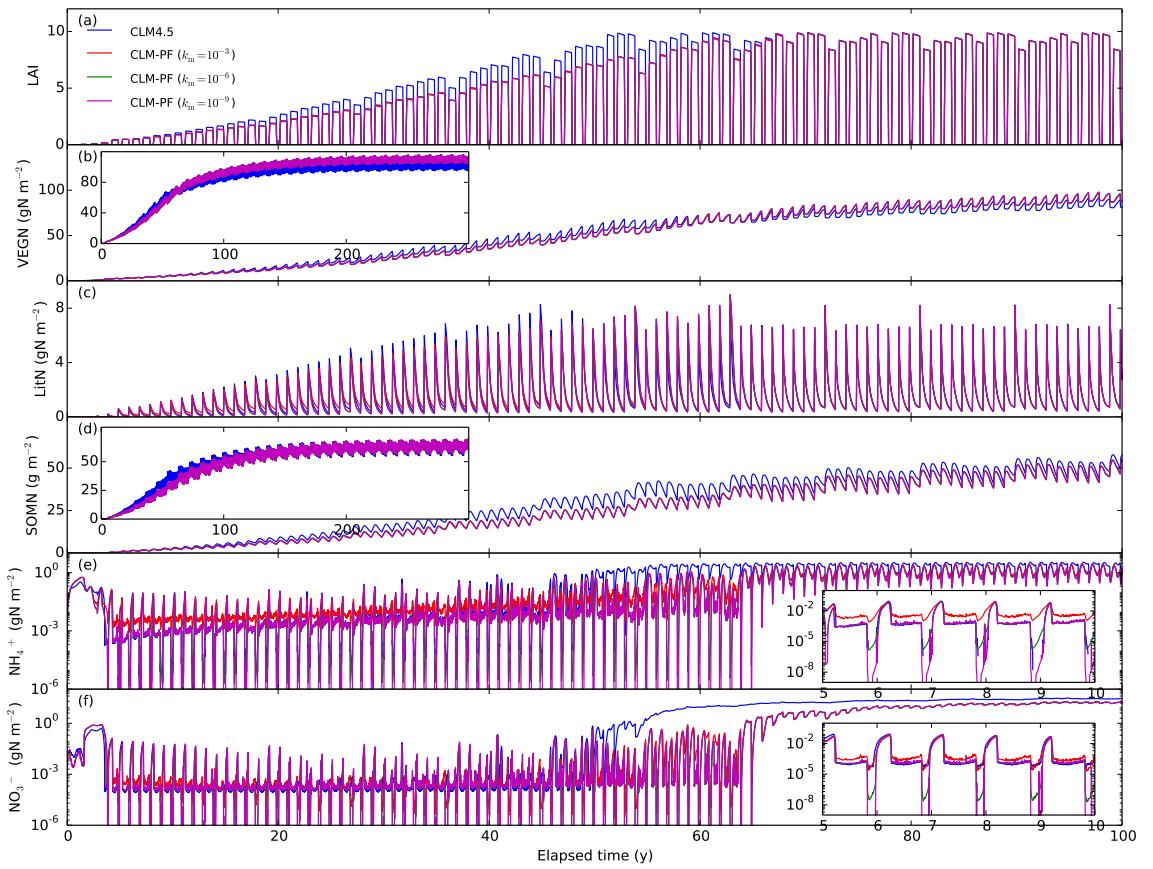


Figure 5. Calculated LAI and nitrogen distribution among vegetation, litter, SOM, NH_4^+ , and NO_3^- pools in spin-up simulations for US-WBW site.

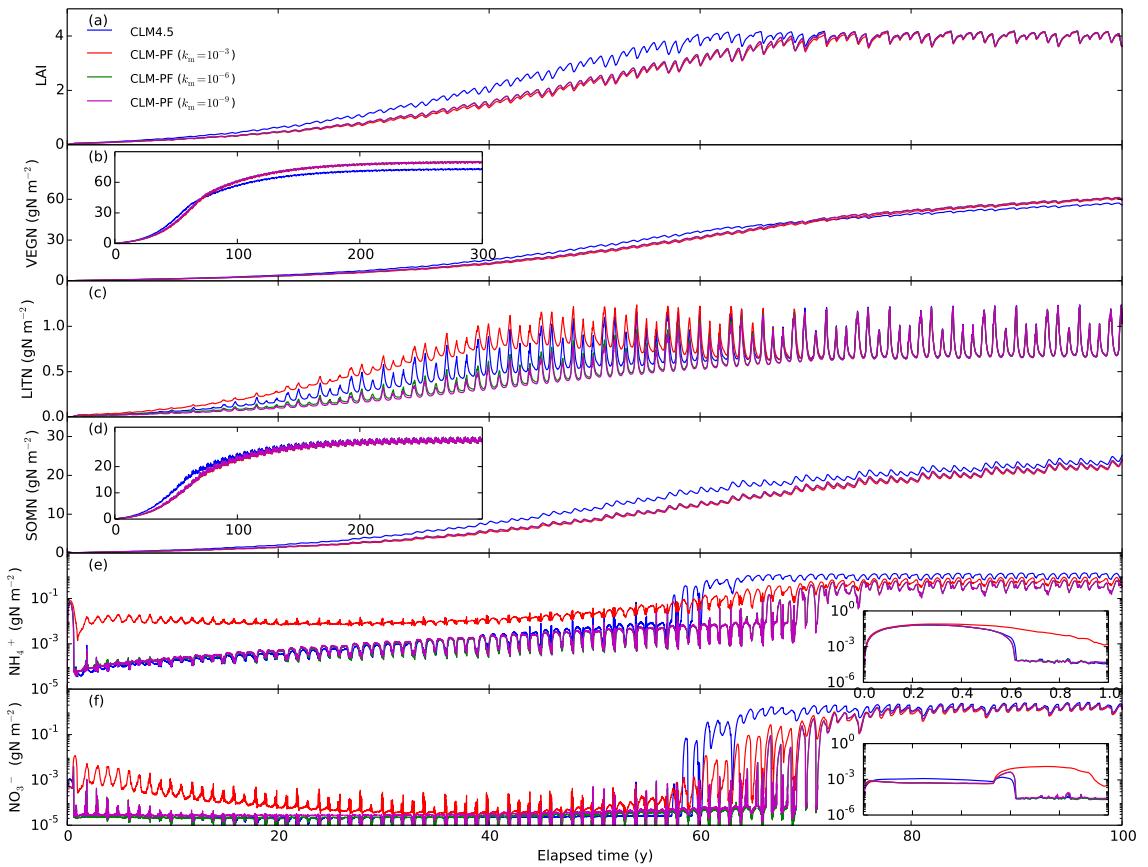


Figure 6. Calculated LAI and nitrogen distribution among vegetation, litter, SOM, NH_4^+ , and NO_3^- pools in spin-up simulations for BR-Cax site.

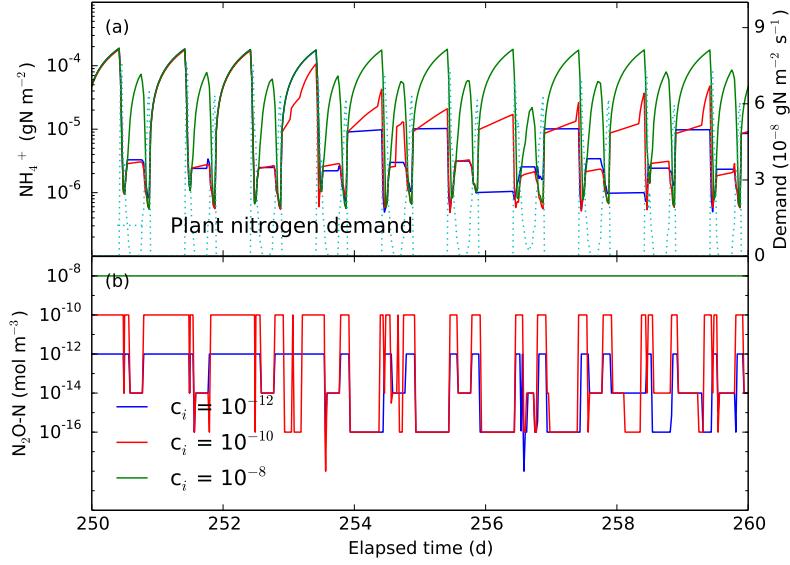


Figure 7. Resetting nitrous oxide concentration to 10^{-8} , 10^{-10} , and $10^{-12} \text{ mol m}^{-3}$ in every CLM 0.5 h time step results in no inhibition to increasing inhibition of reactions when the scaling method is used with $\text{STOL} = 10^{-8}$. $\text{N}_2\text{O}-\text{N}$ concentration in y-axis in (b) is the minimum of the 10 soil layers. Numerical experiments are conducted for the tropical site for the first year with $k_m = 10^{-6} \text{ mol m}^{-3}$. See inset in Fig. (6e) for ammonium concentration in the first year with daily data points.

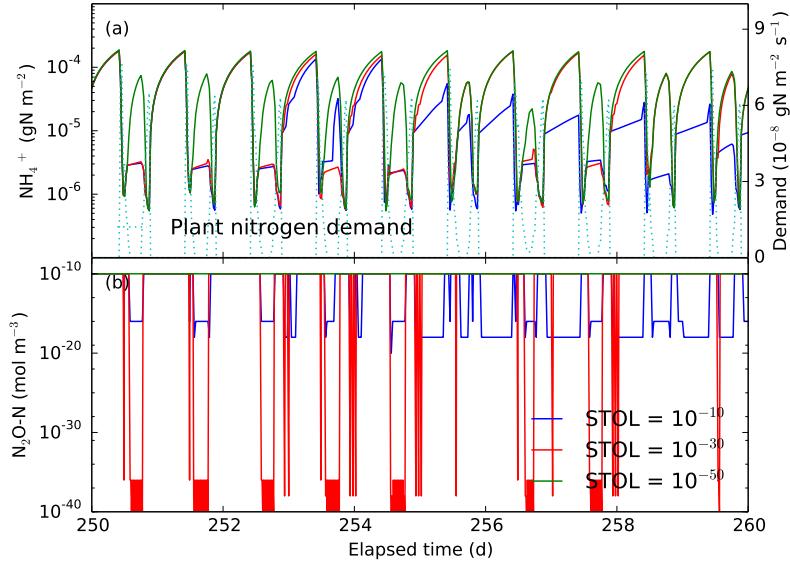


Figure 8. Decreasing STOL can decrease and eliminate the numerical inhibition in the case of $10^{-10} \text{ mol m}^{-3}$ in Fig. (7). $\text{N}_2\text{O}-\text{N}$ concentration in y-axis in (b) is the minimum of the 10 soil layers.

Table 1. Wall time for CLM-PFLOTRAN relative to CLM for spinup simulation on OIC (ORNL Institutional Cluster Phase5)

Site	Clipping			Scaling			Log transformation			
	k_m	10^{-3}	10^{-6}	10^{-9}	10^{-3}	10^{-6}	10^{-9}	10^{-3}	10^{-6}	10^{-9}
US-Brw		1.28	1.30	1.30	1.29	1.29	1.32	1.45	1.49	1.72
US-Pit		1.45	1.47	1.47	1.45	1.45	1.47	1.64	1.68	1.89
BR-Cax		1.43	1.49	1.55	1.44	1.48	1.52	1.62	1.66	1.99

CLM wall time is 29.3, 17.7, and 17.1 hour for the arctic, temperate, and tropical sites for a simulation duration of 1000, 600, and 600 year.

Appendix A: CLM biogeochemical reactions and rates

A1 CLM-CN decomposition

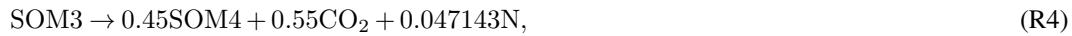
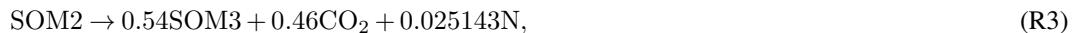
760 The CLM-CN decomposition cascade consists of three litter pools with variable CN ratios, four soil organic matter (SOM) pools with constant CN ratios, and seven reactions (Bonan et al., 2012; Oleson et al., 2013; Thornton and Rosenbloom, 2005). The reaction can be described by



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



775 and

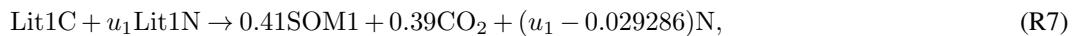


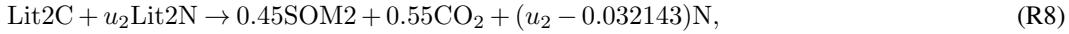
The exact stoichiometric coefficients are calculated in the code using values for respiration factor, CN ratio, and molecular weight specified in the input file.

CLM4.5 has an option to separate N into NH_4^+ and NO_3^- . The N mineralization product is NH_4^+ .
780 As the CN ratio is variable for the three litter pools, litter N pools need to be tracked such that reaction (R1) becomes

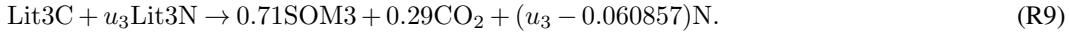


with $u = [\text{LitN}]/[\text{LitC}]$. The three litter decomposition reactions are





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

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)

$$\frac{d[\text{NH}_4^+]}{dt} = -\frac{d[\text{NO}_3^-]}{dt} = -k_n f_T f_w [\text{NH}_4^+]. \quad (\text{A2})$$

The nitrification reaction to produce N_2O is



with one component related to decomposition as

$$\frac{d[\text{NH}_4^+]}{dt} = -2 \frac{d[\text{N}_2\text{O}]}{dt} = -f_{nm} f_T f_w f_{pH} \max(R_{nm}, 0) \quad (\text{A3})$$

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

$$R_{nm} = \sum_i n_i R_i, \quad (\text{A4})$$

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

$$\frac{d[\text{NH}_4^+]}{dt} = -2 \frac{d[\text{N}_2\text{O}]}{dt} = -k_{n2o} f_T f_w f_{pH} (1 - e^{-0.0105[\text{NH}_4^+]}). \quad (\text{A5})$$

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

$$\frac{d[\text{NH}_4^+]}{dt} = -2 \frac{d[\text{N}_2\text{O}]}{dt} = -k_{n2o} f_T f_w f_{pH} [\text{NH}_4^+]. \quad (\text{A6})$$

810 **A3 Denitrification**

The denitrification reaction is



with rate (Dickinson et al., 2002)

$$\frac{d[\text{NO}_3^-]}{dt} = -2\frac{d[\text{N}_2]}{dt} = -k_{\text{deni}} f_T f_w f_{\text{pH}} [\text{NO}_3^-]. \quad (\text{A7})$$

815 **A4 Plant nitrogen uptake**

The plant nitrogen uptake reaction can be written as



and



- 820 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 demand distribution between ammonium and nitrate

Denote $R_{d,p}$ and $R_{d,i}$ as the potential plant, immobilization, nitrification, and denitrification demand (rate); $R_{a,tot} = R_{d,p} + R_{d,i}$ as the total NH_4^+ demand; and $R_{n,tot}$ as the total NO_3^- demand. CLM 825 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 to potential demands, and $R_{n,tot} = 0$; otherwise, the uptakes for NH_4^+ are $[\text{NH}_4^+]R_{d,p}/R_{a,tot}\Delta t$ and $[\text{NH}_4^+]R_{d,i}/R_{a,tot}\Delta t$ for plants and immobilization; $R_{n,tot} = R_{a,tot} - [\text{NH}_4^+]/\Delta t$. If $R_{n,tot}\Delta t < [\text{NO}_3^-]$, all of the remaining demand 830 $R_{n,tot}$ 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.

Appendix B: Implicit time stepping and Newton-Raphson iteration

Ignoring equilibrium reactions and transport for simplicity of discussion in this work, PFLOTRAN solves the ordinary differential equation,

$$835 \quad \frac{dc}{dt} = \mathbf{R}(\mathbf{c}), \quad (\text{B1})$$

with \mathbf{c} as the concentration vector and \mathbf{R} as the kinetic reaction rate. Discretizing Eq. (B1) in time using the backward Euler method,

$$(\mathbf{c}^{k+1} - \mathbf{c}^k)/\Delta t = \mathbf{R}(\mathbf{c}^{k+1}). \quad (\text{B2})$$

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

$$840 \quad \mathbf{f}(\mathbf{c}^{k+1,p}) = (\mathbf{c}^{k+1,p} - \mathbf{c}^k)/\Delta t - \mathbf{R}(\mathbf{c}^{k+1,p}), \quad (B3)$$

and Jacobian as

$$\mathbf{J} = \frac{\partial \mathbf{f}(\mathbf{c}^{k+1,p})}{\partial \mathbf{c}^{k+1,p}}, \quad (B4)$$

the update is

$$\delta \mathbf{c}^{k+1,p} = \mathbf{J}^{-1} \mathbf{f}(\mathbf{c}^{k+1,p}), \quad (B5)$$

$$845 \quad \text{and the iteration equation is}$$

$$\mathbf{c}^{k+1,p+1} = \mathbf{c}^{k+1,p} - \delta \mathbf{c}^{k+1,p}. \quad (B6)$$

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,

$$\| \mathbf{f}(\mathbf{c}^{k+1,p}) \|_2 < \text{ATOL}, \quad (B7)$$

$$850$$

$$\frac{\| \mathbf{f}(\mathbf{c}^{k+1,p}) \|_2}{\| \mathbf{f}(\mathbf{c}^{k+1,0}) \|_2} < \text{RTOL}, \quad (B8)$$

or

$$\frac{\| \delta \mathbf{c}^{k+1,p} \|_2}{\| \mathbf{c}^{k+1,p} \|_2} < \text{STOL}. \quad (B9)$$

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. The default values in PFLOTRAN are $\text{ATOL} = 10^{-50}$, $\text{RTOL} = 10^{-8}$, $\text{STOL} = 10^{-8}$, $\text{MAXIT} = 50$, $\text{MAXF} = 10^4$, and $\text{MAX_CUT} = 16$.

Appendix C: Matrix equation for example 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 $J_{deni} = \frac{dR_{deni}}{d[\text{NO}_3^-]} = k_{deni}$, the matrix equation (Eq. B5) becomes Eq. (C1),

$$\begin{bmatrix} \frac{1}{\Delta t} + J_{at} + J_{nitr} & 0 & 0 & 0 \\ -J_{at} & \frac{1}{\Delta t} & 0 & 0 \\ -J_{nitr} + J_{nt,a} & 0 & \frac{1}{\Delta t} + J_{nt} + J_{deni} & 0 \\ -J_{nt,a} & 0 & -J_{nt,n} & 1/\Delta t \\ 0 & 0 & -0.5J_{deni} & 0 \end{bmatrix} \begin{pmatrix} \delta[\text{NH}_4^+]^{k+1,1} \\ \delta[\text{PlantA}]^{k+1,1} \\ \delta[\text{NO}_3^-]^{k+1,1} \\ \delta[\text{PlantN}]^{k+1,1} \\ \delta[\text{N}_2]^{k+1,1} \end{pmatrix} = \begin{pmatrix} R_{at} + R_{nitr} \\ -R_{at} \\ -R_{nitr} + R_{nt} + R_{deni} \\ -R_{nt} \\ -0.5R_{deni} \end{pmatrix}. \quad (C1)$$