

# Addressing numerical challenges in introducing a reactive transport code into a land surface model: a biogeochemical modeling proof-of-concept with CLM–PFLOTRAN 1.0

Guoping Tang<sup>1</sup>, Fengming Yuan<sup>1</sup>, Gautam Bisht<sup>1,2</sup>, Glenn E. Hammond<sup>3,4</sup>, Peter C. Lichtner<sup>5</sup>, Jitendra Kumar<sup>1</sup>, Richard T. Mills<sup>6</sup>, Xiaofeng Xu<sup>1,7</sup>, Ben Andre<sup>2,8</sup>, Forrest M. Hoffman<sup>1</sup>, Scott L. Painter<sup>1</sup>, and Peter E. Thornton<sup>1</sup>

<sup>1</sup>Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA

<sup>2</sup>Lawrence Berkeley National Laboratory, Berkeley, California, USA

<sup>3</sup>Pacific Northwest National Laboratory, Richland, Washington, USA

<sup>4</sup>Sandia National Laboratories, Albuquerque, New Mexico, USA

<sup>5</sup>OFM Research–Southwest, Santa Fe, New Mexico, USA

<sup>6</sup>Intel Corporation, Hillsboro, Oregon, USA

<sup>7</sup>San Diego State University, San Diego, California, USA

<sup>8</sup>National Center for Atmospheric Research, Boulder, Colorado, USA

*Correspondence to:* P. E. Thornton (thorntonpe@ornl.gov) and S. L. Painter (paintersl@ornl.gov)

**Abstract.** [CE]Note that the abbreviation CLM (Community Land Model) is written as CLMs in the plural form, i.e. Community Land Models. There is no need to make it consistent throughout the text. Please check that the singular and plural forms are used correctly. We explore coupling to a configurable subsurface reactive transport code as a flexible and extensible approach to biogeochemistry in land surface models. A reaction network with the Community Land Model carbon–nitrogen (CLM-CN) decomposition, nitrification, denitrification, and plant uptake is used as an example. We implement the reactions in the open-source PFLOTRAN (massively parallel subsurface flow and reactive transport) code and couple it with the CLM. To make the rate formulae designed for use in explicit time stepping in CLMs compatible with the implicit time stepping used in PFLOTRAN, the Monod substrate rate-limiting function with a residual concentration is used to represent the limitation of nitrogen availability on plant uptake and immobilization. We demonstrate that CLM–PFLOTRAN predictions (without invoking PFLOTRAN transport) are consistent with CLM4.5 for Arctic, temperate, and tropical sites.

Switching from explicit to implicit method increases rigor but introduces numerical challenges. Care needs to be taken to use scaling, clipping, or log transformation to avoid negative concentrations during the Newton iterations. With a tight relative update tolerance (STOL) to avoid false convergence,

an accurate solution can be achieved with about 50 % more computing time than CLM in point mode site simulations using either the scaling or clipping methods. The log transformation method takes 60–100 % more computing time than CLM. The computing time increases slightly for clipping and scaling; it increases 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. The frequent occurrence of very low concentrations (e.g. below nanomolar) can increase the computing time for clipping or scaling by about 20 %, double for log transformation. Overall, the log transformation method is accurate and robust, and the clipping and scaling methods are efficient. When the reaction network is highly nonlinear or the half saturation or residual concentration is very low, the allowable time-step cuts may need to be increased for robustness for the log transformation method, or STOL may need to be tightened for the clipping and scaling methods to avoid false convergence.

As some biogeochemical processes (e.g., methane and nitrous oxide reactions) involve very low half saturation and thresholds, this work provides insights for addressing non-physical negativity issues and facilitates the representation of a mechanistic biogeochemical description in Earth system models to reduce climate prediction uncertainty.

## 1 Introduction

Land surface (terrestrial ecosystem) models (LSMs) calculate the fluxes of energy, water, and greenhouse 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 “bucket”, second generation “biophysical”, and third generation “physiological” models (Seneviratne et al., 2010), current LSMs such as the Community Land Model (CLM) implement comprehensive thermal, hydrological, and biogeochemical processes (Oleson et al., 2013). The importance of accurate representation of soil biogeochemistry in LSMs is suggested by the confirmation that the increase of carbon dioxide ( $\text{CO}_2$ ), methane ( $\text{CH}_4$ ), and nitrous oxide ( $\text{N}_2\text{O}$ ) in the atmosphere since preindustrial time is 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  $\sim 250$  soil biogeochemical models developed in the past  $\sim 80$  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., Riley et al., 2014).

As LSMs usually hardcode the soil biogeochemistry reaction network (pools/species, reactions, rate formulae), substantial effort is often required to modify the source code for testing alternative models and incorporating new process understanding. Fang et al. (2013) demonstrated the use of a reaction-based approach to facilitate implementation of the Community Land Model carbon–nitrogen (CLM-CN) and CENTURY models and incorporation of the 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 multiphase mechanistic carbon and nitrogen models with many speciation and microbial reactions (Maggi et al., 2008; Gu and Riley, 2010; Riley et al., 2014), but it has not been coupled to an 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) among plants and microbes. In CLMs, 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 by soil nitrogen concentration (Oleson et al., 2013; Thornton and Rosenbloom, 2005). This avoids negative concentrations and does not introduce mass balance errors (Tang and Riley, 2016) as CLM uses explicit time stepping.

RTMs often account for limitation of reactant availability on reaction rates for each individual reaction to allow mechanistic representations and flexibility in adding reactions. For maximum robustness, some RTM codes support the use of fully implicit time stepping, usually employing a variant of the Newton–Raphson method. Negative concentration can be introduced during Newton–Raphson iterations, which is not physical and can cause numerical instability and errors (Shampine et al., 2005). In our work with CLMs, 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 nanomolar ( $10^{-9} \text{ M}$ ) level (Conrad, 1996). The redox potential ( $\text{Eh}$ ) needs to be decreased to  $-0.35 \text{ V}$  (oxygen concentration  $< 10^{-22} \text{ M}$ ; Hungate, 1975) for methanogens to grow (Jarrell, 1985).

Three methods are used to avoid negative concentration in RTM 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, 2005; Sonnenthal et al., 2014) in each iteration. Except that log transformation is more computationally demanding (Hammond, 2003), how these methods for enforcing non-negativity affect computational accuracy and efficiency is rarely discussed.

As LSMs need to run under various conditions at the global scale for simulation duration of centuries, it is necessary to resolve accuracy and efficiency issues to use RTM codes for LSMs. The 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 (massively parallel subsurface flow and reactive transport) (Lichtner et al., 2015; Hammond et al., 2014); couple that model to CLM (we refer to the coupled model as CLM–PFLOTRAN); test the implementation at Arctic, temperate, and tropical sites; and examine the implication of using scaling, clipping, and log transformation for avoiding negative concentrations. Although we focus on a number of carbon–nitrogen reactions implemented in PFLOTRAN and CLM, the critical numerical issue of avoiding negative concentrations has broader relevance as biogeochemical representations in LSMs become more mechanistic.

## 2 Methods

Among the many reactions in LSMs are the soil biogeochemical reactions for carbon and nitrogen cycles, in particular the organic matter decomposition, nitrification, denitrifica-

tion, 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 et al., 2006; Schmidt et al., 2011). In this work, we use a network of 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

In CLM–PFLOTRAN, CLM can instruct 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 PFLOTRAN biogeochemistry, with the CLM solving the energy and water flow equations and handling the solute transport (mixing, advection, diffusion, and leaching). Here, we focus on how reactions are implemented and thus only use PFLOTRAN in batch mode (i.e. without transport). However, PFLOTRAN's advection and diffusion capabilities are operational in the CLM–PFLOTRAN coupling described here. In each CLM time step, the CLM provides production rates for Lit1C, Lit1N, Lit2C, Lit2N, Lit3C, Lit3N for litter fall; CWDC and CWDN for coarse woody debris production, nitrogen deposition and fixation, and plant nitrogen demand; and specifies liquid water content, matrix potential, and temperature for PFLOTRAN. PFLOTRAN solves ordinary differential equations for the kinetic reactions and mass action equations for 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_sandbox_example.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; the resulting system of algebraic equations is solved using the Newton–Raphson method (Appendix B). A Krylov subspace method is usually employed to solve the Jacobian systems arising during the Newton–Raphson iterations, but, as the problems are relatively small in this study, we solve them directly via LU (lower upper) factorization.

Unlike the explicit time stepping in CLM, in which only reaction rates need to be calculated, implicit time stepping requires evaluating derivatives. While PFLOTRAN provides an option to calculate derivatives numerically via finite-

differencing, we use analytical expressions for efficiency and accuracy.

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 a flexible number of terms and typical moisture, temperature, and pH response functions is coded in PFLOTRAN. Most of the biogeochemical reactions can be specified in an input file, with a 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 pools and reactions are traditionally hardcoded in CLM. Consequently, any change of the pools, reactions, or rate formula may require source code modification. Therefore, the more general approach used by PFLOTRAN facilitates implementation of increasingly mechanistic reactions and tests of various representations with less code modification.

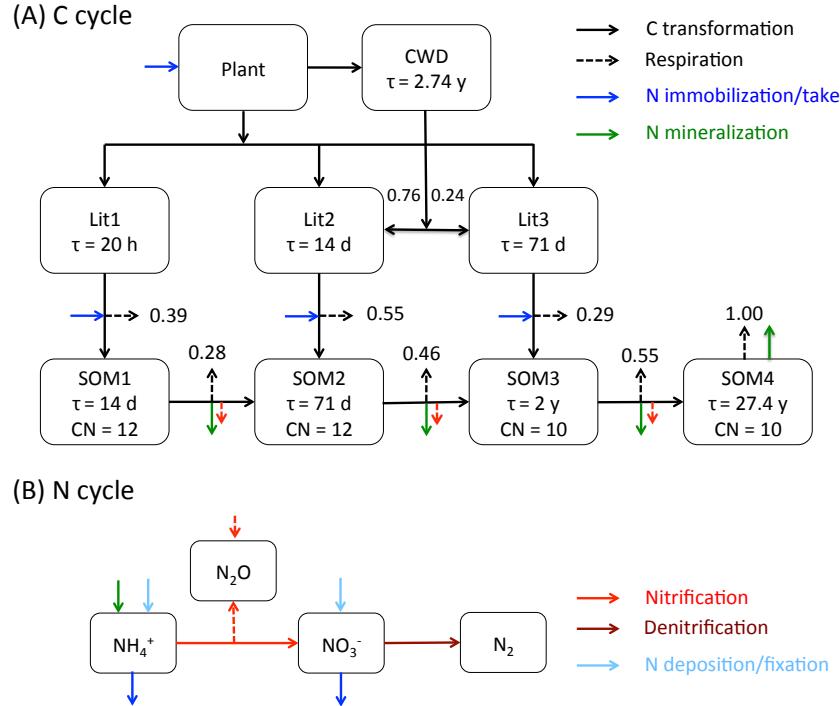
## 2.2 Mechanistic representation of rate-limiting processes

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, and A7): the rate decreases to zero as the concentration decreases to zero. A residual concentration  $[CN_u]_r$  is often added to represent a threshold below which a reaction stops, for example, to the decomposition rate (Eq. A1) as

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

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

For the litter decomposition reactions (Appendix Reactions R7–R9) that immobilize N, the nitrification Reaction (R11) associated with decomposition to produce nitrous oxide, and the plant nitrogen uptake Reactions (R13) and (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 (Fennell and Gossett, 1998), Eq. (1) becomes



**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).  $\tau$  is the turnover time, and CN is the CN ratio in g C over g N.

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

with half saturation  $k_m$  and a mineral nitrogen residual concentration  $[N]_r$ . In the case of  $[N] - [N]_r = k_m$ , the rate-limiting function is equal to 0.5. 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,  $k_m([N] + k_m)^{-2}$ , increases to about  $k_m^{-1}$  as the concentration decreases to below  $k_m$ . This represents 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 ( $NO_3^-$ ), it is necessary to partition the demands between ammonium and nitrate for plant uptake and immobilization. If we simulate the ammonium limitation on plant uptake with

$$R_a = R_p \frac{[NH_4^+]}{[NH_4^+] + k_m}, \quad (3)$$

the plant nitrate uptake can be represented by

$$\begin{aligned} R_n &= (R_p - R_a) \frac{[NO_3^-]}{[NO_3^-] + k_m} \\ &= R_p \frac{k_m}{[NH_4^+] + k_m} \frac{[NO_3^-]}{[NO_3^-] + k_m}. \end{aligned} \quad (4)$$

In this equation  $R_p$ ,  $R_a$ , and  $R_n$  are the plant uptake rates for nitrogen, ammonium (Reaction R13), and nitrate (Reaction R14).  $R_p$  is calculated in the CLM and input to PFLOTRAN as a constant during each CLM time step. Equation (4) essentially assumes an inhibition of ammonium on nitrate uptake, which is consistent with the observation that plant nitrate uptake rate remains low until ammonium concentrations drop below a threshold (Eltrop and Marschner, 1996). However, the form of the rate expression will differ for different plants (Pfautsch et al., 2009; Warren and Adams, 2007; Nordin et al., 2001; Falkengren-Grerup, 1995; Gherardi et al., 2013), which will require different representations in future developments.

CLM uses a demand-based competition approach (Appendix A, Sect. A5) to represent the limitation of 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 limiting the consumption

rates if supply is less than demand 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 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 the Monod function with a residual concentration for individual reactions is easier to implement and allows for more flexibility in adding reactions.

### 2.3 Scaling, clipping, and log transformation for avoiding negative concentration

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

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

where

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

for negative  $\delta c^{k+1,p+1}(i)$  with  $m$  as the number of species times the number of numerical grid cells and  $\alpha$  as a factor with default value of 0.99. A second approach, used by RTM codes STOMP, HYDROGEOCHEM 5.0, and TOUGH-REACT, is clipping (i.e., for any  $\delta c^{k+1,p+1}(i) \geq c^{k+1,p}(i)$ ,  $\delta c^{k+1,p+1}(i) = c^{k+1,p}(i) - \epsilon$  with  $\epsilon$  as a small number, e.g.,  $10^{-20}$ ). This limits the minimum concentration that can be modeled.

A third approach, log transformation, also ensures a 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 state changes without the need to use variable switching. Instead of solving Eq. (B3) for  $c^{k+1}$  using Eqs. (B4)–(B6), this method solves for  $(\ln c^{k+1})$  (Hammond, 2003) with

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

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

and

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

### 3 Tests, results, and discussions

The Newton–Raphson method and scaling, clipping, and log transformation are widely used and extensively tested for RTMs, but not for coupled LSM–RTM applications. The CLM describes biogeochemical dynamics 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<sup>-3</sup> carbon in an 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 simulations reveal some numerical issues (errors, divergence, and small time-step sizes) that were not widely reported. We identify the issues from coupled CLM–PFLOTRAN simulations and reproduce them in simple test problems. We examine remedies in the simple test problems and test them in the coupled simulations.

For tests 1–3, we start with plant ammonium uptake to examine the numerical solution for Monod function, and then add nitrification and denitrification incrementally to assess the implications of adding reactions. For test 4, we check the implementation of mineralization and immobilization in the decomposition reactions. Third, we compare the nitrogen demand partition into ammonium and nitrate between CLM and PFLOTRAN. With coupled CLM–PFLOTRAN spin-up simulations for Arctic, temperate, and tropical sites, we assess the application of scaling, clipping, and log transformation to achieve accurate, efficient, and robust simulations. Spreadsheets, PFLOTRAN input files, and additional materials are provided as Supplement, and archived at Tang et al. (2016).

Our implementation of CLM soil biogeochemistry introduces mainly two parameters: half saturation ( $k_m$ ) and residual concentration. A wide range of  $k_m$  values was reported for ammonium, nitrate, and organic nitrogen for microbes and plants. The median, mean, and standard deviations range from 10 to 100, 50 to 500, and 10 to 200 μM, respectively (Kuzyakov and Xu, 2013). Reported residual concentrations are limited – likely because of the detection limits of the analytical methods – and are considered to be zero (e.g., Høgh-Jensen et al., 1997). The detection limits are usually at the micromolar level, while up to the nanomolar level was reported (Nollet and De Gelder, 2013). In Ecosys, the  $k_m$  is 0.40 and 0.35 g N m<sup>-3</sup>, and the residual concentration is 0.0125 and 0.03 g N m<sup>-3</sup> (Grant, 2013) for ammonium and nitrate for microbes, respectively. We start with  $k_m = 10^{-6}$  M or mol m<sup>-3</sup>, and residual concentration =  $10^{-15}$  M or mol m<sup>-3</sup> for plants and microbes. To further investigate the nonphysical solution negativity for the current study and for future application for other reactants (e.g.,  $H_2$  and  $O_2$ ) where the 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 Simple tests

#### 3.1.1 Plant nitrogen uptake, nitrification, and denitrification

It was observed that plants can decrease nitrogen concentration to below the detection limit in hours (Kamer et al., 2001). 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 ( $R_p$ ) 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 PFLOTRAN. We use the Monod function to represent the limitation of nitrogen availability on uptake. Once the plant nitrogen uptake is calculated in PFLOTRAN, it is returned to CLM, which allocates the uptake among, leaf, live stem, dead stem, etc., and associated storage pools (Oleson et al., 2013). We examine the numerical solutions for the Monod equation at first. 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.

#### Plant ammonium uptake (test 1)

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

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

Discretizing it in time using the backward Euler method for a time-step size  $\Delta t$ , a solution is

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

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 with the Monod function mathematically ensures  $[\text{NH}_4^+]^{k+1} \geq [\text{NH}_4^+]_r$ .

We use a spreadsheet to examine the Newton–Raphson iteration process for solving Eq. (10) and the application of clipping, scaling, and log transformation (test1.xlsx in the Supplement). When an overshoot gets the concentration closer to the negative than the positive root (Eq. 11), the iterations converge to the nonphysical negative semianalytical solution (spreadsheet case3). This can be avoided by using clipping, scaling, or log transformation (spreadsheet case4, case6, case8).

Even though clipping avoids convergence to the negative solution, the ammonium consumption is clipped, but the PlantA (Reaction AR13) production is not clipped (spreadsheet case5), violating the reaction stoichiometry. This results in mass balance errors for explicit time stepping (Tang and Riley, 2016). For implicit time stepping, additional iterations can resolve this violation to avoid mass balance error. However, if a nonreactive species is added with a concentration of  $1000 \text{ mol m}^{-3}$  (e.g., soil organic matter pool 4 SOM4 in organic layer), the relative update  $s_{\text{norm rel}} = \|\delta c^{k+1,p}\|_2 / \|c^{k+1,p}\|_2$  decreases to  $9.3 \times 10^{-9}$  in the iteration (spreadsheet case5a). With a relative update tolerance (STOL; Eq. B9) of  $10^{-8}$ , the iteration would be deemed converged. This false convergence may produce 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 a tight STOL to avoid this false convergence so that additional iterations can be taken to solve the residual equation accurately.

In contrast to clipping, scaling applies the same scaling factor to limit both ammonium consumption and PlantA production following the stoichiometry of the reaction (spreadsheet case6). However, if we add a production reaction that is independent of plant ammonium uptake, say nitrate deposition, which can come from CLM as a constant rate in a time step rather than an internally balanced reaction, scaling reduces the plant ammonium uptake to account for the availability of ammonium as we intend, but the nitrate deposition rate is also reduced by the same scaling factor even though it is not limited by the availability of either ammonium or nitrate (spreadsheet case7). Like clipping, this unintended consequence can be resolved in the subsequent iterations, and a loose STOL may lead to false convergence and mass balance errors.

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 challenges for the log transformation method even though PlantA is only a product. When it is zero, the Jacobian 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 underflow of the exponential function (spreadsheet caseA, as a 64-bit real number (“double precision”), is precise to 15 significant digits and has a range of  $e^{-709}$  to  $e^{709}$  in IEEE 754 floating-point representation; Lemmon and Schafer, 2005). Clipping the update (say to be between  $-5$  and  $5$ ) is needed to prevent numerical overflow or underflow. Like the cases with clipping without log transformation, clipping violates reaction stoichiometry in the clipping iteration, and this violation needs to be resolved in subsequent iterations (spreadsheet 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 or singular when log transformation is used, and clipping is needed to guard against overflow or underflow of the exponential function; (4) clipping limits the consumption, but not the corresponding production, violating reaction stoichiometry in the iteration; (5) production reactions with external sources are inhibited in the iterations when scaling is applied, which is unintended; (6) additional iterations can resolve issues in (4) or (5); and (7) loose update tolerance convergence criteria may cause false convergence and result in mass balance errors for clipping and scaling.

### Plant ammonium uptake and nitrification (test 2)

Adding a nitrification Reaction (R10) with a first-order rate to the plant ammonium uptake Reaction (R13), Eq. (10) becomes

$$\begin{aligned} \frac{d[NH_4^+]}{dt} &= -R_a \frac{[NH_4^+]}{[NH_4^+] + k_m} \\ -k_{nitr}[NH_4^+] &= -R_{at} - R_{nitr}. \end{aligned} \quad (12)$$

A semianalytical solution similar to Eq. (11) can be derived for Eq. (12). With  $J_{at} = dR_{at}/d[NH_4^+] = R_a k_m ([NH_4^+] + k_m)^{-2}$ , and  $J_{nitr} = dR_{nitr}/d[NH_4^+] = k_{nitr}$ , the matrix equation Eq. (B5) becomes

$$\begin{bmatrix} 1/\Delta t + J_{at} + J_{nitr} & 0 & 0 \\ -J_{at} & 1/\Delta t & 0 \\ -J_{nitr} & 0 & 1/\Delta t \end{bmatrix} \begin{pmatrix} \delta[NH_4^+]^{k+1,1} \\ \delta[PlantA]^{k+1,1} \\ \delta[NO_3^-]^{k+1,1} \end{pmatrix} = - \begin{pmatrix} R_{at} + R_{nitr} \\ -R_{at} \\ -R_{nitr} \end{pmatrix}, \quad (13)$$

for the first iteration. As  $R_{at} + R_{nitr} \geq 0$ , the ammonium update is negative even when the ammonium concentration is not very low. The off-diagonal terms for PlantA and nitrate in the Jacobian matrix bring the negative ammonium update into the updates for PlantA and nitrate even though there is no reaction that consumes them. Specifically,

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

$$\begin{aligned} \frac{\delta[NO_3^-]^{k+1,1}}{\Delta t} &= -\frac{J_{nitr}}{\frac{1}{\Delta t} + J_{at} + J_{nitr}} R_{at} \\ + \frac{\frac{1}{\Delta t} + J_{at}}{\frac{1}{\Delta t} + J_{at} + J_{nitr}} R_{nitr}. \end{aligned} \quad (15)$$

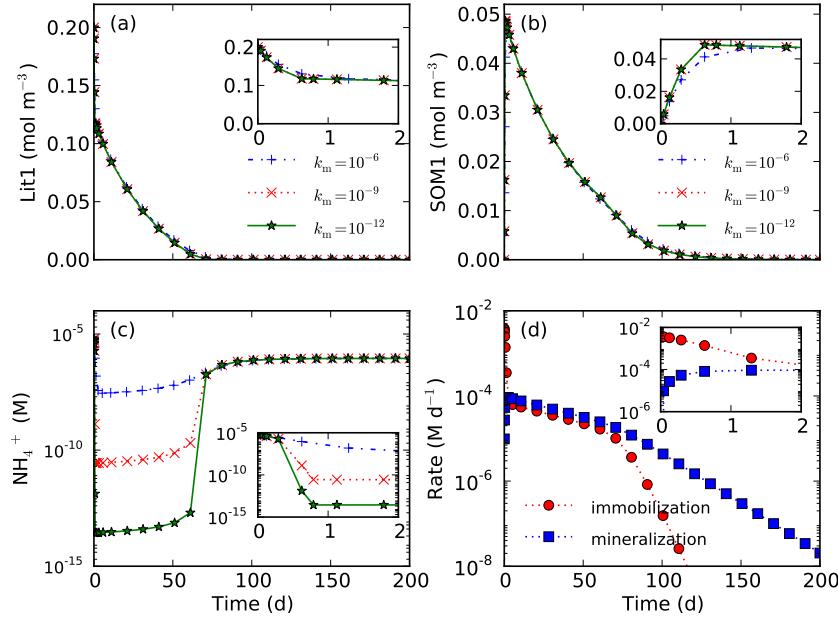
Depending on the rates ( $R_{at}$ ,  $R_{nitr}$ ), derivatives ( $J_{at}$ ,  $J_{nitr}$ ), and time-step size  $\Delta t$ , the update can be negative for PlantA and nitrate, producing a zero-order “numerical consumption”, in which the limitation of availability is not explicitly represented. This has implications for the scaling method.

The scaling factor ( $\lambda$ ) is a function of not only the update but also the concentration (Eq. 6). If a negative update is produced for a zero concentration, the scaling factor is zero, decreasing the scaled update to zero. The iteration converges without any change to the concentrations, numerically stopping 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 test1.xlsx case6 (Supplement) to create test2.xlsx (Supplement) to demonstrate that a small enough initial concentration relative to the negative update may numerically inhibit all of the reactions as well. An update of  $-6.6 \times 10^{-6} \text{ M}$  is produced for nitrate (spreadsheet scale1). When the initial nitrate concentration  $[NO_3^-]_0$  is not too small, say  $10^{-6} \text{ M}$ , the solution converges to the semianalytical solution in six iterations. When  $[NO_3^-]_0$  is decreased to  $10^{-9} \text{ M}$ , the relative update snorm<sub>rel</sub> is  $9.2 \times 10^{-10}$ . If STOL =  $10^{-9}$ , the solution is deemed converged as Eq. (B9) is met, but not to the positive semianalytical 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 update to below STOL to reach false convergence. If we tighten STOL to  $10^{-30}$ , the iterations continue, with decreasing nitrate concentration,  $\lambda$ , and snorm<sub>rel</sub> by 2 orders of magnitude ( $1 - \alpha$  as default  $\alpha = 0.99$ ) in each iteration, until snorm<sub>rel</sub> reaches  $10^{-30}$  (spreadsheet scale2). Unless STOL is sufficiently small, or MAXIT (the maximum number of iterations before stopping the current iterations and cutting the time step) is small (Appendix B), false convergence is likely to occur for the scaling method. The impact of “numerical 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.

In summary, this test problem demonstrates that (1) a negative update can be produced even for products during a Newton–Raphson iteration; and (2) when a negative update is produced for a very low concentration, a very small scaling factor may numerically inhibit all of the reactions due to false convergence even with very tight STOL.

### Plant uptake, nitrification, and denitrification (test 3)

The matrix and update equations with plant nitrate uptake and denitrification added to test 2 are available in Appendix C. In addition to nitrate and PlantA, PlantN (Reaction AR14) and the denitrification product nitrogen gas may have negative 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 nitrate concentration is added in the Jacobian matrix for PlantN (Eq. C1). As a result, a negative update for both



**Figure 2.** 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**).

ammonium and nitrate contributes to negative PlantN update through the two nonzero off-diagonal terms. Therefore, the likelihood for a negative update to PlantN is greater than PlantA as the former is influenced by more rates and derivatives. We add plant nitrate uptake and denitrification into test2.xlsx (Supplement) and assess the implications of increased reactions and complexity in test3.xlsx (Supplement). In addition to nitrate, this introduces a negative update for nitrogen gas in the first iteration (spreadsheet scale1). As the iterations resolve the balance between nitrite production from nitrification, and consumption due to plant uptake and denitrification, update to PlantN becomes negative and eventually leads to false convergence. The time-step size needs to be decreased from 1800 to 15 s to resolve the false convergence (spreadsheet scale2). In contrast, the added reactions have less impact on the clipping and log transformation methods.

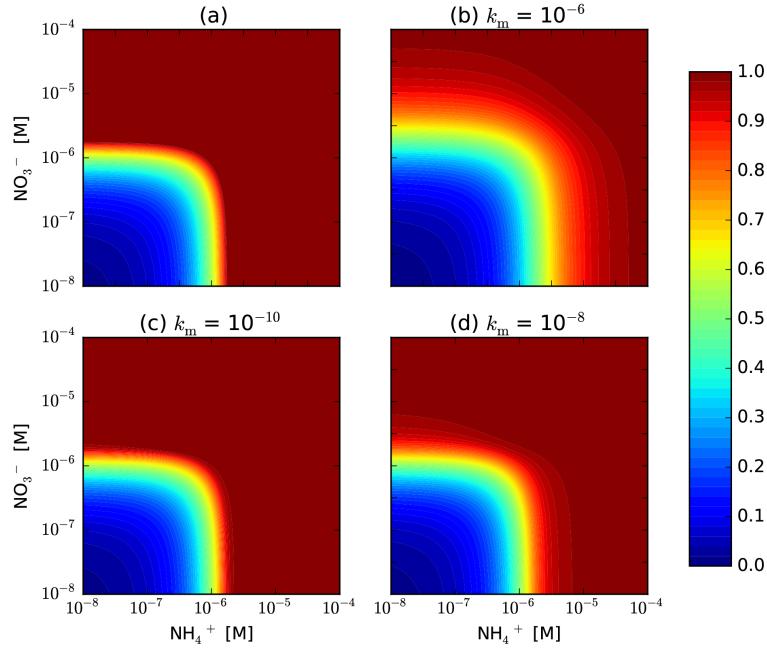
### 3.1.2 Nitrogen immobilization and mineralization during decomposition (test 4)

We examine another part of the reaction network: decomposition, nitrogen immobilization, and mineralization (Fig. 1). We consider a case of decomposing 0.2 M Lit1C + 0.005 M Lit1N to produce SOM1 with initial 4  $\mu$ M ammonium using the reactions (R2) and (R7) in the CLM-CN reaction network (Fig. 1). We use PFLOTRAN with a water-saturated grid cell with porosity of 0.25. At the beginning, Lit1 decreases and SOM1 increases sharply because the rate coefficient for Lit1 is 16 times that for SOM1 (Fig. 2a and b). As ammonium concentration decreases by orders of magnitude

because of the faster immobilization than mineralization rate (Fig. 2c and 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. As the immobilization rate decreases with decreasing Lit1, ammonium concentration rebounds after Lit1 is depleted. For  $k_m$  values of  $10^{-6}$ ,  $10^{-9}$ , and  $10^{-12}$  M, Lit1 and SOM1 dynamics are similar except for slight differences in the early transit periods, but the ammonium values are decreased to  $\sim 10^{-8}$ ,  $10^{-11}$ , and  $10^{-14}$  M, respectively. Smaller  $k_m$  values result in lower ammonium concentrations, which have implications for the clipping, scaling, and log transformation methods as discussed in tests 1–3.

### 3.1.3 Nitrogen demand partitioning between ammonium and nitrate

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 implemented in Eqs. (3) and (4). As an example, we consider uptake  $R_p = 10^{-9}$  M s<sup>-1</sup> from a solution with various  $[NH_4^+]$  and  $[NO_3^-]$  for a 0.5 h time step. With CLM,  $f_{pi} = 1$  when  $[NH_4^+] + [NO_3^-] \geq R_p\Delta t$ ; otherwise, it decreases with decreasing  $[NH_4^+] + [NO_3^-]$  (Fig. 3). The new representation (Eqs. 3 and 4) is generally similar, with  $f_{pi} = 1$  or 0 when  $[NH_4^+] + [NO_3^-] \gg k_m$  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 decreas-



**Figure 3.**  $f_{pi} = \frac{N_{take}}{N_{demand}} = \max \left( 1, \frac{[NH_4^+] + [NO_3^-]}{N_{demand}} \right)$  (a) vs.  $= \frac{[NH_4^+]}{k_m + [NH_4^+]} + \left( 1 - \frac{[NH_4^+]}{k_m + [NH_4^+]} \right) \frac{[NO_3^-]}{k_m + [NO_3^-]}$  (b–d) in a 0.5 h time step with an uptake rate of  $10^{-9} \text{ M s}^{-1}$ .  $f_{pi}$  for the latter representation is less than or equal to that for the first one. The difference decreases with decreasing half saturation  $k_m$ .

ing  $k_m$ , apparently disappearing at  $k_m = 10^{-10}$ . Various levels 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), which is similar to microbial uptake of inorganic and organic nitrogen species (Foulland 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 abundant, nitrate will not be taken by plants. The new scheme allows the level of preference to be adjusted by varying  $k_m$ ; more realistic representations can be implemented relatively easily.

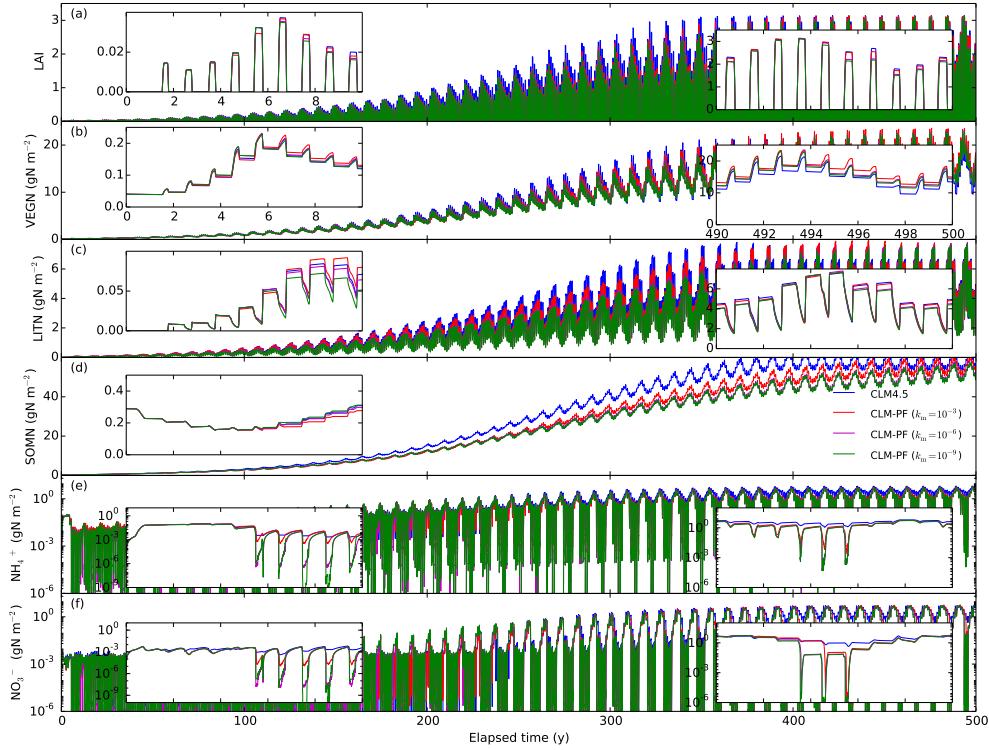
### 3.2 CLM–PFLOTRAN simulations

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 in which 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  $10^6 \text{ m}$ , `o_scalar` to 1, (2) `cryoturb_diffusion`, (3) `som_diffusion`, and (4) `sf_no3` and `sf_sminn` 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 durations are 1000, 600, and 600 years for the Arctic, temperate, and tropical sites, respectively. In the base case,  $k_m = 10^{-6} \text{ mol m}^{-3}$  and residual concentration is  $10^{-15} \text{ mol m}^{-3}$ . 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. The simulations are conducted using the ORNL Institutional Cluster (OIC Phase5, an SGI Altix with dual 8-core AMD Opterons per compute node) with CLM–PFLOTRAN (as well as third-party libraries MPICH, PETSc, NetCDF, HDF5, etc.) compiled with gfortran 4.8.1 with the “-O1” optimization level. Due to the small size of the simulations, our tests use only a single CPU core.

#### 3.2.1 Site descriptions

The US-Brw site ( $71.35^\circ \text{ N}$ ,  $156.62^\circ \text{ W}$ ) is located near Barrow, Alaska. The mean annual temperature, precipitation, and snowfall are  $-12^\circ \text{ C}$ , 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



**Figure 4.** Calculated LAI and nitrogen distribution among vegetation, litter, SOM,  $\text{NH}_4^+$ , and  $\text{NO}_3^-$  pools in spin-up simulations for the US-Brw site.

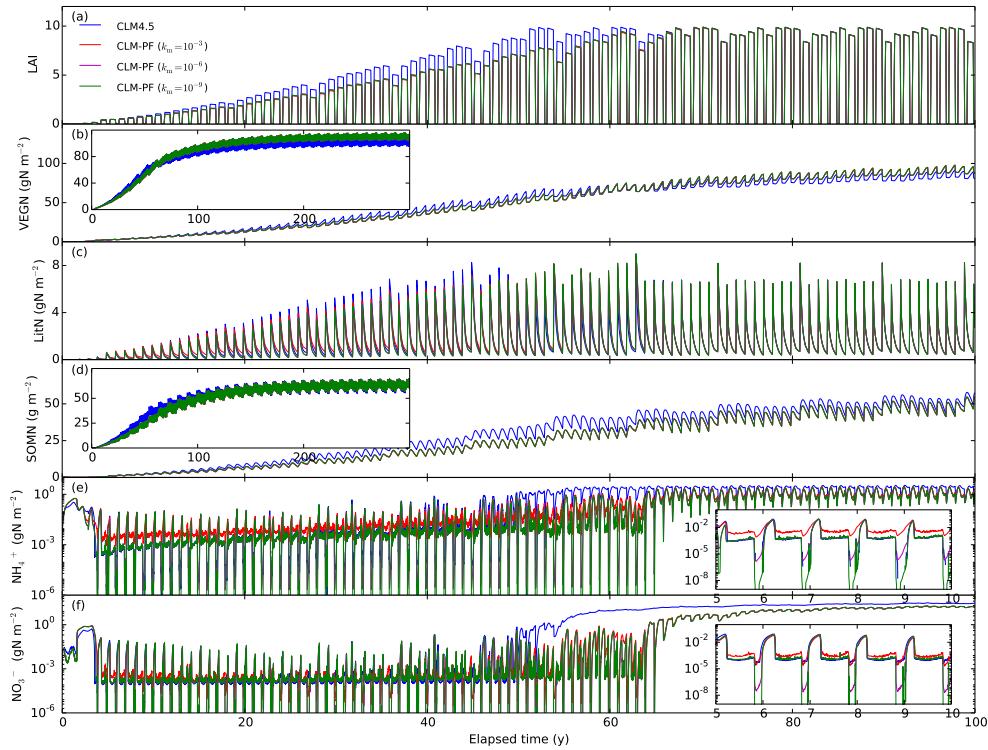
variable in length but generally begins in early June and lasts until early September (Hinkel and Nelson, 2003). The area is composed of several different representative wet–moist coastal sedge tundra types, including wet sedges, grasses, moss, and assorted lichens. The leaf area index (LAI) is  $\sim 1.1$  (AmeriFlux data).

The US-WBW site ( $35.96^\circ \text{ N}$ ,  $84.29^\circ \text{ W}$ ) 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  $\sim 139 \text{ cm}$ , and the mean median temperature is  $14.5^\circ \text{ C}$ . 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 agricultural land 50 years ago. LAI is  $\sim 6.2$  (Hanson et al., 2004).

The BR-Cax site ( $-1.72^\circ \text{ N}$ ,  $-51.46^\circ \text{ W}$ ) is located in the eastern Amazon tropical rainforest. The mean annual rainfall is between 200 and 250 cm, 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  $\sim 4\text{--}6$  (Powell et al., 2013).

### 3.2.2 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. 4–6). Overall, CLM–PFLOTRAN is close to CLM4.5 in predicting LAI and nitrogen distribution among vegetation, litter, SOM (soil organic matter), 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 periods (200–400 years for the Arctic, and 50–70 years for the temperate and tropical sites) for SOMN (soil organic matter nitrogen), ammonium, and nitrate. This is not surprising as (1) the nitrogen demand competition scheme implemented in CLM–PFLOTRAN is different from that in CLM4.5 (Fig. 3), (2) the former solves the reaction network simultaneously while the latter does so sequentially (resolve the plant uptake and decomposition first, then nitrification, then denitrification), and (3) the carbon nitrogen cycle is very sensitive to the nitrogen competition representation. Close to steady state, both CLM4.5 and CLM–PFLOTRAN overpredict the LAI at the Arctic and temperate sites, and underpredict soil organic matter accumulation, which will be resolved in future work.



**Figure 5.** Calculated LAI and nitrogen distribution among vegetation, litter, SOM,  $\text{NH}_4^+$ , and  $\text{NO}_3^-$  pools in spin-up simulations for US-WBW site.

The Arctic site shows a distinct summer growing season (Fig. 4): LAI and VEGN (vegetation nitrogen) jump up at the beginning, then level off, and drop down at the end of the growing season when LITN (litter nitrogen) jumps up due to litter fall. Ammonium and nitrate concentrations drop to very low levels at the beginning of the growing season and accumulate at 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 variation to reach equilibrium.

The higher  $k_m$  of  $10^{-3} \text{ mol m}^{-3}$  results in lower immobilization, higher accumulation of LITN, and higher ammonium and nitrate concentrations than  $k_m$  of  $10^{-6} \text{ mol m}^{-3}$  for the tropical site during the spin-up (Fig. 6). This is not surprising because a higher  $k_m$  of  $10^{-3} \text{ mol m}^{-3}$  numerically poses a stricter limitation on the extent that plants and microbes can take from soils. The range of  $k_m$  values ( $10^{-6}$ , and  $10^{-9} \text{ mol m}^{-3}$ ) 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. 4e and f and 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 low concentrations (test 4), and weak downregulation and steep transition between zero order and first order, it has implications on accuracy and efficiency of the numerical solutions.

### 3.2.3 Accuracy and efficiency

Numerical errors introduced due to false convergence 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 reports  $\geq 10^{-8} \text{ g m}^{-2}$  errors (to reduce log file size as the simulation durations are hundreds of years and the time-step size is half an hour). When log transformation is used, mass balance errors are not reported for the Arctic, temperate, and tropical sites with  $k_m$  values  $10^{-3}$ ,  $10^{-6}$ , and  $10^{-9} \text{ mol m}^{-3}$ . The computing time for CLM–PFLOTRAN is about 60–100 % more than that of CLM (Table 1). This is not unreasonably high as the implicit method involves solving a Jacobian system for each Newton–Raphson step (Eq. B5), and log transformation converts the linear problem into a nonlinear one. 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.

Mass balance errors are reported for  $k_m$  values of  $10^{-6}$ , and  $10^{-9}$  but not for  $10^{-3} \text{ mol m}^{-3}$  when clipping is applied. With  $k_m = 10^{-3} \text{ mol m}^{-3}$ , the plant uptake and immobilization are inhibited at relatively high concentration so that nitrogen concentrations are high. With  $k_m$  decreasing from  $10^{-6}$  to  $10^{-9} \text{ mol m}^{-3}$ , nitrogen concentrations are lowered to a much lower level (Figs. 4–6, similar to Fig. 2c), increasing the likelihood of overshoot. Mass balance errors are reported when the relative update is below STOL, preventing further iterations from resolving the violation of reaction stoichiometry 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  $10^{-8} \text{ g m}^{-3}$  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  $k_m = 10^{-9} \text{ mol m}^{-3}$ . Tightening STOL only slightly increases the computing time. Because clipping often occurs at very low concentrations, the reported mass balance errors are usually small ( $\sim 10^{-8} \text{ g N m}^{-2}$  to  $\sim 10^{-7} \text{ g N m}^{-2}$ ), and do not have substantial impact on the overall simulation results.

The results for scaling is similar to clipping: mass balance errors are reported for  $k_m$  values of  $10^{-6}$  and  $10^{-9}$  but not for  $10^{-3} \text{ mol m}^{-3}$ ; tightening STOL to  $10^{-12}$  eliminates these errors; it takes 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  $10^{-8} \text{ mol m}^{-3}$  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 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 5 months (Fig. 6e and 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 a reset concentration of  $10^{-8} \text{ mol m}^{-3}$ , the minimum nitrous oxide concentration for the 10 layers is  $10^{-8} \text{ mol m}^{-3}$ , and ammonium concentrations show two peaks followed by two drops due to the two plant uptake peaks every day. Decreasing the reset concentration to  $10^{-10} \text{ mol m}^{-3}$ , the minimum concentration drops to  $10^{-12}$ ,  $10^{-14}$ , and  $10^{-16} \text{ mol m}^{-3}$ , corresponding to 1, 2, and 3 scaling iterations with overshoot for nitrous oxide. These result in numerical “inhibition” of nitrogen rebound every day. It worsens with further decrease of the reset concentration to  $10^{-12} \text{ mol m}^{-3}$ . This introduces mass balance errors as reported in CLM because the false convergence numerically inhibits all of the reactions including nitrogen deposition and litter input from CLM to PFLO-

TRAN. Unlike clipping, these false convergences introduce excessive mass balance errors because of the inhibition of productions specified from CLM. If all of the reactions are internally balanced, false convergence does not result in mass balance errors.

The frequent negative update to nitrous oxide is produced because the rate for the nitrification Reaction (R11) is parameterized as a fraction of the net nitrogen mineralization rate to reflect the relationship between labile carbon content and nitrous oxide 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 reactions bring in ammonium and nitrate limitation, and ammonium inhibition on nitrate immobilization. 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. Negative updates to all of these species contribute to negative updates to nitrous oxide. Similar to tests 2–3, a negative update to a low nitrous oxide concentration can decrease snorm<sub>rel</sub> to below STOL, resulting in false convergence and mass balance errors. 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 and computational cost due to the reduced sparsity of the 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 reactions can be caused by negative updates to very low concentrations for any species.

The numerical errors can be decreased and eliminated by decreasing STOL. Similar to the tests 2–3 a small STOL can result in small snorm<sub>rel</sub>, then a very small STOL is needed. For the case of reset concentration  $10^{-10}$  for the 1 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. 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 doubles for the log transformation.

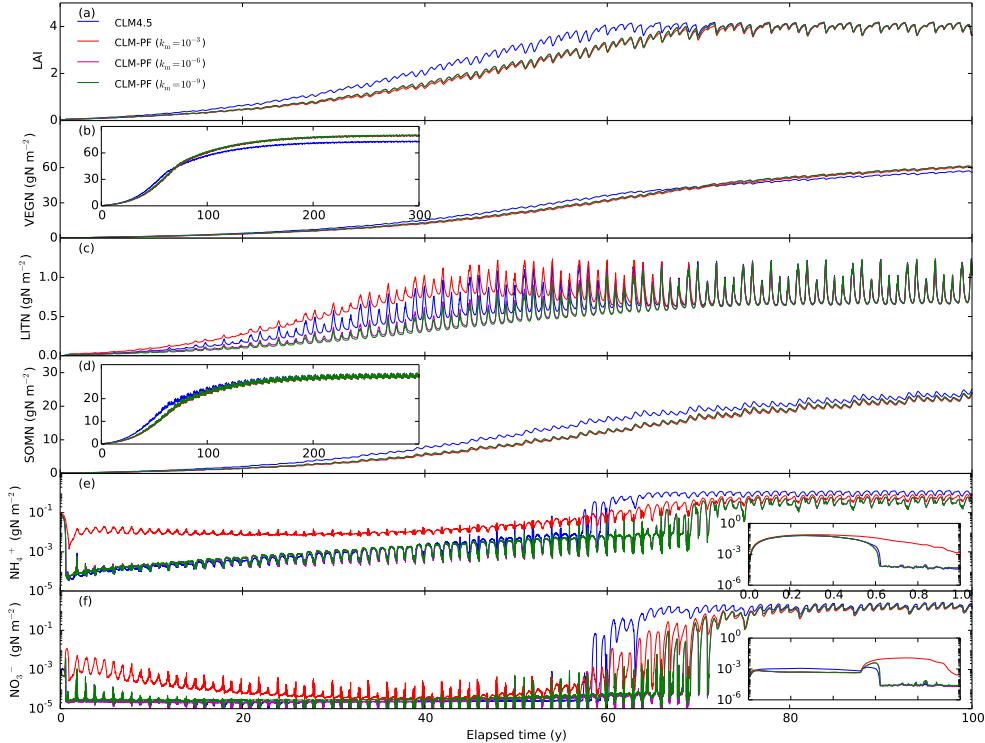
#### 4 Summary and conclusions

Global land surface models have traditionally represented subsurface soil 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 im-

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

Site	Clipping			Scaling			Log transformation		
	$k_m$	$10^{-3}$	$10^{-6}$	$10^{-9}$	$10^{-3}$	$10^{-6}$	$10^{-9}$	$10^{-3}$	$10^{-6}$
US-Brw	1.28	1.30	1.30	1.29	1.29	1.32	1.45	1.49	1.72
US-WBW	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 h for the Arctic, temperate, and tropical sites for a simulation duration of 1000, 600, and 600 years.  $k_m$  is the half saturation ( $\text{mol m}^{-3}$ ).

**Figure 6.** Calculated LAI and nitrogen distribution among vegetation, litter, SOM,  $\text{NH}_4^+$ , and  $\text{NO}_3^-$  pools in spin-up simulations for BR-Cax site.

plement CLM-CN decomposition cascade, nitrification, denitrification, and plant nitrogen uptake reactions in CLM–PFLOTRAN. We illustrate that with implicit time stepping using the Newton–Raphson method, the concentration can become negative during the iterations even for species that have no consumption, which need to be prevented by intervening in the Newton–Raphson iteration procedure.

Simply stopping the iteration with negative concentration and returning to the time-stepping subroutine to cut time-step size can avoid negative concentration, but may result in small time-step sizes and high computational cost. Clipping, scaling, and log transformation can all prevent negative concentration and reduce computational cost but at the risk of accuracy. Our 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 sit-

uations, satisfying STOL does not guarantee satisfying the residual equation, and therefore it may introduce false convergence. Clipping reduces the consumption but not the production in some reactions, violating reaction stoichiometry. Subsequent iterations are required to resolve this violation. A tight STOL is needed to avoid false convergence and prevent mass balance errors. While the scaling method reduces the whole update vector following the stoichiometry of the reactions to maintain mass balance, a small scaling factor caused by a negative update to a small concentration may diminish the update and result in false convergence, numerically inhibiting all reactions, which is not intended for productions with external sources (e.g., nitrogen deposition from CLM to PFLOTRAN). 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 time. The computational cost increases with decreasing concentrations, most substantially for log transformation.

These computational issues arise because we switch from the explicit methods to the implicit methods for soil biogeochemistry. We use small half saturation (e.g.,  $10^{-9}$ ), residual concentration (e.g.,  $10^{-15}$ , slightly above machine zero), and large initial time-step size (e.g., 0.5 h) to exemplify the causes of the accuracy, efficiency, and robustness issues. For the zero-order rate formulae in CLM, the limitation of reactant (nitrogen in this work) availability needs to be explicitly represented for robustness and flexibility. With mechanistic representations, reaction stops or reverses when the rate-limiting reactants decrease to a threshold, or when the reaction becomes thermodynamically unfavorable, nullifying the need for half saturation and residual concentration. Before our representations are sufficiently mechanistic, a small residual concentration (say  $10^{-15}$ ) serves as a safeguard to avoid failure for the log transformation method and unnecessary efficiency and accuracy loss for the clipping and scaling methods unless a smaller residual concentration is dictated by physical and chemical conditions.

For reactions with very low half saturation and residual concentrations, e.g., redox reactions involving  $O_2$ ,  $H_2$ , and  $CH_4$ , STOL can be set to zero to avoid false convergence, at the risk of potential increased computational cost. Increasing STOL (say to  $10^{-12}$ ) decreases computing time at the risk of potential accuracy loss. To cover a wide range of many orders of magnitude concentrations in soil biogeochemistry and for accuracy and robustness, the modelers can start with the log transformation method. If it is desirable to reduce the computational time, STOL can be relaxed, and clipping can be used without log transformation. The scaling method is another option but with strict STOL requirement. As the accuracy is checked and logged in CLM for carbon and nitrogen mass balance, the modelers can assess the tradeoff between efficiency and accuracy in CLM–PFLOTRAN and select their optimum.

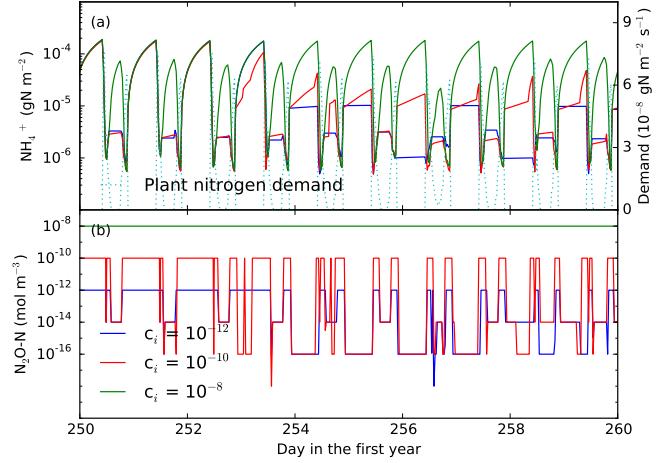
Our CLM–PFLOTRAN spin-up simulations at Arctic, temperate, and tropical sites produce results similar to CLM4.5, and indicate that accurate and robust solutions can be achieved with clipping, scaling, or log transformation. The computing time is 50 to 100 % more than CLM4.5 for a range of half saturation 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 indicate that accurate, efficient, and robust solutions for current CLM soil biogeochemistry can be achieved using CLM–PFLOTRAN. We thus demonstrate the feasibility of using an open-source, general-purpose reactive transport code with CLM, enabling significantly more complicated 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 advect-

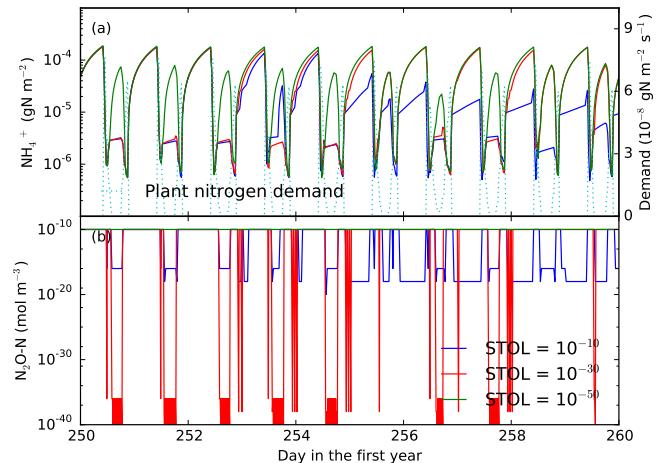
tion, diffusion, and reaction equations directly in the LSM. This has been done using explicit time stepping and operator splitting to simulate the transport and transformation of carbon, nitrogen, and other species in CLM (Tang et al., 2013). An advantage of our approach of using a community RTM to 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 advection–dispersion–reaction equations (Yeh and Tripathi, 1989; Zheng and Bennett, 2002; Steefel et al., 2015). 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.

## Code availability

PFLOTRAN is 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/pfplotran>. CLM–PFLOTRAN is under development and will be made available according to the NGEE-Arctic Data Management Policies and Plans (<http://ngee-arctic.ornl.gov/content/ngee-arctic-data-management-policies-and-plans>). Before it becomes publicly available, please contact the corresponding authors to obtain access to the code.



**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 the  $y$  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.

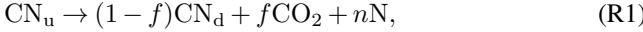


**Figure 8.** Decreasing  $\text{STOL}$  can decrease and eliminate the numerical inhibition in the case of  $c_i = 10^{-10} \text{ mol m}^{-3}$  in Fig. 7.  $\text{N}_2\text{O}-\text{N}$  concentration in the  $y$  axis in (b) is the minimum of the 10 soil layers.

## 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 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



where  $\text{CN}_u$  and  $\text{CN}_d$  are the upstream and downstream pools (molecular formula, for 1 mol upstream and downstream pool, there is  $u$  and  $d$  mol N),  $N$  is either  $\text{NH}_4^+$  or  $\text{NO}_3^-$ ,  $f$  is the respiration fraction, and  $n = u - (1-f)d$ . The rate is

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

where  $k_d$  is the rate coefficient, and  $f_T$  and  $f_w$  are the temperature and moisture response functions. With a constant CN ratio, the decomposition reactions for the four SOM pools are



and



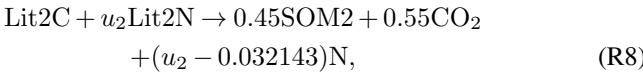
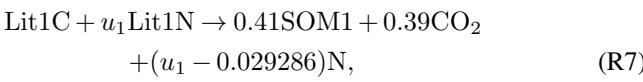
The exact stoichiometric coefficients are calculated in the code using values for respiration factor, CN ratio, and molecular weight specified in the input file.

CLM4.5 has an option to separate N into  $\text{NH}_4^+$  and  $\text{NO}_3^-$ . The N mineralization product is  $\text{NH}_4^+$ .

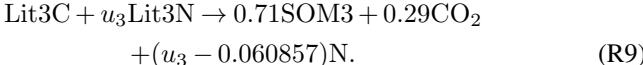
As the CN ratio is variable for the three litter pools, litter N pools need to be tracked such that Reaction (R1) becomes



where  $u = [\text{LitN}] / [\text{LitC}]$ . 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  $\text{NH}_4^+$ ,  $\text{NO}_3^-$ , or both.

### A2 Nitrification

The nitrification reaction to produce  $\text{NO}_3^-$  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  $\text{N}_2\text{O}$  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 Reactions (R2)–(R5) and (R7)–(R9). The second component is (Parton et al., 1996)

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

Ignoring the high-order terms and moving the unit conversion factor into  $k_{\text{N}_2\text{O}}$ , it can be simplified as a first-order rate as

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

### A3 Denitrification

The denitrification reaction is



with rate (Dickinson et al., 2002)

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

### A4 Plant nitrogen uptake

The plant nitrogen uptake reaction can be written as



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 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  $\text{NH}_4^+$  demand; and  $R_{n,tot}$  as the total  $\text{NO}_3^-$  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 to potential demands, and  $R_{n,tot} = 0$ ; otherwise, the uptakes for  $\text{NH}_4^+$  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  $R_{n,tot}$  is met with available  $\text{NO}_3^-$ . Otherwise, available  $\text{NO}_3^-$  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,

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

where  $c$  is the concentration vector,  $R$  is the kinetic reaction rate, and  $F$  is the fluxes (e.g., nitrogen deposition). Discretizing Eq. (B1) in time using the backward Euler method,

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

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

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

and Jacobian as

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

where  $p$  is the iteration counter, the update is

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

and the iteration equation is

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

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

Specifically,

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

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

or

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

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 test 3

Adding to test 2 a plant  $\text{NO}_3^-$  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 a 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

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

**The Supplement related to this article is available online at doi:10.5194/gmd-0-1-2016-supplement.**

**Author contributions.** G. Bisht, B. Andre, R. T. Mills, J. Kumar, and F. M. Hoffman developed the CLM-PFLOTRAN framework that this work is built upon. F. Yuan, G. Tang, G. Bisht, and X. Xu added biogeochemistry to the CLM-PFLOTRAN interface. F. Yuan proposed the nitrification and denitrification reactions and rate formulae. G. Tang, F. Yuan, and X. Xu implemented the CLM soil biogeochemistry in PFLOTRAN under guidance of G. E. Hammond, P. C. Lichtner, S. L. Painter, and P. E. Thornton. G. Tang prepared the manuscript with contributions from all co-authors. G. Tang, F. Yuan, G. Bisht, and G. E. Hammond contributed equally to the work.

**Acknowledgements.** Thanks to Nathaniel O. Collier at ORNL for many discussions that contributed significantly to this work. Thanks to Kathie Tallant and Kathy 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. Development of CLM-PFLOTRAN was partially supported by the ORNL Laboratory Directed Research and Development (LDRD) program. ORNL is managed by UT-Battelle, LLC, for the U.S. Department of Energy under contract DE-AC05-00OR22725.

**Disclaimer.** 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, worldwide 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>).

Edited by: C. Sierra

## References

- 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), *Glob. Change Biol.*, 19, 957–974, doi: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 Gross, S. J. D.: Modeling denitrification in terrestrial and aquatic ecosystems at regional scales, *Ecol. Appl.*, 16, 2123–2142, doi:10.1890/1051-0761(2006)016[2123:MDITAA]2.0.CO;2, 2006.

Conrad, R.: Soil microorganisms as controllers of atmospheric trace gases ( $\text{H}_2$ , CO,  $\text{CH}_4$ , OCS,  $\text{N}_2\text{O}$ , and NO), *Microbiol. Rev.*, 60, 609–640, available at: <http://mmbi.asm.org/content/60/4/609.full.pdf> (last access: 14 December 2015), 8987358[pmid], 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 an eastern Amazonian rainforest, *New Phytol.*, 187, 579–591, doi: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, *J. Climate*, 15, 278–295, doi:10.1175/1520-0442(2002)015<0278:NCOCME>2.0.CO;2, 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 Phytol.*, 133, 469–478, doi:10.1111/j.1469-8137.1996.tb01914.x, 1996.

Falkengren-Grerup, U.: Interspecies differences in the preference of ammonium and nitrate in vascular plants, *Oecologia*, 102, 305–311, doi: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, 2013.

Fennell, D. E. and Gossett, J. M.: Modeling the production of and competition for hydrogen in a dechlorinating culture, *Environ. Sci. Technol.*, 32, 2450–2460, doi:10.1021/es980136l, 1998.

Fouilland, E., Gosselin, M., Rivkin, R. B., Vasseur, C., and Mostajir, B.: Nitrogen uptake by heterotrophic bacteria and phytoplankton in Arctic surface waters, *J. Plankton Res.*, 29, 369–376, doi:10.1093/plankt/fbm022, 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, 2013.

Grant, R. F.: Modelling changes in nitrogen cycling to sustain increases in forest productivity under elevated atmospheric  $\text{CO}_2$  and contrasting site conditions, *Biogeosciences*, 10, 7703–7721, doi:10.5194/bg-10-7703-2013, 2013.

Gu, C. and Riley, W. J.: Combined effects of short term rainfall patterns and soil texture on soil nitrogen cycling – a modeling analysis, *J. Contam. Hydrol.*, 112, 141–154, doi:10.1016/j.jconhyd.2009.12.003, 2010.

Hammond, G. E.: Innovative Methods for Solving Multicomponent Biogeochemical groundwater Transport on Supercomputers, Thesis, University of Illinois, Urbana-Champaign, 2003.

Hammond, G. E., Lichtner, P. C., and Mills, R. T.: Evaluating the performance of parallel subsurface simulators: an illustrative example with PFLOTRAN, *Water Resour. Res.*, 50, 208–228, doi:10.1002/2012WR013483, 2014.

- Hanson, P. and Wullschleger, S.: North American Temperate Deciduous Forest Responses to Changing Precipitation Regimes, Springer Verlag, 2003.
- 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, *Ecol. Monogr.*, 74, 443–489, doi: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, *Ecol. Model.*, 200, 493–510, doi:10.1016/j.ecolmodel.2006.09.001, 2007.
- 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, *J. Geophys. Res.-Atmos.*, 108, 8168, doi:10.1029/2001JD000927, 2003.
- Høgh-Jensen, H., Wollenweber, B., and Schjoerring, J. K.: Kinetics of nitrate and ammonium absorption and accompanying H<sup>+</sup> fluxes in roots of *Lolium perenne* L. and N<sub>2</sub>-fixing *Trifolium repens* L., *Plant Cell Environ.*, 20, 1184–1192, doi:10.1046/j.1365-3040.1997.d01-145.x, 1997.
- Hungate, R.: The rumen microbial ecosystem, *Annu. Rev. Ecol. Syst.*, 6, 39–66, doi:10.1146/annurev.es.06.110175.000351, 1975.
- 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, UK and New York, NY, USA, doi:10.1017/CBO9781107415324, 2013.
- Jarrell, K. F.: Extreme oxygen sensitivity in methanogenic archaeabacteria, *BioScience*, 35, 298–302, doi:10.2307/1309929, 1985.
- Kamer, K., Kennison, R. L., and Fong, P.: Rates of inorganic nitrogen uptake by the estuarine green macroalgae *Enteromorpha intestinalis* and *Ulva expansa*, 2003, 130–141, 2001.
- Kirchman, D. L.: The uptake of inorganic nutrients by heterotrophic bacteria, *Microb. Ecol.*, 28, 255–271, doi: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 Res. Pt. I*, 45, 347–365, doi:10.1016/S0967-0637(97)00075-7, 1998.
- Kuzyakov, Y. and Xu, X.: Competition between roots and microorganisms for nitrogen: mechanisms and ecological relevance, *New Phytol.*, 198, 656–669, doi: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, *Environ. Res. Lett.*, 7, 015507, doi:10.1088/1748-9326/7/1/015507, 2012.
- Lemmon, D. R. and Schafer, J. L.: Developing Statistical Software in Fortran 95, Statistics and Computing, Springer, 2005.
- 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, *J. Geophys. Res.-Biogeosci.*, 113, G02016, doi:10.1029/2007JG000578, 2008.
- Manzoni, S. and Porporato, A.: Soil carbon and nitrogen mineralization: theory and models across scales, *Soil Biol. Biochem.*, 41, 1355–1379, doi:10.1016/j.soilbio.2009.02.031, 2009.
- Middelburg, J. J. and Nieuwenhuize, J.: Nitrogen uptake by heterotrophic bacteria and phytoplankton in the nitrate-rich Thames Estuary, *Mar. Ecol.-Prog. Ser.*, 203, 13–21, doi:10.3354/meps203013, 2000.
- Nollet, L. M. L. and De Gelder, L. S. P.: Handbook of Water Analysis, 3rd Edn., 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, 2001.
- 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, 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.: Generalized model for N<sub>2</sub> and N<sub>2</sub>O production from nitrification and denitrification, *Global Biogeochem. Cy.*, 10, 401–412, doi:10.1029/96GB01455, 1996.
- Parton, W. J., Holland, E. A., Del Grosso, S. J., Hartman, M. D., Martin, R. E., Mosier, A. R., Ojima, D. S., and Schimel, D. S.: Generalized model for NO<sub>x</sub> and N<sub>2</sub>O emissions from soils, *J. Geophys. Res.-Atmos.*, 106, 17403–17419, doi:10.1029/2001JD900101, 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 Physiol.*, 29, 389–399, doi:10.1093/treephys/tpn033, 2009.
- 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 Phytol.*, 200, 350–365, doi: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 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, 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.: Persistence of soil organic matter as an ecosystem property, *Nature*, 478, 49–56, doi: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 exchanges of energy, water, and carbon between continents and the atmosphere, *Science*, 275, 502–509, doi:10.1126/science.275.5299.502, 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-Sci. Rev.*, 99, 125–161, doi:10.1016/j.earscirev.2010.02.004, 2010.
- Shampine, L. F., Thompson, S., Kierzenka, J. A., and Byrne, G. D.: Non-negative solutions of ODEs, *Appl. Math. Comput.*, 170, 556–569, doi:10.1016/j.amc.2004.12.011, 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 Transport, LBNL, Report, available at: [http://esd1.lbl.gov/files/research/projects/tough/documentation/TOUGHREACT\\_V3-OMP\\_RefManual.pdf](http://esd1.lbl.gov/files/research/projects/tough/documentation/TOUGHREACT_V3-OMP_RefManual.pdf) (last access: 29 February 2016), 2014.
- Steefel, C. I., Appelo, C. A. J., Arora, B., Jacques, D., Kalbacher, T., Kolditz, O., Lagneau, V., Lichtner, P. C., Mayer, K. U., Meeussen, J. C. L., Molins, S., Moulton, D., Shao, H., Šimůnek, J., Spycher, N., Yabusaki, S. B., Yeh, G. T.: Reactive transport codes for subsurface environmental simulation, *Comput. Geosci.*, 19, 445–478, doi:10.1007/s10596-014-9443-x, 2015.
- Tang, G., Yuan, F., Bisht, G., Hammond, G. E., Lichtner, P. C., Kumar, J., Mills, R. T., Xu, X., Andre, B., Hoffman, F. M., Painter, S. L. and Thornton, P. E.: Addressing numerical challenges in introducing a reactive transport code into a land surface model: A biogeochemical modeling proof-of-concept with CLM-PFLOTRAN 1.0: Modeling Archive. Next Generation Ecosystem Experiments Arctic Data Collection, Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA. doi:10.5440/1239799, 2016.
- Tang, J. Y. and Riley, W. J.: Technical Note: A generic law-of-the-minimum flux limiter for simulating substrate limitation in biogeochemical models, *Biogeosciences*, 13, 723–735, doi:10.5194/bg-13-723-2016, 2016.
- Tang, J. Y., Riley, W. J., Koven, C. D., and Subin, Z. M.: CLM4-BeTR, a generic biogeochemical transport and reaction module for CLM4: model development, evaluation, and application, *Geosci. Model Dev.*, 6, 127–140, doi:10.5194/gmd-6-127-2013, 2013.
- Thornton, P. E. and Rosenbloom, N. A.: Ecosystem model spin-up: estimating steady state conditions in a coupled terrestrial carbon and nitrogen cycle model, *Ecol. Model.*, 189, 25–48, doi:10.1016/j.ecolmodel.2005.04.008, 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. Coast. Shelf S.*, 61, 507–515, doi:10.1016/j.ecss.2004.06.014, 2004.
- Warren, C. R. and Adams, P. R.: Uptake of nitrate, ammonium and glycine by plants of Tasmanian wet eucalypt forests, *Tree Physiol.*, 27, 413–419, doi: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.
- Yeh, G. T. and Tripathi, V. S.: A critical evaluation of recent developments in hydrogeochemical transport models of reactive multi-chemical components, *Water Resour. Res.*, 25, 93–108, doi:10.1029/WR025i001p00093, 1989.
- 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.
- Zheng, C. and Bennett, G. D.: Applied Contaminant Transport Modeling, 2nd Edn., Wiley-Interscience, 2002.