Accelerating the spin-up of the coupled carbon and nitrogen cycle model in CLM4

Yilin Fang[1*], Chongxuan Liu[2], L. Ruby Leung[3]

[1]Hydrology Group, Energy and Environment Directorate, Pacific Northwest National Laboratory, Richland, WA 99352, USA

[2]Geochemistry Group, Fundamental and Computational Sciences Directorate, Pacific Northwest National Laboratory, Richland, WA 99352, USA

[3]Atmospheric Sciences and Global Change Division, Fundamental and Computational Sciences Directorate, Pacific Northwest National Laboratory, Richland, WA 99352, USA

[*]Corresponding Author:

Mailing address:

Pacific Northwest National Laboratory

902 Battelle Boulevard, Richland, Washington 99352, USA

Telephone: 1-509-372-4051

Fax: 1-509-372-6089

e-mail: yilin.fang@pnnl.gov

Abstract

The commonly adopted biogeochemistry spin-up process in earth system model is to run the model for hundreds to thousands of years subject to periodic atmospheric forcing to reach dynamic steady state of the carbon-nitrogen (CN) models. A variety of approaches have been proposed to reduce the computation time of the spin-up process. Significant improvement in computational efficiency has been made recently. However, a long simulation time is still required to reach the common convergence criteria of the coupled carbon/nitrogen model. A gradient projection method was proposed and used to further reduce the computation time after examining the trend of the dominant carbon pools. The Community Land Model version 4 (CLM4) with carbon and nitrogen component was used in this study. From point scale simulations we found that the method can reduce the computation time by 20-69% compared to one of the fastest approaches in the literature. We also found that the cyclic stability of total carbon for some cases differs from that of the periodic atmospheric forcing, and some cases even showed instability. Close examination showed that one case has a carbon periodicity much longer than that of the atmospheric forcing due to the annual fire disturbance that is longer than half a year. The rest was caused by the instability of water table calculation in the hydrology model of CLM4. The instability issue is resolved after we replaced the hydrology scheme in CLM4 with a flow model for variably saturated porous media.

1. Introduction

The Initial starting values of carbon/nitrogen (CN) models are not commonly available, especially for large scale applications, but they have important influence on the subsequent C/N states simulated by the models. Typically, Earth system model (ESM) simulations are initialized

in the pre-industrial period to allow sufficient time for the coupled system to respond to the various forcing. Initialization of the CN model is usually achieved by a spin-up run of the CN model given an arbitrary initial condition until an approximate C equilibrium is reached. This time marching of the model requires several hundreds to thousands of years of model simulations before a dynamic steady state is reached. The length of the transient state to reach a dynamic steady state is dependent on the initial conditions of the system. It has long been recognized that the spin-up process of CN models is time consuming due to the slow turnover rates of the soil carbon pools, which significantly affect the computational efficiency for global modeling. A number of approaches have been proposed in the past to improve upon the explicit forward time integration of ordinary differential equations in their native form and rate parameters for CN models. These approaches include the initialization of soil organic matter carbon pools with observations [*Zhang et al.*, 2002], accelerated decomposition method using a higher decomposition rate for litter and soil carbon pools [*Thornton and Rosenbloom*, 2005], decelerated bulk denitrification and leaching method [*Shi et al.*, 2013], and semi-analytical steady-state solution for soil organic C and N pools [*Xia et al.*, 2012]. Except for the semi-analytical approach, other approaches mentioned above have been summarized and compared in Shi et al. [2013]. The semi-analytical model needs initial spin-up values of net primary productivity (NPP), which still requires long simulation time for stabilization because C and N are coupled in CN models. We had previously restructured the CN model in Community Land Model version 4 (CLM4-CN) [*Lawrence et al*. 2011] and developed a steady-state solution directly using annually averaged rate parameters [*Fang et al.*, 2013; *Fang et al*. 2014]. Using our approach, we were able to implement the semi-analytical method in *Xia et al*. [2012]. Our

numerical experiment showed that the semi-analytical method is not necessarily faster compared to the modified form of the "accelerated decomposition" approach in Koven et al. [2013].

Recently Koven et al. [2013] used a modified form of the "accelerated decomposition" (hereafter referred as AD approach) by numerically increasing the decomposition rates for the two slowest soil carbon pools (named as Soil3C and Soil4C) to a level so that their turnover rates are similar to the fast pools during the initialization. Numerical evaluation found that the approach significantly reduced the spin-up time [*Koven et al.*, 2013]. Figure 1 shows the structure of the soil C pool represented in CLM4-CN. Note that heterotrophic respiration fractions are not shown. The reason that the AD approach can accelerate the spin-up is because these two slowest pools are essentially decoupled from the rest of the ordinary differential equations, in that all other pools do not need input from them. The approach, however, cannot be applied to the coarse woody debris (CWD) pool even though its turnover rate is on the same order of Soil3C, because it is an input to the litter pools. Changing the rate of CWD will give a different solution of other pools during each integration step using the same initial condition, which will lead to a state far from equilibrium if the state is used in a restart simulation.

In the AD approach, once the solution is obtained from the accelerated run, the state of Soil3C and Soil4C can be analytically solved. From Fig. 1, the flux of Soil3C and Soil4C pools can be described by the following equations:

$$\frac{dC_{Soil3}}{dt} = -k_{S3}C_{Soil3} + k_{S2}C_{Soil2} + k_{L3}C_{Litr3} \quad (1)$$

$$\frac{dC_{Soil4}}{dt} = -k_{S4}C_{Soil4} + k_{S3}C_{Soil3} \quad (2)$$

where k_{L3}, k_{S2}, k_{S3}, and k_{S4} are the turnover rates of Litter 3, Soil 2, Soil 3, and Soil 4 C pool shown in Fig. 1, respectively. C_{Litr3}, C_{Soil2}, C_{Soil3} and C_{Soil4} are the amount of C in Litter 3, Soil 2, Soil 3 and Soil 4 C pools, respectively. The first term on the right hand side of Eqs. (1) and (2)

84 includes heterotrophic respiration. At the steady state, the left hand side of Eqs. (1) and (2)
85 becomes 0, the amount of Soil3C and Soil4C can then be solved:

86 $$C_{Soil3} = \frac{k_{S2}}{k_{S3}} C_{Soil2} + \frac{k_{L3}}{k_{S3}} C_{Litr3} \quad (3)$$

87 $$C_{Soil4} = \frac{k_{S3}}{k_{S4}} C_{Soil3} \quad (4)$$

88 Eqs. 3 and 4 are applicable regardless whether AD or native run was used (the native run was
89 defined here as the simulations without changing the decomposition rates of Soil3 and Soil4 C
90 pools). Therefore, multiplying Eqs. (3) and (4) by their corresponding accelerator, the results
91 should be close to the native runs. That is

92 $$C_{Soil3,N} = \frac{k_{S2}}{k_{S3,N}} C_{Soil2} + \frac{k_{L3}}{k_{S3,N}} C_{Litr3} = A_{S3} C_{soil3} \quad (5)$$

93 $$C_{Soil4,N} = \frac{k_{S3}}{k_{S4,N}} C_{Soil3} = A_{S4} C_{Soil4} \quad (6)$$

94 where "$_{,N}$" denotes the native run, A is the accelerator.

95 Even with this modified accelerator approach, long simulation time cannot be avoided at dry
96 and cold places because decomposition scaling factor is associated with soil water potential and
97 temperature. Hence new methods are needed to address the spin-up problem. In implicit time
98 integration approaches, based on knowledge about the trajectory of the solution of the initial
99 value problem, linear extrapolation from time integration was often used to find a good initial
100 value for iterative multirate multidisciplinary processes [*Birken et al*., 2014 and references
101 therein]. A number of explicit Euler steps with small time steps followed by an explicit Euler
102 step with large time step when the change in components due to fast processes become negligible
103 has been shown to efficiently solve stiff ordinary differential equations [*Eriksson et al.*, 2003;
104 *Gear and Kevrekidis*, 2003]. We made use of these concepts, referred to as the Gradient

Projection (GP) approach in this study, to further improve the computational efficiency of the biogeochemistry spin-up processes.

2. Methods

2.1 *Model Description*

Community land model, CLM4, is the land component of the Community Earth System Model (CESM) [*Lawrence et al.*, 2011]. Processes simulated in CLM4 include biogeophysics (solar and longwave radiation, momentum, heat transfer in soil and snow, hydrology of canopy, soil, and snow, and stomatal physiology and photosynthesis) and biogeochemistry (phenology, autotrophic respiration, heterotrophic respiration, carbon and nitrogen allocation, and nitrogen source/sink). The vegetation structures (leaf area index, stem area index and height) in CLM4-CN are represented through the predictive state variables of leaf and stem carbon, which are coupled to simulate fluxes of carbon and nitrogen state variables in vegetation, litter, and soil organic matter [*Lawrence et al.*, 2011; *Thornton and Zimmermann*, 2007]. The tree, shrub and grass plant functional types (PFTs) are divided into tropical, temperate and boreal climate groupings using the PFT physiology and climate rules of Nemani and Running [1996] and C3/C4 photosynthetic pathways in the case of grasses [*Lawrence and Chase*, 2007]. For this study, we used CLM4-CN in offline mode, which is not coupled to an atmosphere model.

2.2 *Gradient Projection Method*

If m_c is the number of years (one cycle) of atmospheric forcing which will be used repeatedly in the spin-up run, we use a spin-up time of [$(n+1)\ m_c$] years as a stop point for the accelerated decomposition (AD) run, where $n = 300/m_c$ is an integer. For example, if the number of years of

127 forcing is m_c = 7, the stop time will be at year 301. A stop point of ~300 years for the modified
128 AD approach was selected based on the model results in *Koven et al.* [2013], but it is not an
129 absolute requirement. The best approach is to stop when NPP reaches a dynamic steady-state.

130 At the end of the accelerated run, a dynamic steady state water table should be reached in the
131 soil column. Due to the slow turnover rates, the total soil carbon gradually approaches steady
132 state from one cycle to the next (Fig. 2(a)). We can approximate C at a future time t_n (Fig. 2(b))
133 using the C gradient between two consecutive cycles expressed in the following equation:

134 $$C(t_n) = C(t_1) + \frac{C(t_1)-C(t_0)}{m_c}(t_n - t_1) \quad (7)$$

135 where t_0 is the beginning of the first cycle, t_1 is the beginning of the next cycle, and $t_1 - t_0 = m_c$;
136 $t_n - t_1 = \tau\, m_c$, τ is an integer close to the turnover years (reciprocal of turnover rate) of Soil4 C
137 pool to satisfy the stability requirement of forward or explicit time integration that is used in
138 CLM4-CN to solve the time-dependent ordinary differential equations. The explicit method can
139 be numerically unstable (convergence of solution is not guaranteed) if the time step is too big
140 [*LeVeque*, 2007]. For the first order kinetic type problem, i.e., $u'(t) = ku(t)$, the stability
141 requirement is $|1 + kh| \leq 1$, in which k is the rate constant and h is the time step.

142 We call the method shown in Eq. (7) the gradient projection (GP) method. This method
143 is analogous to that described in Eriksson et al. [2003], which uses a large time step that satisfies
144 stability requirement to integrate the slowest processes once the contributions from fast processes
145 become negligible. We allow j_p to be chosen based on the time period needed to stabilize the
146 components from fast processes between cycles after perturbation, or set as an integer equals to
147 $m_c \times (100/m_c + 1)$) years of simulation after restart from the accelerated run before using this
148 approach, and also perform j_p years of simulation followed by each projection until the solution

149 meets the common convergence criteria of 0.5 g m^{-2} for total soil C during two consecutive
150 cycles [*Shi et al.*, 2013; *Thornton and Rosenbloom*, 2005]. During each projection, balance
151 check for C and N is turned off. The GP method is only applied to the dominant C and N pools,
152 i.e., coarse wood debris, dead stem, dead coarse root and Soil4 pool.

153 **3. Results**

154 A total of 38 single point tower sites from the FLUXNET [*Baldocchi et al.*, 2001] were
155 selected to assess the gradient projection method. These sites include temperate, boreal, tropical,
156 and subtropical climatic environments and four ecosystem types (tropical forests, temperate
157 forests, boreal forest, grasslands, and Mediterranean-type ecosystems) (Table 1).

158 The meteorological forcing, site information such as soil texture, vegetation cover, and
159 satellite-derived phenology at each site are provided by the North American Carbon Program
160 (NACP) site synthesis team for the sites located in North America and by the Large Scale
161 Biosphere-Atmosphere Experiment in Amazônia Model Intercomparison Project (LBA-MIP) for
162 the sites located in South America. The NACP site synthesis and LBA-MIP datasets are detailed
163 in Schwalm et al. [2010] and at http://www.climatemodeling.org/lba-mip. Each site has two
164 runs, one using the AD method and the other using the GP method. The available forcing (Table
165 1) is applied repeatedly during the simulation for each site.

166 Table 2 shows the comparison of total simulation years till a certain convergence criterion is
167 met. Three convergence threshold values in ΔC_{TOC} (3.0, 1.0, and 0.5 g m^{-2} yr^{-1}) were compared.
168 The quality of total soil C is better when the threshold value is smaller [*Thornton and*
169 *Rosenbloom*, 2005]. Compared to the modified AD approach, the reduction of computation cost
170 is shown in Fig. 3. Fig. 3 shows that when high quality solution ($\Delta C_{TOC} \leq 0.5$ g m^{-2} yr^{-1}) is

171 required, the average total reduction in computation cost is 40%. Average 23% of computation
172 time is reduced for achieving the low quality solution ($\Delta C_{TOC} \leq 3$ g m^{-2} yr^{-1}).

173 Note that the computation cost reduction for sites US-Me2, RJA, US-IB1 and US-SO2 is not
174 shown in Fig. 3. Site US-Me2 met the convergence criteria before the GP method is applied.
175 Sites RJA and US-IB1 show oscillation of the annual average total C from one full length
176 (multiple years) of forcing cycle to the next, and site US-SO2 shows a carbon periodicity much
177 longer (81 years) than that of the atmospheric forcing (9 years) (Fig. 4). The oscillation noted in
178 the simulations at RJA and US-IB1 differs from the variability within the forcing cycle, which
179 happens when soil C has a fast turnover rate such that soil C dynamics are primarily controlled
180 by variability of the forcing [*Luo et al*. 2014]. Due to the aforementioned reasons, the GP
181 method failed at those three sites.

182 We first checked whether the oscillation and longer periodicity were caused by fire
183 disturbance. However this can only explain the oscillation at site US-SO2. The annual fire
184 disturbance at site US-SO2 is longer than half a year, while it is less than a month at the other
185 two sites. In the original CLM4, soil water is calculated first for the top ten soil layers (3.8 m
186 below the ground surface) and one aquifer layer using water content based formulation for water
187 mass conservation and groundwater table as bottom boundary condition [*Oleson et al.*, 2010];
188 the Niu et al. [*Niu et al.*, 2005; *Niu et al.*, 2007] parameterizations are then used to simulate
189 groundwater-soil water interaction and update the water table depth. If the water table is below
190 3.8 m, groundwater does not contribute to the soil moisture in the overlaying soil layers. We
191 found that after a hundred years, the water table calculation scheme in CLM4 has resulted in
192 significantly different evolution of water table depth from one cycle to the next when repeatedly
193 forcing the model with atmospheric data at sites RJA and US-IB1. The issue has also been found

194 previously and effort has been made to eliminate the oscillations [*Oleson et al.*, 2010], but such
195 oscillations can still occur under specific conditions such as at RJA and US-IB1. When we
196 turned off the groundwater component, i.e., applying zero flux boundary condition at the bottom
197 of the soil column, we didn't see oscillations in SOC at RJA and US-IB1. In Niu et al. [2007]
198 groundwater model, after solving the mass conservation equations (Richards' equation) in the
199 top ten layers, water is then moved around to account for recharge and subsurface runoff and in
200 the meantime satisfy two conditions for water content in each layer, i.e., the water content has to
201 be greater than the minimum content and smaller than the effective porosity of the layer. By
202 moving water mass around after the Richards' equation is solved, Richards' equation at each
203 node is no longer satisfied if its moisture deviates from its previous solution. We have
204 confirmed the local mass conservation error of water in the original model of CLM4. The error
205 is large when recharge or subsurface runoff is high. The water content formulation itself has
206 been previously shown to cause solution instability for soils near saturation [*Hills et al.*, 1989].
207 Instead of solving the soil water and groundwater separately, we use a flow model for variably
208 saturated porous media, STOMP (Subsurface Transport Over Multiple Phases) [*White and*
209 *Oostrom*, 2000], to see if it can resolve the oscillation in the total soil C.

210 The STOMP simulator was developed to predict non-isothermal hydrological flow and
211 reactive transport in variably saturated subsurface environments. In STOMP, the water mass
212 conservation equation balances the time rate of change of water mass within a control volume
213 with the flux of water mass crossing the control volume surface. The nonlinear equations
214 describing mass conservation are discretized spatially on structured orthogonal grids using the
215 integral finite difference approach of Patankar [1980], which is locally and globally mass
216 conserving. The equations are discretized temporally using first-order backward Euler

differencing or implicit time stepping that is suitable for the solution of the equations that are numerically unstable [*LeVeque,* 2007]. Newton–Raphson iteration is used to resolve the nonlinearities from the constitutive equations that relate the primary and secondary variables.

Detailed information regarding STOMP, such as user's guide, theory guide and code availability can be found at http://stomp.pnnl.gov. For each soil column, the number of vertical grids used for STOMP is 15 and it is the same as that in CLM4. In CLM4, the top ten grids (3.8 m below ground) are used in the soil water scheme. The same initial saturation condition as that in CLM4 is prescribed. For the grid at the top, the Neumann boundary condition is used. For the bottom (42 m below ground), a zero flux boundary condition is used. Because the aquifer is unconfined, we use the bottom node pressure to calculate water table depth. Figs. 5 and 6 shows the model comparison at the beginning of the first 3 years between the simulations using the original soil hydrology scheme in CLM4 and the simulation after replacing the soil water and groundwater-soil water interaction scheme with STOMP at sites RJA and US-IB1. Using STOMP, mass conservation is improved, and the moisture content calculated is more accurate, resulting in wetter and cooler soil (Figs. 5 (b), (c) and Figs. 6 (b), (c)).

Fig. 7 (a) and (b) shows the oscillations of water table depth resolved at both RJA and US-IB1, i.e., the oscillations between forcing cycles noted in the original hydrology scheme in CLM4 are caused by the local water mass balance error. Each cycle of atmospheric forcing at both sites has 3 years. The GP method is successful at those two sites. The little jumps in Fig. 7 (c) and (d) are where the GP method is applied. Both sites show higher total soil C predicted compared to Figs. 4 (a) and (b) because of the new flow model. The issue of the original groundwater model in CLM4 might explain why it took so long (> 4000 years) for some of the grids in the global simulation to converge as shown in Shi et al. [2013]. In addition, the gradient

projection model is not recommended for sites where the length of fire season is too long. For those sites, the overall time that takes for the spin-up run to steady-state is much shorter compared to others, therefore no improvement on the spin-up time is necessary.

4. Conclusions

We described a gradient projection method to further speedup the spin-up process based on the slow nature of soil organic C decomposition. Comparison between our approach and the modified accelerator approach showed that 20-69% of simulation years can be reduced with our approach. While the approach was specifically evaluated using CLM4CN, it can also be readily applied to other CN models in earth system models. No matter what modification is made to improve the spin-up efficiency, a final spin-up is always needed to reach a converged solution due to disequilibrium caused by the modification. Our approach is especially useful when new model formulation is proposed and high quality solution (small convergence threshold) is needed for a fair comparison.

In addition, we also found that the original numerical hydrology scheme, especially the water table calculation in CLM4 creates numerical oscillations in simulated water table, leading to a challenge to achieve the common convergence criteria for soil C. To resolve the issue, we replaced the hydrological model using a flow model for variably saturated porous media. The new flow model caused about 10% increase in computation time, but gives more accurate results that corrected the oscillation behavior of the original hydrological model. Comparing the total C predicted by the old and new flow models, we also see more C being predicted using the new flow model. Whether the prediction of more C is realistic depends on other factors besides the hydrology so we have not attempted to evaluate the simulated C using observations.

Nevertheless, a correct implementation of numerical schemes is always desirable for reducing uncertainty in model prediction.

5. Code availability

The source code of CLM4.0 and STOMP can be requested through http://www.cesm.ucar.edu/models/cesm1.0/ and http://stomp.pnnl.gov/licensing.stm, respectively. The method implemented in this study can be obtained upon request. Contact: yilin.fang@pnnl.gov.

ACKNOWLEDGMENT

This research has been accomplished through funding support from Pacific Northwest National Laboratory's Laboratory Directed Research and Development Program. We thank the North American Carbon Program Site-Level Interim Synthesis team, the Large Scale Biosphere-Atmosphere Experiment in Amazônia Model Intercomparison Project team, and the site investigators for collecting, organizing, and distributing the data required for this analysis. We thank the anonymous referee and Yiqi Luo for comments that improved the manuscript. A portion of this research was performed using PNNL Institutional Computing at Pacific Northwest National Laboratory. PNNL is operated by Battelle for the U.S. Department of Energy under Contract DE-AC05-76RL01830.

References:

Arain, A. A., and N. Restrepo-Coupe (2005), Net ecosystem production in a temperate pine plantation in southeastern Canada, *Agricultural and Forest Meteorology*, *128*(3-4), 223-241, doi: 10.1016/j.agrformet.2004.10.003.

Baldocchi, D., Falge, E., Gu, L. H., Olson, R., Hollinger, D., Running, S., Anthoni, P., Bernhofer, C., Davis, K., Evans, R., Fuentes, J., Goldstein, A., Katul, G., Law, B., Lee, X. H., Malhi, Y., Meyers, W., Oechel, W., U, K.T.P., Pilegaards, K., Schmid, H.Pl, Valentini, R., Verma, S., Vesala, T., Wilson, K. and Wofsy, S. (2001), FLUXNET: A new tool to study the temporal and spatial variability of ecosystem-scale carbon dioxide, water vapor, and energy flux densities, *Bulletin of the American Meteorological Society*, *82*(11), 2415-2434, doi: 10.1175/1520-0477(2001)082<2415:Fantts>2.3.Co;2.

Baldocchi, D. D., S. Y. Ma, S. Rambal, L. Misson, J. M. Ourcival, J. M. Limousin, J. Pereira, and D. Papale (2010), On the differential advantages of evergreenness and deciduousness in mediterranean oak woodlands: a flux perspective, *Ecol Appl*, *20*(6), 1583-1597, doi: 10.1890/08-2047.1.

Baldocchi, D. D., L. K. Xu, and N. Kiang (2004), How plant functional-type, weather, seasonal drought, and soil physical properties alter water and energy fluxes of an oak-grass savanna and an annual grassland, *Agricultural and Forest Meteorology*, *123*(1-2), 13-39, doi: 10.1016/j.agrformet.2003.11.006.

Birken P., Gleim T., Kuhl D., and Meister A. (2014), Fast Solvers for Unsteady Thermal Fluid Structure Interaction, arXiv:1407.0893v1.

Chen, J. M., A. Govind, O. Sonnentag, Y. Q. Zhang, A. Barr, and B. Amiro (2006), Leaf area index measurements at Fluxnet-Canada forest sites, *Agricultural and Forest Meteorology*, *140*(1-4), 257-268, doi: 10.1016/j.agrformet.2006.08.005.

Curtis, P. S., C. S. Vogel, C. M. Gough, H. P. Schmid, H. B. Su, and B. D. Bovard (2005), Respiratory carbon losses and the carbon-use efficiency of a northern hardwood forest, 1999-2003, *New Phytol*, *167*(2), 437-455, doi: 10.1111/j.1469-8137.2005.01438.x.

Davis, K. J., P. S. Bakwin, C. X. Yi, B. W. Berger, C. L. Zhao, R. M. Teclaw, and J. G. Isebrands (2003), The annual cycles of CO2 and H2O exchange over a northern mixed forest as observed from a very tall tower, *Global Change Biology*, *9*(9), 1278-1293, doi: 10.1046/j.1365-2486.2003.00672.x.

Desai, A. R., P. V. Bolstad, B. D. Cook, K. J. Davis, and E. V. Carey (2005), Comparing net ecosystem exchange of carbon dioxide between an old-growth and mature forest in the upper Midwest, USA, *Agricultural and Forest Meteorology*, *128*(1-2), 33-55, doi: 10.1016/j.agrformet.2004.09.005.

Fang, Y., M. Huang, C. Liu, H. Li, and R. Leung (2013), A generic biogeochemical module for Earth system models: Next Generation BioGeoChemical Module (NGBGC), version 1.0, Geoscientific Model Development(6), 1977-1988, doi:10.5194/gmd-6-1977-2013.

Fang, Y., C. Liu, M. Huang, H. Li, and R. Leung (2014), Steady state estimation of soil organic carbon using satellite-derived canopy leaf area index, Journal of Advances in Modeling Earth Systems (6), 1049-1064, DOI: 10.1002/2014MS000331.

Eriksson, K., C. Johnson, and A. Logg (2003), Explicit time-stepping for stiff ODES, Siam J Sci Comput, 25(4), 1142-1157, doi:Doi 10.1137/S1064827502409626.

Fischer, M. L., D. P. Billesbach, J. A. Berry, W. J. Riley, and M. S. Torn (2007), Spatiotemporal variations in growing season exchanges of CO2, H2O, and sensible heat in agricultural fields of the Southern Great Plains, *Earth Interact*, *11*, doi: 10.1175/EI231.1.

Flanagan, L. B., L. A. Wever, and P. J. Carlson (2002), Seasonal and interannual variation in carbon dioxide exchange and carbon balance in a northern temperate grassland, *Global Change Biology*, *8*(7), 599-615, doi: 10.1046/j.1365-2486.2002.00491.x.

Gear, C. W., and I. G. Kevrekidis (2003), Projective methods for stiff differential equations: Problems with gaps in their eigenvalue spectrum, Siam J Sci Comput, 24(4), 1091-1106, doi:Pii S1064827501388157

Goulden, M. L., J. W. Munger, S. M. Fan, B. C. Daube, and S. C. Wofsy (1996), Measurements of carbon sequestration by long-term eddy covariance: Methods and a critical evaluation of accuracy, *Global Change Biology*, *2*(3), 169-182, doi: 10.1111/j.1365-2486.1996.tb00070.x.

Griffis, T. J., T. A. Black, K. Morgenstern, A. G. Barr, Z. Nesic, G. B. Drewitt, D. Gaumont-Guay, and J. H. McCaughey (2003), Ecophysiological controls on the carbon balances of three southern boreal forests, *Agricultural and Forest Meteorology*, *117*(1-2), 53-71, doi: 10.1016/S0168-1923(03)00023-6.

Gu, L. H., Meyers, T., Pallardy, S. G., Hanson, P. J., Yang, B., Heuer, M., Hosman, K. P., Liu, Q., Riggs, J. S., Sluss, D. and Wullschleger, S. D. (2007), Influences of biomass heat and biochemical energy storages on the land surface fluxes and radiative temperature,D02107 *Journal of Geophysical Research-Atmospheres*, *112*(D6), doi: 10.1029/2007jd008509.

Hills, R. G., I. Porro, D. B. Hudson, and P. J. Wierenga (1989), Modeling One-Dimensional Infiltration into Very Dry Soils .1. Model Development and Evaluation, *Water Resour Res*, *25*(6), 1259-1269, doi: 10.1029/Wr025i006p01259.

Hollinger, D. Y., S. M. Goltz, E. A. Davidson, J. T. Lee, K. Tu, and H. T. Valentine (1999), Seasonal patterns and environmental control of carbon dioxide and water vapour exchange in an ecotonal boreal forest, *Global Change Biology*, *5*(8), 891-902, doi: 10.1046/j.1365-2486.1999.00281.x.

Hudiburg, T. W., B. E. Law, and P. E. Thornton (2013), Evaluation and improvement of the Community Land Model (CLM4) in Oregon forests, *Biogeosciences*, *10*(1), 453-470, doi: 10.5194/bg-10-453-2013.

Humphreys, E. R., T. A. Black, K. Morgenstern, T. B. Cai, G. B. Drewitt, Z. Nesic, and J. A. Trofymow (2006), Carbon dioxide fluxes in coastal Douglas-fir stands at different stages of

development after clearcut harvesting, *Agricultural and Forest Meteorology*, *140*(1-4), 6-22, doi: 10.1016/j.agrformet.2006.03.018.

Katul, G., R. Leuning, and R. Oren (2003), Relationship between plant hydraulic and biochemical properties derived from a steady-state coupled water and carbon transport model, *Plant Cell Environ*, *26*(3), 339-350, doi: 10.1046/j.1365-3040.2003.00965.x.

Koven, C. D., W. J. Riley, Z. M. Subin, J. Y. Tang, M. S. Torn, W. D. Collins, G. B. Bonan, D. M. Lawrence, and S. C. Swenson (2013), The effect of vertically resolved soil biogeochemistry and alternate soil C and N models on C dynamics of CLM4, *Biogeosciences*, *10*(11), 7109-7131, doi: 10.5194/bg-10-7109-2013.

Lawrence, D. M., Oleson, K. W., Flanner, M. G., Thornton, P. E., Swenson, S. C., Lawrence, P. J., Zeng, X. B., Yang, Z. L., Levis, S., Sakaguchi, K., Bonan, G. B., and Slater, A. G. , Parameterization Improvements and Functional and Structural Advances in Version 4 of the Community Land Model, *Journal of Advances in Modeling Earth Systems*, *3*, M03001, doi: 10.1029/2011ms000045.

Lawrence, P. J., and T. N. Chase (2007), Representing a new MODIS consistent land surface in the Community Land Model (CLM 3.0), *Journal of Geophysical Research-Biogeosciences*, *112*, *G01023*, doi: 10.1029/2006jg000168.

LeVeque, R. J. (2007), Finite Difference Methods for Ordinary and Partial Differential Equations: Steady-State and Time-Dependent Problems Society for Industrial and Applied Mathematics, Philadelphia, PA.

Lipson, D. A., R. F. Wilson, and W. C. Oechel (2005), Effects of elevated atmospheric CO2 on soil microbial biomass, activity, and diversity in a chaparral ecosystem, *Appl Environ Microb*, *71*(12), 8573-8580, doi: 10.1128/Aem.71.12.8573-8580.2005.

Luo, Y., T. F. Keenan, and M. Smith (2014), Predictability of the terrestrial carbon cycle, Global Change Biology (2014), doi: 10.1111/gcb.12766.

Matamala, R., J. D. Jastrow, R. M. Miller, and C. T. Garten (2008), Temporal changes in C and N stocks of restored prairie: Implications for C sequestration strategies, *Ecol Appl*, *18*(6), 1470-1488, doi: 10.1890/07-1609.1.

McCaughey, J. H., M. R. Pejam, M. A. Arain, and D. A. Cameron (2006), Carbon dioxide and energy fluxes from a boreal mixedwood forest ecosystem in Ontario, Canada, *Agricultural and Forest Meteorology*, *140*(1-4), 79-96, doi: 10.1016/j.agrformet.2006.08.010.

Nemani, R., and S. W. Running (1996), Implementation of a hierarchical global vegetation classification in ecosystem function models, *Journal of Vegetation Science*, *7*(3), 337-346, doi: 10.2307/3236277.

Niu, G. Y., Z. L. Yang, R. E. Dickinson, and L. E. Gulden (2005), A simple TOPMODEL-based runoff parameterization (SIMTOP) for use in global climate models, *Journal of Geophysical Research-Atmospheres*, *110, D21106*, doi: 10.1029/2005jd006111.

Niu, G. Y., Z. L. Yang, R. E. Dickinson, L. E. Gulden, and H. Su (2007), Development of a simple groundwater model for use in climate models and evaluation with Gravity Recovery and Climate Experiment data, *Journal of Geophysical Research-Atmospheres*, *112*, *D07103*, doi: 10.1029/2006jd007522.

Oleson, K. W., Lawrence, D. M., Bonan, G. B., Flanner, M. G., Kluzek, E., Lawrence, P. J., Levis, S., Swenson, S. C., Thornton, P. E., Dai, A., Decker, M., Dickinson, R., Feddema, J., Heald, C. L., Hoffman, F., Lamarque, J-F., Mahowald, N., Niu, G-Y., Qian, T., Randerson, J., Running, S., Sakaguchi, K., Slater, A., Stöckli, R., Wang, A., Yang, Z-L., Zeng, X., and Zeng, X. (2010), Technical Description of version 4.0 of the Community Land Model (CLM). report, 266 pp., Natl. Cent. for Atmos. Res., Boulder, Colo.,*Rep.*

Patankar, S. V. (1980), Numerical Heat Transfer and Fluid Flow, Hemisphere Publishing Corporation, Washington, D.C.

Schwalm, C. R., Williams, C. A., Schaefer, K., Anderson, R., Arain, M. A., Baker, I., Barr, A., Black, T. A., Chen, G. S., Chen, J. M., Ciais, P., Davis, K. J., Desai, A., Dietze, M., Dragoni, D., Fischer, M. L., Flanagan, L. B., Grant, R., Gu, L. H., Hollinger, D., Izaurralde, R. C., Kucharik, C., Lafleur, P., Law, B. E., Li, L. H., Li, Z. P., Liu, S. G., Lokupitiya, E., Luo, Y. Q., Ma, S. Y., Margolis, H., Matamala, R., McCaughey, H., Monson, R. K., Oechel, W. C., Peng, C. H., Poulter, B., Price, D. T., Riciutto, D. M., Riley, W., Sahoo, A. K., Sprintsin, M., Sun, J. F., Tian, H. Q., Tonitto, C., Verbeeck, H., and Verma, S. B. (2010), A model-data intercomparison of CO2 exchange across North America: Results from the North American Carbon Program site synthesis, *Journal of Geophysical Research-Biogeosciences*, *115*, G00h05, doi: 10.1029/2009jg001229.

Shi, M. J., Z. L. Yang, D. M. Lawrence, R. E. Dickinson, and Z. M. Subin (2013), Spin-up processes in the Community Land Model version 4 with explicit carbon and nitrogen components, *Ecological Modelling*, *263*, 308-325, doi: 10.1016/j.ecolmodel.2013.04.008.

Sims, D. A., Rahman, A. F., Cordova, V. D., Baldocchi, D. D., Flanagan, L. B., Goldstein, A. H., Hollinger, D. Y., Misson, L., Monson, R. K., Schmid, H. P., Wofsy, S. C., and Xu, L. K. (2005), Midday values of gross CO2 flux and light use efficiency during satellite overpasses can be used to directly estimate eight-day mean flux, *Agricultural and Forest Meteorology*, *131*(1-2), 1-12, doi: 10.1016/j.agrformet.2005.04.006.

Suyker, A. E., and S. B. Verma (2012), Gross primary production and ecosystem respiration of irrigated and rainfed maize-soybean cropping systems over 8 years, *Agricultural and Forest Meteorology*, *165*, 12-24, doi: 10.1016/j.agrformet.2012.05.021.

Suyker, A. E., S. B. Verma, and G. G. Burba (2003), Interannual variability in net CO2 exchange of a native tallgrass prairie, *Global Change Biology*, *9*(2), 255-265, doi: 10.1046/j.1365-2486.2003.00567.x.

Thornton, P. E., and N. A. Rosenbloom (2005), Ecosystem model spin-up: Estimating steady state conditions in a coupled terrestrial carbon and nitrogen cycle model, *Ecological Modelling*, *189*(1-2), 25-48, doi: 10.1016/j.ecolmodel.2005.04.008.

Thornton, P. E., and N. E. Zimmermann (2007), An improved canopy integration scheme for a land surface model with prognostic canopy structure, *Journal of Climate*, *20*(15), 3902-3923, doi: 10.1175/Jcli4222.1.

Turnipseed, A. A., P. D. Blanken, D. E. Anderson, and R. K. Monson (2002), Energy budget above a high-elevation subalpine forest in complex topography, *Agricultural and Forest Meteorology*, *110*(3), 177-201, doi: 10.1016/S0168-1923(01)00290-8.

White, M. D., and M. Oostrom (2000), STOMP Subsurface Transport Over Multiple Phases, Version 2.0, Theory Guide, PNNL-12030, UC-2010, Pacific Northwest National Laboratory, Richland, WA.

Xia, J. Y., Y. Q. Luo, Y. P. Wang, E. S. Weng, and O. Hararuk (2012), A semi-analytical solution to accelerate spin-up of a coupled carbon and nitrogen land model to steady state, *Geoscientific Model Development*, *5*(5), 1259-1271, doi: 10.5194/gmd-5-1259-2012.

Yi, C., K. J. Davis, P. S. Bakwin, A. S. Denning, N. Zhang, A. Desai, J. C. Lin, and C. Gerbig (2004), Observed covariance between ecosystem carbon exchange and atmospheric boundary layer dynamics at a site in northern Wisconsin, *Journal of Geophysical Research-Atmospheres*, *109, D08302*, doi: 10.1029/2003jd004164.

Zhang, Y., C. S. Li, X. J. Zhou, and B. Moore (2002), A simulation model linking crop growth and soil biogeochemistry for sustainable agriculture, *Ecological Modelling*, *151*(1), 75-108, doi: 10.1016/S0304-3800(01)00527-0.

Table 1. Location, PFT, soil type and number of years of atmospheric forcing for each site

ID	Site	Longitude (^{o}E)	Latitude (^{o}N)	Elev (m)	Height[a] (m)	CLM4 PFT[b]	CLAY (%)	SAND (%)	SILT (%)	m_c[d] Years
1	US-Ha1 [*Goulden et al.*, 1996]	-72.1715	42.5378	343	30	BDTmp	6	66	29	1991-2006
2	US-WCr [*Yi et al.*, 2004]	-90.0799	45.8059	520	30	BDTmp	20.17	42.52	37.32	1998-2006
3	US-Syv [*Desai et al.*, 2005]	-89.3477	46.242	450	37	BDTmp	16.43	46.56	37.01	2001-2006
4	US-PFa [*Davis et al.*, 2003]	-90.2723	45.9459	470	122	BDTmp	20.17	42.52	37.32	1995-2005
5	US-UMB [*Curtis et al.*, 2005]	-84.7138	45.5598	234	50	BDTmp	0.6	92.6	6.8	1998-2006
6	US-MOz [*Gu et al.*, 2007]	-92.2	38.7441	219	30	BDTmp	24.68	46.38	28.94	2004-2007
7	US-Dk2 [*Katul et al.*, 2003]	-79.1004	35.9736	163	42	BDTmp	21.62	54.43	23.95	2003-2005
8	US-MMS [*Sims et al.*, 2005]	-86.4131	39.3231	275	48	BDTmp	63	34	3	1999-2006
9	US-Ton [*D D Baldocchi et al.*, 2010]	-120.966	38.4316	169	23	BDTmp and C3NAGrs	15	41	44	2001-2007
10	BAN[c]	-50.1591	-9.82442	120	40	BDTrop	37	24	39	2004-2006
11	CA-Oas [*Griffis et al.*, 2003]	-106.198	53.6289	530	30	BDBorl	18.8	50.32	30.87	1997-2006
12	CA-Gro [*McCaughey et al.*, 2006]	-82.1556	48.2167	300	30	BDBorl	20	65	25	2004-2006
13	US-Ho1 [*Hollinger et al.*, 1999]	-68.7403	45.2041	79	29	NETmp	15.9	50.35	33.75	1996-2004
14	CA-Ca1	-125.334	49.8672	300	45	NETmp	2.63	84.42	12.94	1998-2006

	[*Humphreys et al.*, 2006]									
15	CA-TP4 [*Arain and Restrepo-Coupe*, 2005]	-80.3574	42.7098	219	30	NETmp	0	98	2	2002-2007
16	US-NR1 [*Turnipseed et al.*, 2002]	-105.546	40.0329	3050	26	NETmp	21.43	43.13	35.45	1998-2007
17	US-Dk3 [*Katul et al.*, 2003]	-79.0942	35.9782	163	21	NETmp	13.66	51.59	34.81	1998-2005
18	US-Me2 [*Hudiburg et al.*, 2013]	-121.557	44.4524	1310	30	NETmp	7	67	26	2002-2007
19	CA-Obs [*Griffis et al.*, 2003]	-105.118	53.9872	629	30	NEBorl	4.12	80.89	14.97	2000-2006
20	CA-Qfo [*Chen et al.*, 2006]	-74.3421	49.6925	382	25	NEBorl	4	51.5	29	2004-2006
21	CA-Ojp [*Griffis et al.*, 2003]	-104.692	53.9163	579	30	NEBorl	2.5	94.47	3.02	2000-2006
22	K67[c]	-54.9589	-2.85667	130	63	BETrop	90	2	8	2002-2004
23	K83[c]	-54.9714	-3.01803	130	64	BETrop	80	18	2	2001-2003
24	RJA[c]	-61.9309	-10.0832	191	60	BETrop	10	80	10	2000-2002
25	K77[c]	-54.8944	-3.01983	130	18	Crop	80	18	2	2001-2005
26	FNS[c]	-62.3572	-10.7618	306	8.5	Crop	10	80	10	1999-2001
27	US-Ne2 [*Suyker and Verma*, 2012]	-96.4701	41.1649	362	6	Crop	31.68	30.7	37.62	2001-2006
28	US-Ne1 [*Suyker and Verma*, 2012]	-96.4766	41.1651	361	6	Crop	31.68	30.7	37.62	2001-2006
29	US-IB1 [*Matamala et al.*, 2008]	-88.2227	41.8593	225	4	Crop	37.2	7.8	55.4	2005-2007
30	US-Ne3 [*Suyker and Verma*, 2012]	-96.4397	41.1649	363	6	Crop	31.68	30.7	37.62	2001-2006
31	US-ARM [*Fischer et al.*, 2007]	-97.4884	36.605	311	65	Crop	43.1	27.98	28.92	2000-2007
32	PDG[c]	-47.6499	-21.6195	690	21	C3NAGrs and C4Grs	3	85	12	2001-2003
33	US-IB2 [*Matamala et al.*, 2008]	-88.241	41.8406	226	4	C3NAGrs and C4Grs	34.8	12.18	53	2004-2007

34	CA-Let [*Flanagan et al.*, 2002]	-112.94	49.7093	960	4	C3NAGrs	35.6	28.1	34.8	1997-2006
35	US-Var [*D D Baldocchi et al.*, 2004]	-120.951	38.4133	129	2	C3NAGrs	12.5	29.5	58	2001-2007
36	US-Shd [*Suyker et al.*, 2003]	-96.6833	36.9333	350	5	C3NAGrs	38.4	5.1	56	1997-2000
37	US-Los [*Yi et al.*, 2004]	-89.9792	46.0827	480	10	BEShr	16.43	46.56	37.01	2000-2006
38	US-SO2 [*Lipson et al.*, 2005]	-116.623	33.3739	1406	5	BEShr	21.31	43.94	34.75	1998-2006

[a] Approximate height of the wind/temperature and flux measurements above the surface.
[b] Abbreviated plant functional types (PFTs) are : BDBorl – broadleaf deciduous boreal tree; BDTmp – broadleaf deciduous temperate tree; BDTrop – broadleaf deciduous tropical tree; BEShr – broadleaf evergreen shrub; BETrop – broadleaf evergreen tropical tree; crop – C3 crop; C3NAGrs – C3 non-arctic grass; C4Grs- C4 grass; NEBorl – needleleaf evergreen boreal tree; NETmp – needleleaf evergreen temperate tree.
[c] The site information and meteorological forcing is from the LBA-MIP dataset.
[d] m_c is the number of years of atmospheric forcing.

Table 2. Comparison between the gradient projection method (in bold) and the accelerated spin-up method.

ID	Site	Number of simulation years to reach		
		$\Delta C_{TOC} \leq 3$ (g m^{-2} yr^{-1})	$\Delta C_{TOC} \leq 1$ (g m^{-2} yr^{-1})	$\Delta C_{TOC} \leq 0.5$ (g m^{-2} yr^{-1})
1	US-Ha1	416/**416**	816/**480**	1024/**624**
2	US-WCr	828/**558**	1395/**729**	1809/**828**
3	US-Syv	930/**600**	1314/**744**	1680/**924**
4	US-PFa	1375/**617**	2057/**891**	2255/**946**
5	US-UMB	387/**387**	855/**567**	1116/**567**
6	US-MOz	772/**596**	1400/**924**	1776/**1096**
7	US-Dk2	453/**408**	888/**696**	1215/**849**
8	US-MMS	872/**552**	1136/**704**	1400/**744**
9	US-Ton	413/**402**	749/**574**	959/**672**
10	BAN	1281/**1050**	1677/**1284**	1959/**1452**
11	CA-Oas	1870/**860**	2170/**1090**	2470/**1240**
12	CA-Gro	351/**351**	966/**774**	1455/**1023**
13	US-Ho1	972/**585**	1503/**756**	1845/**864**
14	CA-Ca1	597/**468**	972/**549**	1215/**558**
15	CA-TP4	798/**528**	1170/**636**	1224/**828**
16	US-NR1	1310/**740**	1910/**870**	2470/**1000**
17	US-Dk3	640/**432**	848/**472**	992/**488**
18	US-Me2	312/**312**	318/**318**	354/**354**
19	CA-Obs	509/**441**	1029/**700**	1351/**770**
20	CA-Qfo	384/**384**	819/**612**	1143/**816**
21	CA-Ojp	520/**441**	812/**539**	1143/**651**
22	K67	543/**447**	720/**510**	831/**612**
23	K83	354/**354**	477/**411**	633/**510**
24	RJA	NA[a]/**348**	NA/**510**	NA/**606**
25	K77	310/**310**	440/**425**	585/**445**
26	FNS	468/**432**	1008/**750**	1350/**954**
27	US-Ne2	954/**726**	1368/**942**	1620/**1098**
28	US-Ne1	732/**534**	1200/**714**	1506/**858**
29	US-IB1	NA/**309**	NA/**333**	NA/**459**
30	US-Ne3	312/**312**	648/**474**	948/**612**
31	US-ARM	864/**552**	1192/**616**	1400/**672**
32	PDG	309/**309**	663/**540**	1077/**807**
33	US-IB2	536/**440**	804/**588**	900/**628**
34	CA-Let	630/**450**	960/**490**	1160/**560**
35	US-Var	608/**427**	881/**651**	1568/**679**

36	US-Shd	1784/**1168**	2128/**1368**	2320/**1472**
37	US-Los	490/**427**	903/**539**	1169/**651**
38	US-SO2	NA/NA		

[a]NA – not evaluated

FIGURE CAPTIONS

Figure 1. Soil Carbon pool structure of CLM4-CN. The arrows represent the decomposition pathways, and *k* is the turnover rate of each pool.

Figure 2. Annual average total soil carbon change with respect to time (a) and the gradient projection over a shorter time interval (b).

Figure 3. Stacked bar chart of percent reduction in computation cost for three convergence threshold values.

Figure 4. Annual average total soil C with respect to time at site RJA, US-IB1 and US-SO2.

Figure 5. Comparison of water table depth (a), average soil moisture content (b), and average soil temperature (c) using original soil hydrology model and STOMP in CLM4 at site RJA.

Figure 6. Comparison of water table depth (a), average soil moisture content (b), and average soil temperature (c) using original soil hydrology model and STOMP in CLM4 at site US-IB1.

Figure 7. Comparison of water table depth simulated by original soil hydrology scheme in CLM4 (solid line) and STOMP (dashed line) at site RJA (a) and site US-IB1(b) at the last 42 years of the simulation; (c) and (d) are annual average total soil carbon at site RJA and US-IB1 using STOMP in CLM4 and the GP method.

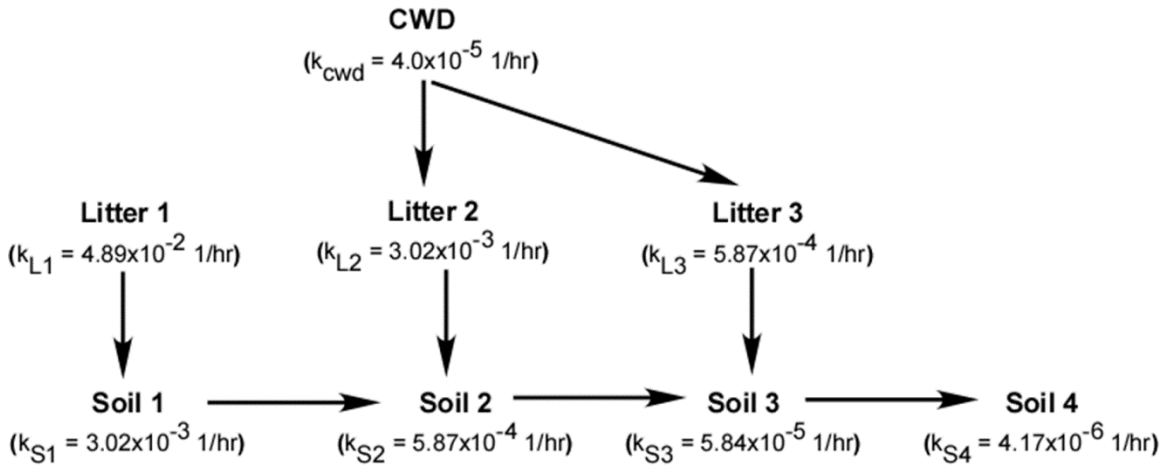

Figure 1. Soil Carbon pool structure of CLM4-CN. The arrows represent the decomposition pathways, and *k* is the turnover rate of each pool.

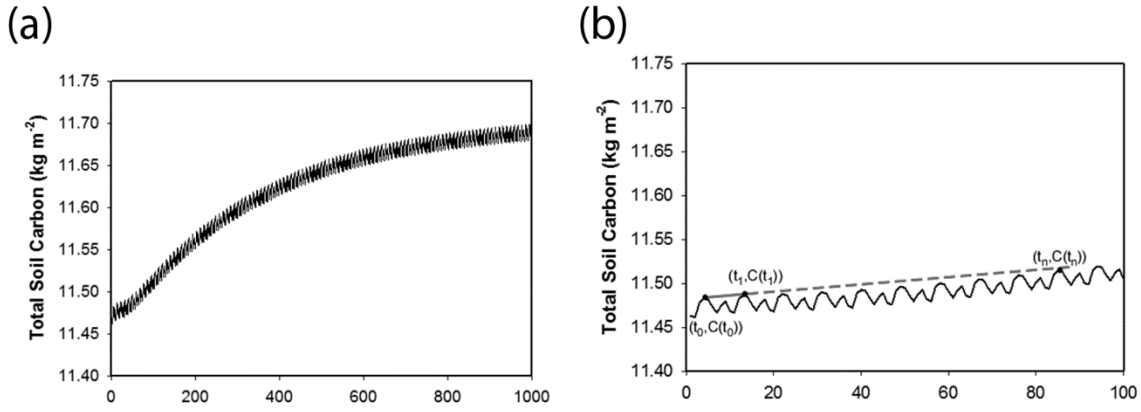

Figure 2. Annual average total soil carbon change with respect to time (a) and the gradient projection over a shorter time interval (b).

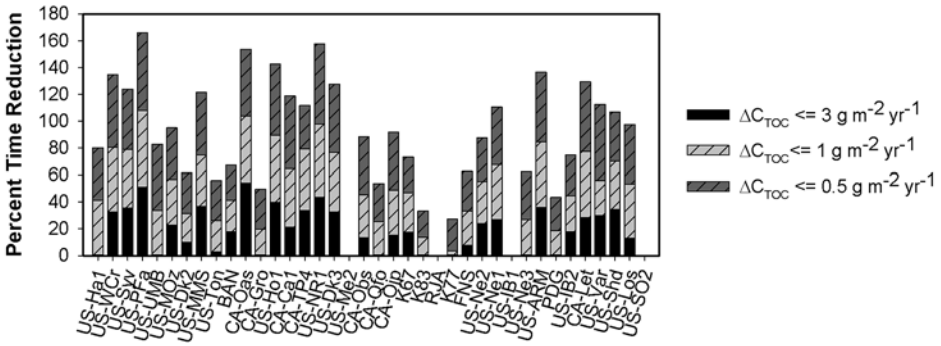

Figure 3. Stacked bar chart of percent reduction in computation cost for three convergence threshold values.

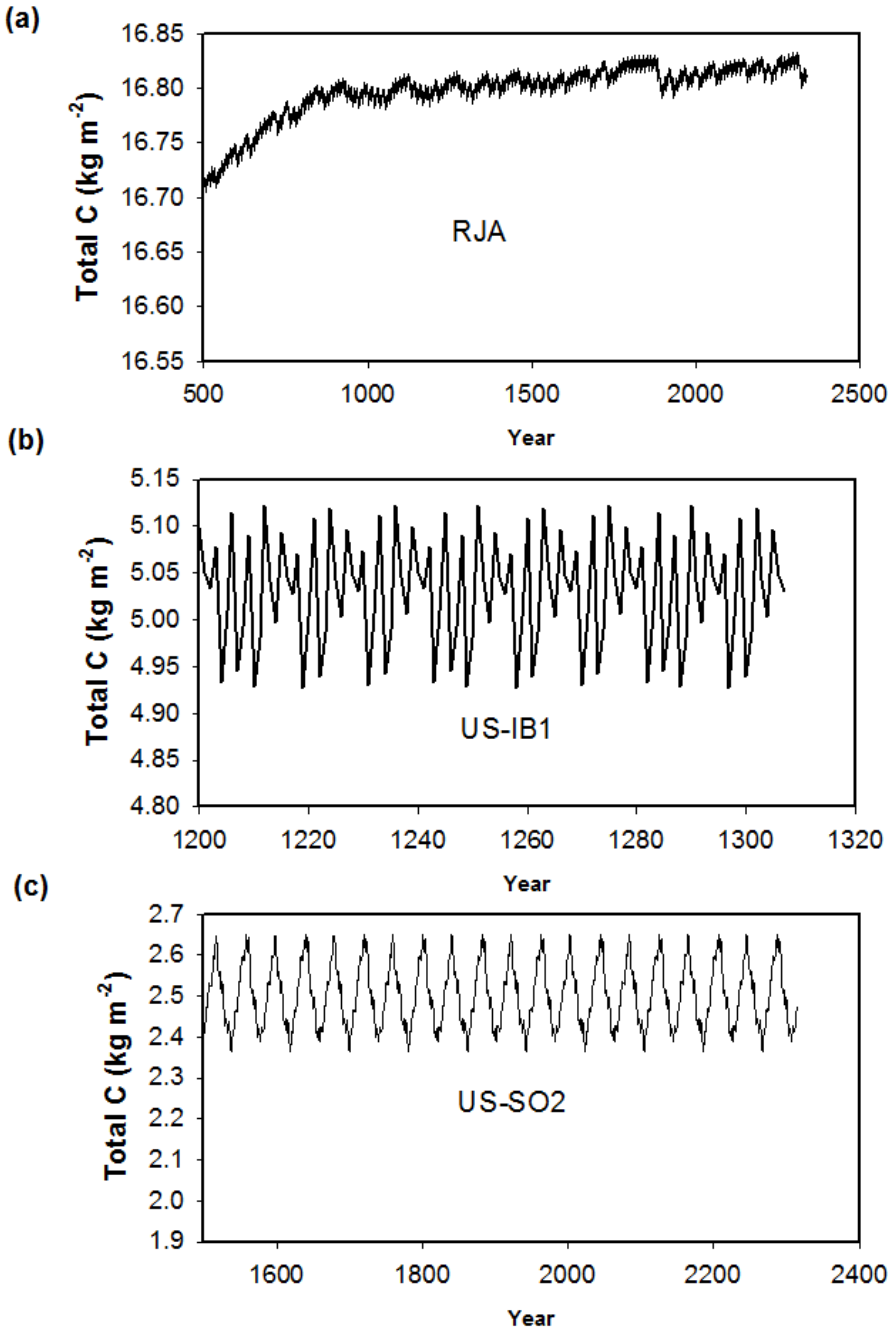

Figure 4. Annual average total soil C with respect to time at site RJA, US-IB1 and US-SO2.

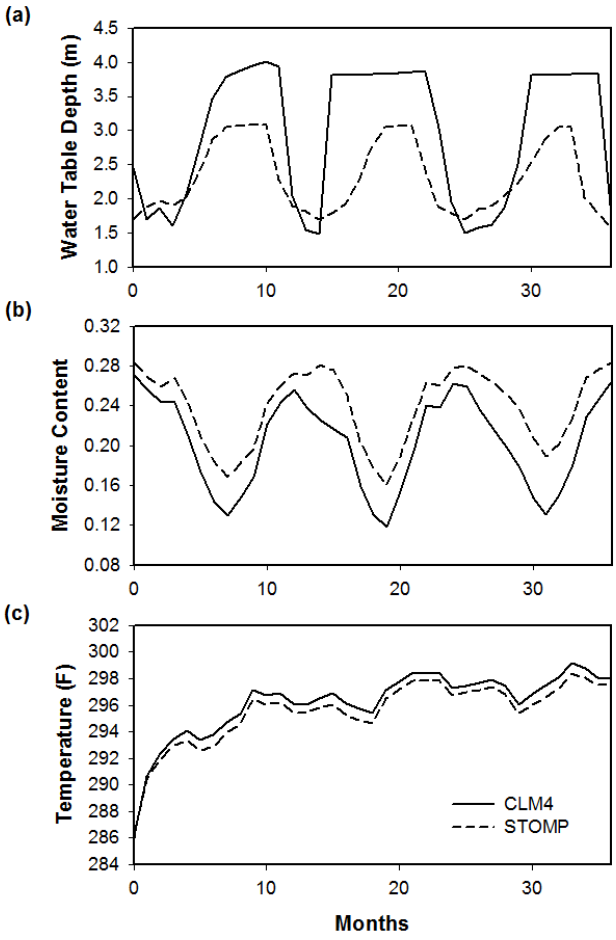

Figure 5. Comparison of water table depth (a), average soil moisture content (b), and average soil temperature (c) using original soil hydrology model and STOMP in CLM4 at site RJA.

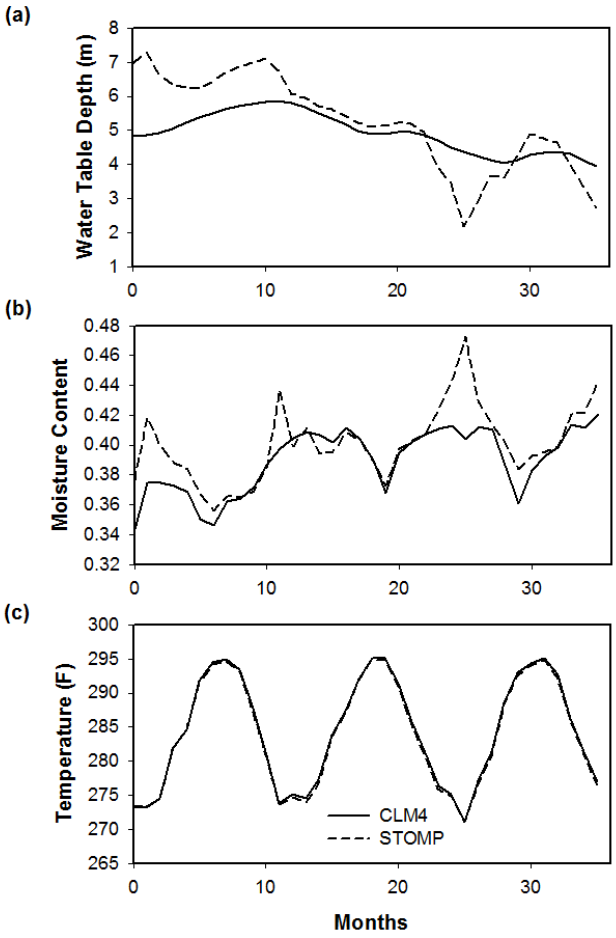

Figure 6. Comparison of water table depth (a), average soil moisture content (b), and average soil temperature (c) using original soil hydrology model and STOMP in CLM4 at site US-IB1.

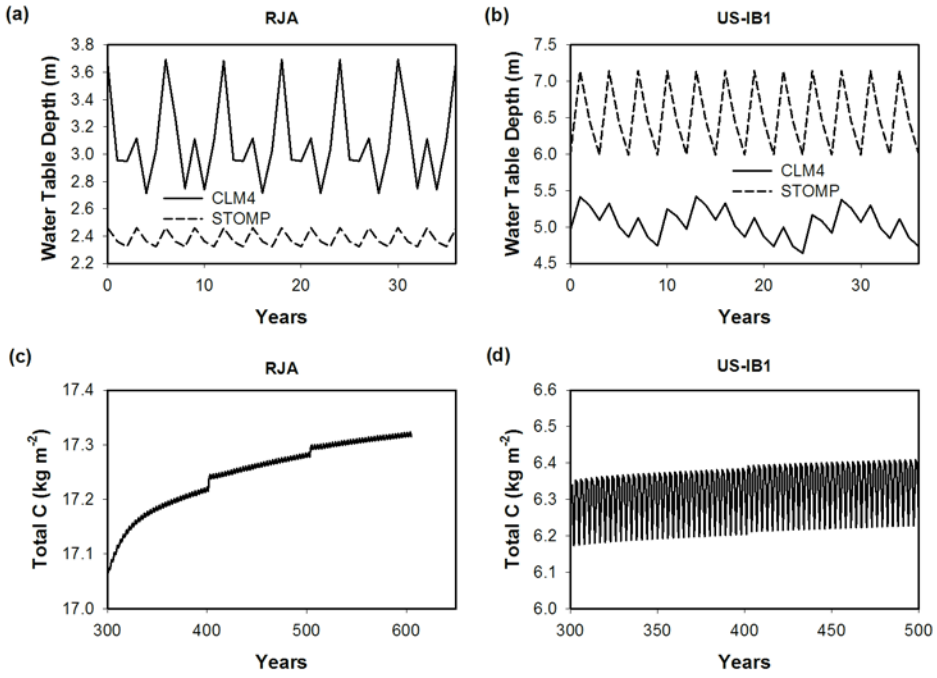

Figure 7. Comparison of water table depth simulated by original soil hydrology scheme in CLM4 (solid line) and STOMP (dashed line) at site RJA (a) and site US-IB1(b) at the last 42 years of the simulation; (c) and (d) are annual average total soil carbon at site RJA and US-IB1 using STOMP in CLM4 and the GP method.