



Global Biogeochemical Cycles

RESEARCH ARTICLE

10.1002/2015GB005239

Key Points:

- First-order decay functions capture macroscopic SOC dynamics, but their underpinnings need more study
- Optimal parameter calibration through data assimilation is a high priority to reduce model biases
- Traceability analyses are needed to understand the consequences of variation in driving variables

Correspondence to:

Y. Luo,
yluo@ou.edu

Citation:

Luo, Y., et al. (2016), Toward more realistic projections of soil carbon dynamics by Earth system models, *Global Biogeochem. Cycles*, 30, 40–56, doi:10.1002/2015GB005239.

Received 7 JUL 2015

Accepted 17 DEC 2015

Accepted article online 18 DEC 2015

Published online 21 JAN 2016

Toward more realistic projections of soil carbon dynamics by Earth system models

Yiqi Luo^{1,2}, Anders Ahlström^{3,4}, Steven D. Allison^{5,6}, Niels H. Batjes⁷, Victor Brovkin⁸, Nuno Carvalhais^{9,10}, Adrian Chappell¹¹, Philippe Ciais¹², Eric A. Davidson¹³, Adien Finzi¹⁴, Katerina Georgiou^{15,16}, Bertrand Guenet¹², Oleksandra Hararuk¹⁷, Jennifer W. Harden¹⁸, Yujie He⁶, Francesca Hopkins⁶, Lifan Jiang¹, Charlie Koven¹⁶, Robert B. Jackson³, Chris D. Jones¹⁹, Mark J. Lara²⁰, Junyi Liang¹, A. David McGuire^{21,20}, William Parton²², Changhui Peng²³, James T. Randerson⁶, Alejandro Salazar²⁴, Carlos A. Sierra⁹, Matthew J. Smith²⁵, Hanqin Tian²⁶, Katherine E. O. Todd-Brown²⁷, Margaret Torn¹⁶, Kees Jan van Groenigen²⁸, Ying Ping Wang²⁹, Tristram O. West³⁰, Yaxing Wei³¹, William R. Wieder³², Jianyang Xia³³, Xia Xu³⁴, Xiaofeng Xu³⁵, and Tao Zhou³⁶

¹Department of Microbiology and Plant Biology, University of Oklahoma, Norman, Oklahoma, USA, ²Center for Earth System Science, Tsinghua University, Beijing, China, ³Department of Earth System Science, Stanford University, Stanford, California, USA, ⁴Department of Physical Geography and Ecosystem Science, Lund University, Lund, Sweden, ⁵Department of Ecology and Evolutionary Biology, University of California, Irvine, California, USA, ⁶Department of Earth System Science, University of California, Irvine, California, USA, ⁷ISRIC-World Soil Information, Wageningen, Netherlands, ⁸Max Planck Institute for Meteorology, Hamburg, Germany, ⁹Max Planck Institute for Biogeochemistry, Jena, Germany, ¹⁰CENSE, Departamento de Ciências e Engenharia do Ambiente, Faculdade de Ciências e Tecnologia, Universidade NOVA de Lisboa, Caparica, Portugal, ¹¹CSIRO Land and Water National Research Flagship, Canberra, Australia, ¹²Laboratoire des Sciences du Climat et de l'Environnement, CEA CNRS UVSQ, Gif-sur-Yvette, France, ¹³Appalachian Laboratory, University of Maryland Center for Environmental Science, Frostburg, Maryland, USA, ¹⁴Department of Biology, Boston University, Boston, Massachusetts, USA, ¹⁵Department of Chemical and Biomolecular Engineering, University of California, Berkeley, California, USA, ¹⁶Earth Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California, USA, ¹⁷Pacific Forestry Centre, Canadian Forest Service, Victoria, British Columbia, Canada, ¹⁸U.S. Geological Survey, Menlo Park, California, USA, ¹⁹Met Office Hadley Centre, Exeter, UK, ²⁰Institute of Arctic Biology, University of Alaska Fairbanks, Fairbanks, Alaska, USA, ²¹U.S. Geological Survey, Alaska Cooperative Fish and Wildlife Research Unit, Fairbanks, Alaska, USA, ²²Natural Resource Ecology Laboratory, Colorado State University, Fort Collins, Colorado, USA, ²³Institute of Environment Sciences, University of Quebec at Montreal, Montreal, Quebec, Canada, ²⁴Department of Biological Sciences, Purdue University, West Lafayette, Indiana, USA, ²⁵Computational Science Laboratory, Microsoft Research, Cambridge, UK, ²⁶International Center for Climate and Global Change Research, School of Forestry and Wildlife Sciences, Auburn University, Auburn, Alabama, USA, ²⁷Biological Sciences Division, Pacific Northwest National Laboratory, Richland, Washington, USA, ²⁸Center for Ecosystem Science and Society, Northern Arizona University, Flagstaff, Arizona, USA, ²⁹CSIRO Ocean and Atmosphere Flagship, Aspendale, Victoria, Australia, ³⁰Joint Global Change Research Institute, College Park, Maryland, USA, ³¹Environmental Sciences Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA, ³²National Center for Atmospheric Research, Boulder, Colorado, USA, ³³School of Ecological and Environmental Sciences, East China Normal University, Shanghai, China, ³⁴Department of Ecology, Evolution and Organismal Biology, Iowa State University, Ames, Iowa, USA, ³⁵Department of Biological Science, University of Texas at El Paso, El Paso, Texas, USA, ³⁶State Key Laboratory of Earth Surface Processes and Resource Ecology, Beijing Normal University, Beijing, China

Abstract Soil carbon (C) is a critical component of Earth system models (ESMs), and its diverse representations are a major source of the large spread across models in the terrestrial C sink from the third to fifth assessment reports of the Intergovernmental Panel on Climate Change (IPCC). Improving soil C projections is of a high priority for Earth system modeling in the future IPCC and other assessments. To achieve this goal, we suggest that (1) model structures should reflect real-world processes, (2) parameters should be calibrated to match model outputs with observations, and (3) external forcing variables should accurately prescribe the environmental conditions that soils experience. First, most soil C cycle models simulate C input from litter production and C release through decomposition. The latter process has traditionally been represented by first-order decay functions, regulated primarily by temperature, moisture, litter quality, and soil texture. While this formulation well captures macroscopic soil organic C (SOC) dynamics, better understanding is needed of their underlying mechanisms as related to microbial processes, depth-dependent environmental controls, and other processes that strongly affect soil C dynamics. Second, incomplete use of observations in model parameterization is a major cause of bias in soil C projections from ESMs. Optimal parameter calibration with both pool- and flux-based data sets through data assimilation is

among the highest priorities for near-term research to reduce biases among ESMs. Third, external variables are represented inconsistently among ESMs, leading to differences in modeled soil C dynamics. We recommend the implementation of traceability analyses to identify how external variables and model parameterizations influence SOC dynamics in different ESMs. Overall, projections of the terrestrial C sink can be substantially improved when reliable data sets are available to select the most representative model structure, constrain parameters, and prescribe forcing fields.

1. Introduction

Soils contain the largest terrestrial pool of organic carbon (C) [Jobbágy and Jackson, 2000; Tarnocai *et al.*, 2009], yet their representation in Earth system models (ESMs) currently contributes substantial uncertainty to C cycle and climate projections in the Earth system [Jones *et al.*, 2003, 2006; Arora *et al.*, 2013; Todd-Brown *et al.*, 2013]. When C cycle feedbacks are expressed in a common framework with other feedbacks in the climate system, such as clouds or ice albedo effects, the magnitude of their uncertainty is comparable to cloud feedbacks, which have long been regarded as the most significant uncertainty in climate modeling [Gregory *et al.*, 2009]. Particularly relevant to the subject of our analysis in this paper, Jones and Falloon [2009] found a strong relationship between changes in soil organic C (SOC) and the strength of simulated C-climate feedbacks within ESMs. Thus, it is important to improve representation of soil processes and feedbacks in ESMs. This paper aims to reveal possible ways of improving model projections.

Representation of soil C dynamics in ESMs requires knowledge in soil science and model development. In the discipline of soil science, climate has long been recognized as one of the primary factors determining distributions of soil C and nitrogen (N) stocks [Jenny, 1961]. Soil microorganisms were recognized as the main agents of decomposition of soil organic matter, and the effects of substrate quality on decomposition rates, including C-N interactions, have been studied for nearly a century [Tenney and Waksman, 1929; Waksman, 1952]. However, we continue to struggle with the challenge of applying our knowledge about these basic principles as we develop sophisticated numerical models to represent the feedbacks between climate and soil C stocks. For example, soil C stabilization, although long recognized as a key emergent property for understanding C stocks and flows, is still a subject of much research and has not yet been explicitly integrated into numerical soil biogeochemistry models [Kleber *et al.*, 2007; Schmidt *et al.*, 2011; Davidson *et al.*, 2014].

The initial development of soil organic C models, such as the RothC and CENTURY models [Jenkinson *et al.*, 1987; Parton *et al.*, 1987], began in the 1980s to simulate the effects of different agricultural practices on crop yields, soil C and N dynamics, and nutrient cycling for long-term agricultural experimental sites. During the 1990s, when the effects of environmental change (e.g., elevated CO₂ and climatic change) on ecosystem dynamics became major research questions, soil C models were integrated into many ecosystem models to simulate plant production, soil C dynamics, trace gas fluxes, and nutrient cycling in response to climate change for all of the major ecosystems in the world [Schimel *et al.*, 1996]. Global models of the terrestrial C cycle were developed about two decades ago with the primary goal of projecting changes in terrestrial C storage under increasing atmospheric CO₂ concentration and climate change [Melillo *et al.*, 1993; Haxeltine and Prentice, 1996; McGuire *et al.*, 1997]. Simulation of SOC dynamics has been mainly based on existing models such as CENTURY [Parton *et al.*, 1993], together with dynamic global vegetation models [Cramer *et al.*, 2001], land use scenarios [Hurtt *et al.*, 2011], and N cycling [Edburg *et al.*, 2011] in the Coupled Model Intercomparison Project Phase 5 (CMIP5). Those models have been used to determine the strength of climate-C feedbacks [Arora *et al.*, 2013] and anthropogenic CO₂ emissions compatible with a target CO₂ pathway [Jones *et al.*, 2013].

Various evaluation studies of CMIP5 results show that ESMs give widely different projections of soil C dynamics and poorly fit observations [Todd-Brown *et al.*, 2013, 2014; Carvalhais *et al.*, 2014; Yan *et al.*, 2014]. Furthermore, the simulated contemporary soil C stock varies from 510 to 3040 GtC among the 11 ESMs [Todd-Brown *et al.*, 2013]. Similar magnitudes of uncertainties were also found for the overall land C models in other studies [Friedlingstein *et al.*, 2006; Arora *et al.*, 2013]. This variability is not well constrained by modern benchmarks and is reflected in the future scenarios [Todd-Brown *et al.*, 2014]. On the other hand, theoretical analysis suggests that many of the soil C processes are intrinsically predictable given knowledge of the initial conditions, carbon input rates, soil carbon residence times, and their environmental sensitivities [Luo *et al.*, 2015]. Thus, it is still possible that projections of soil C dynamics by the current ESMs can be significantly improved through reduction of biases among models and better fit with observations.

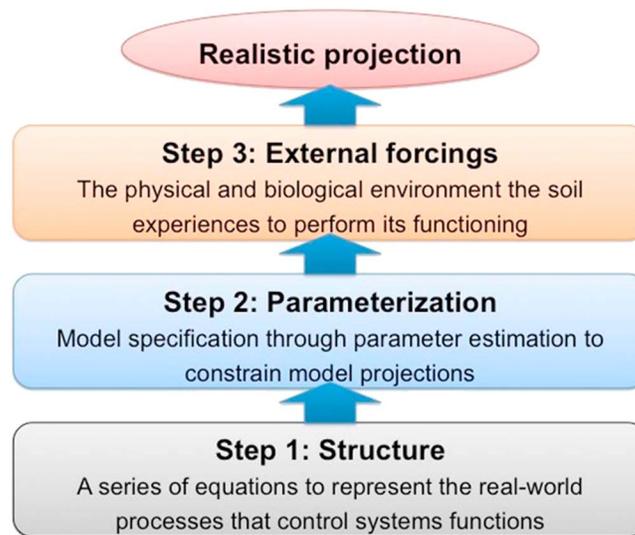


Figure 1. Three modeling components toward realistic projections of soil C dynamics. Model structure formulation determines patterns of model behavior, parameterization completes model specification to constrain model simulations, and forcing variables specify the environments that soil experiences.

2013; Wang *et al.*, 2014a]. The basic patterns described by a model must be consistent with empirical evidence, as ESMs are ultimately used to predict real-world phenomena. Once the model structure is defined, coefficients in the equations are assigned through parameterization, a process necessary to complete model specification. While model structure tends to define the range of possible trajectories, the choice of parameter values for a given model defines the quantitative accuracy in specific simulations of SOC dynamics. Thus, even with the same model structure, differences in parameter values can generate divergent modeling results. Realistic projections of soil C dynamics also require more reliable external variables, representing the physical and biological environment the soil experiences. As C cycle processes are highly sensitive to environmental change, biases in soil forcings can lead to biases in model projections. Note that while some variables, such as temperature and precipitation, are treated as forcings when land models alone are used for simulation, they can be prognostic variables in coupled simulations of ESMs. To focus the discussion of this paper on modeling SOC dynamics, we refer to all variables that are exogenous to the soil C cycle as external variables no matter how they are represented in a model.

This paper examines each of the three modeling components in terms of their representation in the current generation of ESMs, recent advances, and our recommendations for future research directions. We first show that extensive exploration of alternative model structures in the past three decades has mostly conformed to a similar formulation with C input from litter production and soil C transformation being processed through a series of first-order decay processes. Vast results from experimental and observational studies need to be synthesized to evaluate whether the model formulation should be altered or whether the model parameters should be varied with soil attributes and external variables. Second, we argue that advances in parameter estimation methods make them ready for use in model calibration as standard practice to reduce systematic biases among models. Third, we demonstrate that external variables are represented in ESMs in different ways and potentially cause major variations among model projections. To disentangle the complex impacts of external variables, it is necessary to trace what external variables are represented and how they are manifested in the models.

2. Structure of Soil C Models

2.1. Current Model Structures

The major processes governing SOC dynamics in an ecosystem are C inputs, transformations, and losses from heterotrophic respiration. The primary C inputs are through litterfall at the soil surface, rhizodeposition, and

Fundamentally, model projections of SOC dynamics rely on three components: model structure, parameterization (including initial values of C pools), and external variables (Figure 1) [Raupach and Lu, 2004; Luo *et al.*, 2011]. The poor performance of current ESMs can result from biases in any of the three components across time, space, and soil depth. Model structure is the set of equations used to describe dynamic patterns of SOC processes. For example, a first-order decay function with fixed coefficients describes the monotonic decrease over time of SOC that has entered a soil pool. More complex behaviors can arise when the decay coefficients vary with time- and site-dependent microbial activities or alternative functional forms, such as Michaelis-Menten kinetics, are incorporated [Li *et al.*,

root and rhizome death in the soil. The C inputs are usually simulated via the processes of litter production in all the SOC models, which mostly equal net primary production (NPP) after losses to fire, harvesting, grazing, and photodegradation. Once in the soil, organic C is transformed through the processes of decomposition, mineralization, and stabilization. *Henin and Dupuis* [1945] first described the decomposition of organic C with a first-order decay function of the C mass. Combining C inputs (denoted $\mu(t)$) with decomposition yields a model of soil C dynamics as

$$\frac{dX(t)}{dt} = u(t) - kX(t), \quad (1)$$

where $X(t)$ is the mass remaining at time t and k is a decomposition coefficient. When organic matter in soil is divided into multiple pools to represent different physicochemical properties, the C transfer among pools can be summarized by a matrix equation [*Bolker et al.*, 1998; *Luo et al.*, 2003, 2015; *Sierra et al.*, 2012] as

$$\begin{cases} \frac{dX(t)}{dt} = Bu(t) - A\zeta(t)KX(t) \\ X(t=0) = X_0 \end{cases}, \quad (2)$$

where $X(t)$ is a vector of pool sizes, B is a vector of partitioning coefficients among plant pools, A is a square matrix of transfer coefficients, $\zeta(t)$ is a diagonal matrix of environmental scalars, K is a diagonal matrix of exit rates, and X_0 is a vector of initial pool sizes. The exit rate is the first-order decomposition rate. The sum of all the column elements of each row of matrix A multiplied by -1 corresponds to mineralization of decomposed C to CO_2 via respiration for each pool [*Luo et al.*, 2001, 2003].

Overall, equation (2) can conceptually express all of the soil C transformation processes and summarize structures of classic SOC models, such as the CENTURY [*Parton et al.*, 1987, 1988, 1993] and RothC models [*Jenkinson et al.*, 1987], as well as those embedded in ESMs [*Ciais et al.*, 2013]. Also, this model structure is generally consistent with five fundamental properties of the terrestrial C cycle: compartmentalization, C input through photosynthesis and subsequent plant tissue senescence and mortality, partitioning among pools, donor pool-dominant transfers, and first-order decay [*Luo and Weng*, 2011; *Luo et al.*, 2015]. Thousands of data sets published in the literature from litter decomposition and soil incubation studies have been used to obtain first-order decay parameters that can be used in ESMs [*Zhang et al.*, 2008; *Schädel et al.*, 2013, 2014; *Liang et al.*, 2015]. The scalar function, $\zeta(t)$, in equation (2) represents the environmental modifier for decomposition and transfer rates with respect to changes in temperature, moisture, litter quality, and soil texture. Empirical studies have also indicated that temperature, moisture, litter quality, and soil texture are primary factors that control soil C decomposition and stabilization [*Burke et al.*, 1989; *Adair et al.*, 2008; *Zhang et al.*, 2008; *Xu et al.*, 2012; *Wang et al.*, 2013a].

2.2. Recent Advances

If the fundamental processes of SOC dynamics are straightforward (i.e., C inputs and transformations before release as CO_2) and can be described by a relatively simple set of equations, why do ESMs disagree so much in projecting global soil C dynamics? To answer this question, we need to understand what we have learned from empirical studies on SOC transformation and C input.

SOC transformations are regulated by environmental variables (e.g., temperature, moisture, oxygen, N, phosphorus, and acidity varying with soil profile, space, and time), litter quality (e.g., lignin, cellulose, N, or their relative content), organomineral properties of SOC (e.g., complex chemical compounds, aggregation, physicochemical binding and protection, reactions with inorganic, reactive surfaces, and sorption), and microbial attributes (e.g., community structure, functionality, priming, acclimation, and other physiological adjustments) (Figure 2). Meanwhile, C inputs are also regulated by plant allocation strategies (e.g., root/shoot ratio) and root biology (e.g., rooting depth, rhizodeposition, and symbiotic relationship with mycorrhizae). In addition, both C inputs and transformations are influenced by soil erosion, mineralogy [*Egli et al.*, 2008; *Dümig et al.*, 2011; *Sistla and Schimel*, 2013; *Doetterl et al.*, 2015], topography, land management, land use change, and other disturbances. While we apparently cannot incorporate all those factors and processes into one model, the art of soil modeling is to determine what should be explicitly represented in models and what can be ignored. As no consensus has been developed on what should be explicitly represented in models, the modeling community has explored several types of alternative model structures from equation (2).

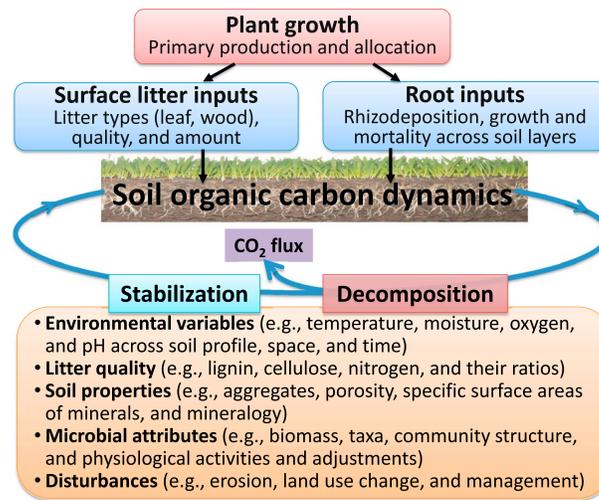


Figure 2. Major processes and factors that determine SOC dynamics. The complex behaviors of SOC dynamics primarily arise from C input and SOC transformations being regulated by environmental and biological variables.

First, fundamentally different from the first-order, linear C transfer model are the nonlinear microbe-explicit models that have been developed to account for microbial roles in decomposition and stabilization of SOC [Wieder *et al.*, 2015]. These models have the potential to explain priming effects [Kuzyakov, 2010], acclimation [Luo *et al.*, 2001; Bradford *et al.*, 2010], and pulse responses of soil respiration to wet-dry cycles of precipitation [Liu *et al.*, 2002; Lawrence *et al.*, 2009]. Many microbial models are based on forward or reverse Michaelis-Menten equations [Schimel and Weintraub, 2003; Allison *et al.*, 2010; Wang *et al.*, 2013b; Wieder *et al.*, 2013, 2014] that mathematically couple microbial biomass to C substrate pools. Such

model formulations generate markedly different patterns of soil C dynamics, such as oscillatory responses to perturbation and insensitivity of soil C storage to C input, in comparison with classic models [Li *et al.*, 2014; Wang *et al.*, 2014b; Hararuk *et al.*, 2015]. Such patterns may exist at the microbial reaction sites but have not been observed in litter decomposition and soil incubation studies. Other ways to represent microbial processes include multiplication of decomposition coefficients by microbial biomass [Fujita *et al.*, 2014], making decomposition a function of substrate chemistry and enzyme-based microbial guilds [Moorhead and Sinsabaugh, 2006], and embedding soil enzyme dynamics in an ecosystem model [Sistla and Schimel, 2013]. Such models may avoid undesirable oscillations of nonlinear microbial models yet account for microbial roles in SOC decomposition. Nevertheless, it has also been argued that decomposition of SOC is more likely to be limited by substrate availability than microbial activities [Schimel and Schaeffer, 2012].

The second category of alternative model structures is related to differences in the way that pools are defined. Soil C pools, which are conceptually devised with different decomposability in the classic SOC models [Jenkinson, 1990; Parton *et al.*, 1993], have been replaced by measurable SOC fractions [Smith *et al.*, 2002; Stewart *et al.*, 2008; Luo *et al.*, 2014] or continuous functions [Ågren and Bosatta, 1996]. A key motivation for this change in definitions was the need to reconcile models with observations of ¹³C and data. Many long-lived compounds are in principle highly decomposable, implying that the idea of a fixed distribution of turnover times as in classic soil models is too simplistic [Schmidt *et al.*, 2011]. A vertically resolved soil biogeochemical scheme has been introduced with mixing of soil C and N among soil layers due to bioturbation, cryoturbation, and diffusion [Koven *et al.*, 2013]. The soil C dynamics in each layer can still be represented with the classic model structure, and as a result, inclusion of the vertical dimension does not alter the fundamental behaviors of the models. The soil C storage in vertically resolved models is still determined jointly by C influx and residence times.

Third, while the model formulation to simulate C transfer may be similar to the classic models, different response functions (i.e., different $\zeta(t)$ in equation (2)) are used to simulate C cycle responses to external variables. For example, temperature modifies almost all processes in the C cycle. A variety of formulations, including exponential, Arrhenius, and optimal response functions, have been used to describe C cycle responses to temperature changes in different models [Lloyd and Taylor, 1994; Jones *et al.*, 2005; Sierra *et al.*, 2015a]. Similarly, influences of soil water content on the C cycle are represented by empirically derived coefficients to modify temperature response functions [e.g., Falloon *et al.*, 2011; Moyano *et al.*, 2012, 2013] or limitations of substrate supply for enzymatic processes due to diffusion through water films of varying thickness as soil moisture varies [Davidson *et al.*, 2012]. Different response functions are used to link C cycle processes with nutrient availability, soil clay content, litter quality, and many other environmental conditions.

Differences in the use of individual response functions may not change basic dynamics but can substantially contribute to the uncertainty in model projections [Exbrayat *et al.*, 2013].

Fourth, disturbance events may be represented in models in different ways [Grosse *et al.*, 2011; West *et al.*, 2011; Goetz *et al.*, 2012; Hicke *et al.*, 2012]. Soil erosion, for example, can be modeled with horizontal movement of C, adding a third dimension to classic two-dimensional models [Rosenbloom *et al.*, 2006]. Other disturbances, such as fire, extreme drought, insect outbreaks, land management, and land cover and land use change can be represented in models to influence soil C dynamics via (1) modifying soil and microclimatic environments; (2) transferring C from one pool to another (e.g., from live to dead pools during storms or release to the atmosphere with fire); and (3) altering rates of C processes, for example, gross primary productivity (GPP), growth, tree mortality, or heterotrophic respiration [Kloster *et al.*, 2010; Thonicke *et al.*, 2010; Luo and Weng, 2011; Prentice *et al.*, 2011; Weng *et al.*, 2012]. Although many disturbance events can be incorporated into classic models without changing the basic formulation (i.e., equation (2)) [see, e.g., Weng *et al.*, 2012], the structure of a model with disturbances represented is different from that without, leading to different simulated SOC dynamics.

Lastly, model structures are commonly considered altered even if the model formulation is the same when different processes are represented to simulate environments that soils experience. Most of the ESMs involved in CMIP5 underestimate C storage in wetland and peatland regions [Limpens *et al.*, 2008] because C dynamics in wetlands or peatlands are not simulated. Different rates of decomposition under anaerobic conditions in the wetlands and peatlands are eventually expressed through changes in parameters of equation (2) instead of its formulation. Similarly, dramatically increased decomposition rates of thawed soil C in permafrost regions can be represented by different parameter values in models that explicitly simulate permafrost processes [Hobbie *et al.*, 2000; Koven *et al.*, 2013; Schädell *et al.*, 2014]. More generally, model structural uncertainty can arise from many different external variables being represented although all the models may use a similar formulation as in equation (2).

In sum, past studies have extensively explored alternative model structures. Except the nonlinear microbial models and the three-dimension erosion models, all the alternative structures conform the basic formulation: C inputs from litter production and soil C transformation are processed through a series of first-order kinetic transfer functions between pools (equation (2)). Those processes and factors as depicted in Figure 2, or even different ways of defining pools, all influence soil C dynamics mainly through their effects on rate parameters and state variables of equation (2). However, how the rate parameters and state variables are affected by those variables is not well specified.

2.3. Future Research Directions

It is urgent to synthesize experimental and observational results to better represent soil carbon cycling processes in ESMs. Thousands of papers have been published from observational and experimental studies on many processes and factors that influence soil carbon dynamics. Although empirical results from litter decomposition and soil incubation studies have been partly synthesized to verify the first-order kinetic transfer [Zhang *et al.*, 2008; Xu *et al.*, 2016], the majority of data published in the literature has not been integrated to help modelers decide which processes should be explicitly represented.

Data synthesis is particularly crucial to evaluate how microbial processes should be incorporated into ESMs (Figure 3). Microorganisms have long been known to catalyze almost all the SOC transformation processes (e.g., decomposition, stabilization, and mineralization). Developing microbial models to project SOC dynamics will require firm empirical evidence to address two questions: (1) how do microbial functions vary with environmental factors? and (2) does that variation significantly affect decomposition and other key soil processes [Schimel, 2001]? Although many microbial models have been proposed to explore possible microbial roles in SOC dynamics [Wieder *et al.*, 2015], these models need rigorous evaluation with observations before they can be incorporated into ESMs.

In addition, more and more studies indicate that multilayer soil C models are needed to account for depth-dependent variations in C inputs and decomposition. For instance, radiocarbon ages of soils in most climate zones increase rapidly with depth [Mathieu *et al.*, 2015], implying that some properties of soil C dynamics may be different in shallow versus deep layers. In addition, decomposition rates in permafrost-affected soils drop off abruptly with depth as soils transition from a seasonally thawed active layer at the surface to permanently frozen layers.

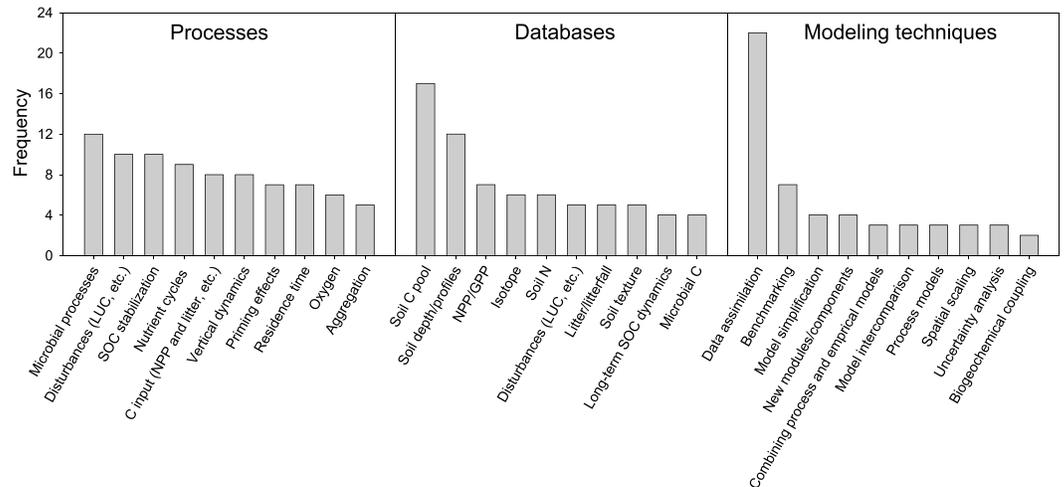


Figure 3. Consensus on recommended top 10 processes, databases, and modeling techniques to be studied in the coming 5–10 years to improve model projections of SOC dynamics. The survey was conducted with 49 participants at the workshop “Representing soil carbon dynamics in global land models to improve future IPCC assessments” held at Breckenridge, CO, USA on 12–14 June 2014. Each participant was asked to make five or less recommendations that should be easily implemented yet effective to (1) improve fitness between modeled and observed grid-level values of soil C density, (2) reduce model-model differences, and (3) represent real-world soil C dynamics in ESMs as realistically as possible.

Moreover, many processes associated with disturbances, land management, dynamic vegetation, and nutrients all potentially have strong effects on soil C dynamics. Soil C processes in wetlands and peatlands are also strongly regulated by anaerobic environmental conditions, leading to distinctive SOC dynamics from uplands. Ideally, all the important processes should be incorporated into soil C models. Practically, it is impossible for any models to do so. Thus, it is a challenge to decide what processes are to be included versus omitted.

3. Parameterization of Soil C Models

3.1. Current Practices

Parameterization is a major cause of the mismatch between modeled and observed soil C stocks. For example, global soil C stocks varied from 510 Gt C to 3040 Gt C among 11 models involved in CMIP5 [Todd-Brown *et al.*, 2013] and ranged from 425 to 2111 Gt C among 10 terrestrial biosphere models in the Multi-scale Synthesis and Terrestrial Model Intercomparison Project [Tian *et al.*, 2015]. The large differences in modeled soil C stocks were attributed to threefold differences in C inputs and fourfold differences in residence times [Todd-Brown *et al.*, 2013; Carvalhais *et al.*, 2014; Yan *et al.*, 2014]. Moreover, ESMs simulated different responses of land C storage and emissions to land use, partly due to different parameterizations of land surface processes [Brovkin *et al.*, 2013].

Parameterization is among the least studied components of modeling and has been analyzed relatively crudely [Luo *et al.*, 2001]. It is well known among modelers that a model with well-calibrated parameters at one site may not reliably estimate SOC dynamics at other sites unless the parameters are adjusted again [Xiao *et al.*, 2014]. Parameter calibration has been traditionally used to fit model output to observations in simulation models. It is sometimes based on data from observations and experiments so that parameter values can be specified within observed ranges when data can be directly converted to parameter values, such as specific rates of litter decomposition [Zhang *et al.*, 2008]. Parameter values can alternatively be set with an educated guess when they are not measurable in experiments due to technical limitations, such as root exudation. In addition, parameter values can also be set to generate reasonable pool sizes, such as fine root biomass as a result of root growth and mortality. Or parameter values can sometimes be chosen to ensure model stability.

Even though it has long been a common practice in the ecological modeling community, the calibration of parameters in ESMs often is not practical. It becomes very difficult to identify parameters in a complicated model that can be effectively calibrated to fit data well across diverse landscapes. Without rigorous parameter calibration, together with the omission of some processes, all ESMs or global land models show

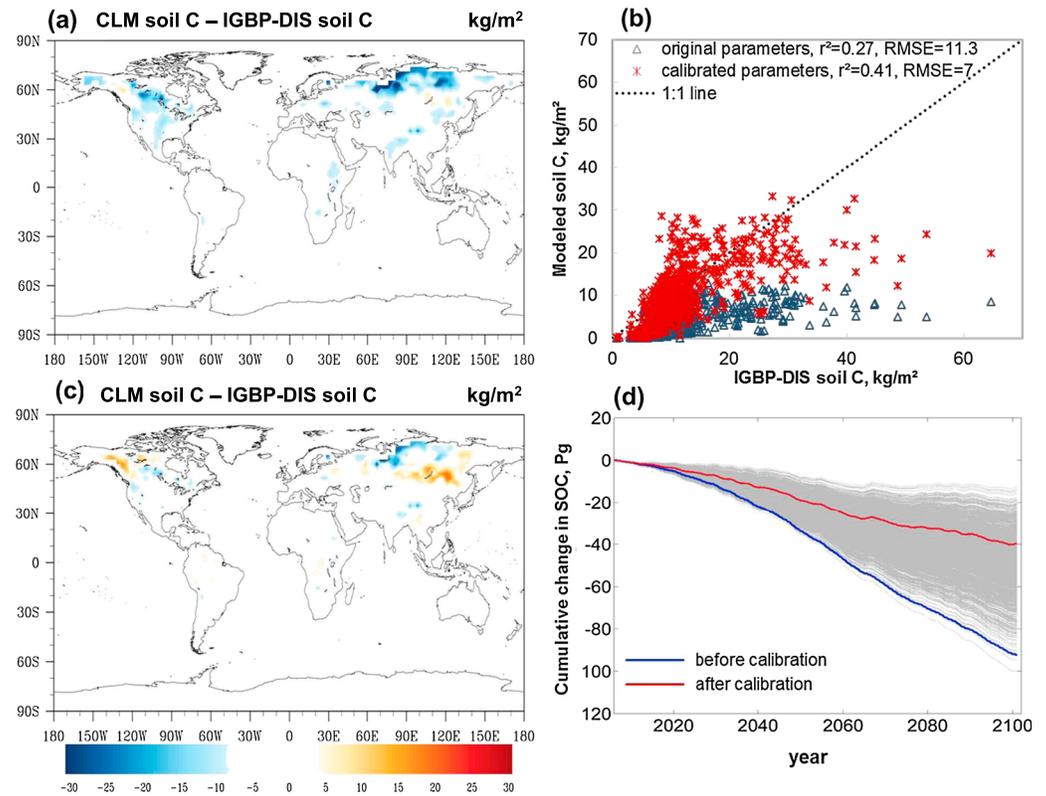


Figure 4. Spatial correspondence of CLM-CASA' produced SOC to the IGBP-DIS reported SOC (a, b) before and (b, c) after data assimilation and (d) SOC responses to RCP8.5 climate change scenario before and after calibration. The points in Figure 4b represent the grid cell values. Model with default parameters explained 27% of variation in the observed soil C, whereas model with calibrated parameters explained 41% of variability in the observed soil C (adopted from Hararuk *et al.* [2014]).

systematic biases in modeled soil C stocks [Luo *et al.*, 2015]. Spatially, SOC estimates from individual models are poorly correlated with the harmonized world soil database (HWSD) [Food and Agriculture Organization *et al.*, 2012], showing either systematic underestimation or overestimation [Todd-Brown *et al.*, 2013; Tian *et al.*, 2015], although it should be noted that there is considerable uncertainty in the HWSD.

3.2. Recent Advances

Model parameters can be calibrated with data using statistically rigorous methods, such as data assimilation [Luo *et al.*, 2011]. Data assimilation integrates multiple sources of information from field observations to constrain parameters of C cycle models at ecosystem, regional, and global scales [Xu *et al.*, 2006; Zhou and Luo, 2008; Weng and Luo, 2011; Zhou *et al.*, 2012; Niu *et al.*, 2014]. At the global scale, for example, Hararuk *et al.* [2014, 2015] applied data assimilation methods to estimate the optimal parameters of the Community Land Model (CLM3.5) and two nonlinear microbial models with global databases of soil C content and microbial biomass. The optimized model explained 41% of the global variability in the observed SOC in comparison with 27% with the original parameters of CLM3.5 (Figure 4). The estimated SOC at the regional and global scales still have high mismatches at least partly due to the assumption that the training data represents soils in steady state [Carvalho *et al.*, 2008, 2010; Weng *et al.*, 2011] and partly due to model structural errors (e.g., no representation of peatlands or wetlands). When parameters that quantify nonsteady states of the C cycle are estimated, the mismatches significantly decrease for both vegetation and soil C pools [Zhou *et al.*, 2013].

The research community at large has recently made substantial progress in understanding parameters and their estimation with data assimilation in several aspects. First, it is known that soil C dynamics are quantitatively determined by relative changes in C inputs and residence times even if there are tens or hundreds of parameters in a typical soil C cycle model [Todd-Brown *et al.*, 2014; Luo *et al.*, 2015]. Thus, the behavior of

Table 1. Example Data Sets and Their Potential Uses for Constraining Soil Carbon Model Parameters

Category	Data Type	Temporal Frequency	Spatial Coverage	Model Constraints	
Observation	Litter production	Most annually	DPM ^a	C input to soil pools	
	Litter mass	Most monthly	DPM	Litter pools and transfer coefficients	
	Litter decomposition	Most monthly	DPM	Rates of litter decomposition	
	Root growth	Irregular	DPM	Rates of root growth	
	Root biomass	Irregular	DPM	Root pools and transfer coefficients	
	Soil carbon	Once every few years	DPM	Soil carbon pools and transfer coefficients	
	Soil microbial biomass	Irregular	DPM	Microbial pools and transfer coefficients	
	Soil respiration	Most monthly	DPM	Rates of soil carbon decomposition	
	Isotopes	¹³ C in soil and efflux	Irregular	DPM	Carbon transfer
		¹⁴ C in soil and efflux	Irregular	DPM	Residence times of soil carbon
Vertical profiles	Root, soil C, and ¹⁴ C	Irregular	DPM	Vertical transfer and properties of soil carbon	
Global change experiment (elevated CO ₂ , warming, precipitation, and nitrogen)	Litter, soil, and root dynamics	Weekly to yearly	Mostly in temperate and boreal regions and a few from tropical regions	Response functions of soil carbon processes to global change factors	

^aDPM = distributed point measurement.

modeled SOC dynamics at any point in time, space, and soil depth can be analyzed according to variations in C influx and residence time as they are related to environmental variables, litter quality, SOC properties, microbial attributes, and disturbances [Xia *et al.*, 2013]. When parameters related to C residence times were calibrated against a global soil C database with data assimilation, model agreement with observations was substantially improved (Figure 4) [Hararuk *et al.*, 2014].

Second, it is important to recognize that individual data sets contain information that may be useful in constraining only a subset of parameters [Weng and Luo, 2011]. For example, flux data can constrain flux-related parameters but contain little information with which to constrain pool-based parameters and vice versa [Keenan *et al.*, 2013; Du *et al.*, 2015]. Thus, data on soil respiration and net ecosystem exchange from flux measurements may not be useful to constrain soil C dynamics unless combined with pool-related data sets [Sierra, 2012; Keenan *et al.*, 2013]. A few C cycle data assimilation systems have been developed to assimilate data mostly from eddy flux, atmospheric, and satellite measurements [Rayner *et al.*, 2005; Kaminski *et al.*, 2013], whereas several studies use pool-based data sets only to constrain C transfer coefficients among pools [Hararuk *et al.*, 2014]. It will become more effective to improve global land models when flux- and pool-based global data sets are combined to constrain both C input and residence times [e.g., Smith *et al.*, 2013].

Third, most parameter estimation studies have shown that the number of parameters constrained by observational data sets is limited, typically only a few parameters from each data set [Wang *et al.*, 2001; Xu *et al.*, 2006]. When a few data sets are used to calibrate parameters of complex models, such as ESMs, the calibrated models with very different structures are able to fit equally well the existing observations but project largely different responses to future scenarios [He *et al.*, 2014]. This is an issue of equifinality with respect to the available information [Luo *et al.*, 2009; Sierra *et al.*, 2015b]. To avoid the equifinality issue, multiple sources of high-quality diverse observations are necessary (Table 1). Because soil C content is a complex product of litter production and SOC decomposition [Luo *et al.*, 2003; Zhou and Luo, 2008; Xia *et al.*, 2013], all parameters related to C inputs, allocation, and decomposition need to be constrained by multiple observations [Smith *et al.*, 2013]. In addition, observations that constrain parameters controlling soil physics (for example, thermal conductivity and water holding capacity) will indirectly help constrain SOC decomposition. Data products that are most effective for constraining parameters of soil C models include soil C pools, litter C pools, and root biomass [Jackson *et al.*, 2000] (Table 1). NPP and litterfall can be used to constrain C inputs to the soil system. Radiocarbon data are very informative as joint constraints for parameters of slow processes of soil C models [Gaudinski *et al.*, 2000; Trumbore, 2009; Baisden and Canessa, 2013; Koven *et al.*, 2013].

Fourth, the quality of the data sets matters for calibrated parameters. Higher-quality data sets used in data assimilation result in more representative parameters estimates. One of the most popular soil databases used to benchmark ESMs is the HWSO. The HWSO presents estimates of SOC content for 30 × 30 arc sec grid cells (~1 × 1 km), using class transfer functions that take into account regional differences in soil types [Omuto *et al.*, 2013]. It is important that these data sets be accompanied by soil reference data that encompass factors

important in soil formation. Recently, digital soil mapping techniques that include machine learning algorithms have been developed that draw on large soil profile databases and analyses of environmental covariates representing soil forming factors [Arrouays *et al.*, 2014; Hengl *et al.*, 2014], leading to a much improved accuracy (e.g., $r^2 = 0.61$ for SOC content in Africa at 250 m resolution) [Hengl *et al.*, 2015]. Alternately, data assimilation can directly use spatially distributed data points [Zhou *et al.*, 2013], avoiding the uncertainty introduced by harmonization as required for HWSD-type mapping approaches.

Fifth, parameters are not necessarily constants as often assumed in traditional simulation models. More and more syntheses have shown that model parameters vary between sites of measurements, often change with time, and may be better represented as probability distributions [Medlyn *et al.*, 1999; Lebauer *et al.*, 2013]. For example, fine root allocation tends to follow a log-normal distribution [Saugier *et al.*, 2001]. Moreover, parameters estimated from data sets derived from global change experiments vary with those global change factors [Luo *et al.*, 2003; Xu *et al.*, 2006; Shi *et al.*, 2015]. For example, comparison of posterior probability density functions from data assimilation studies showed that estimated C turnover in foliage and fine root pools was much higher at elevated than ambient CO₂ at the Duke forest free-air CO₂ enrichment site [Luo *et al.*, 2003; Xu *et al.*, 2006]. Those differences in estimated parameters propagate through models, leading to differences in simulated soil C pools.

In sum, technical development in the past decade or so makes it possible to use data assimilation to rigorously calibrate parameters of soil C models as standard practice. Systematic biases of soil C models can be effectively reduced when many sets of high-quality data from both flux- and pool-based measurements over long time and large space are used to calibrate two synthetic parameters: C influx and residence time.

3.3. Future Research Directions

Data assimilation is among the highest priorities for near-term research to reduce systematic biases (Figures 3 and 4) that pervade almost all soil C models. As a cornerstone for data assimilation, research efforts are needed to develop high-quality diverse data sets capability of effectively constraining parameters in soil C models (Table 1). Parameter calibration through data assimilation with common high-quality databases is expected to be especially effective when inputs, external forcings, and parameters can be constrained simultaneously.

Parameter calibration for ESMs with multiple data streams at the global scale has to tackle several challenges, such as the compatibility of multiple, heterogeneous data sets that constrain different model aspects across a diverse range of temporal and spatial scales, intractability of structural complexity of big models, equifinality of model structure selection and parameter estimation, and computational demand of global optimization with complex models. To effectively tackle the challenges of data assimilation with complex models, we have to develop more innovative approaches through multidisciplinary collaboration with mathematicians, statisticians, and computer scientists. One example is the development of the semianalytic spin-up method by Xia *et al.* [2012] that greatly reduces the computing time for global parameter estimation. The traceability framework developed by Xia *et al.* [2013] can help isolate various components and develop high-fidelity emulators of the complex C models for data assimilation [Hararuk *et al.*, 2014].

While data assimilation is expected to reduce model biases, parameterization also needs to represent uncertainty arising from subgrid heterogeneity in models. Many of the dominant processes incorporated in ESMs are characterized by very high subgrid-scale variability in space and time. This variability is always much higher than any ESMs can resolve. These highly variable processes must be parameterized using the variables represented in ESMs, together with ancillary data on boundary conditions such as soil and vegetation properties, sometimes through stochastic upscaling models. This is an important issue that the soil modeling community needs to work on.

4. External Variables to Soil C Cycling

4.1. Representation of External Variables in ESMs

Soil C dynamics are subject to changes in external variables. The external variables that influence soil C dynamics include climate conditions (e.g., temperature and precipitation), edaphic conditions (e.g., soil texture, mineralogy, and soil depth), soil thermal conditions, hydrological conditions (soil moisture, water table in wetlands, frozen versus liquid versus vapor state in seasonally and perennially frozen soils), oxygen and nutrient levels (e.g., redox state and N and phosphorus availability), and vegetation characteristics

(e.g., rooting depth and litter types) (Figure 2). Those external variables regulate various aspects of soil C dynamics, cause spatial or temporal variability, and need to be appropriately represented in models.

External variables are represented in ESMs at least in three different ways: (1) parameters that do not evolve over time but directly control the system dynamics, (2) boundary conditions that evolve over time as forcing but are not part of the system being modeled, and (3) prognostic variables that are allowed to evolve over time as part of the modeled system. Traditionally, prognostic soil environmental variables are considered to be part of the modeled system but treated as external variables to SOC processes in this paper to facilitate analysis of SOC modeling results. Indeed, isolating endogenous processes from exogenous variables helps not only diagnose causes of model uncertainty but also understand fundamental properties of the terrestrial C cycle [Luo *et al.*, 2015].

Clay content and mineralogy, which influence soil C stabilization and decomposition, are parameters that are usually assigned from observations. Regional and global data sets can constrain information on physical properties (soil texture and mineralogy) [Gulde *et al.*, 2008; Feng *et al.*, 2009; Journet *et al.*, 2014]. Soil depth, which is required for multisoil layer models, is a forcing parameter that is directly derived from measurement [Journet *et al.*, 2014]. Vegetation and land use for stand-alone, site-specific models are usually set as parameters to control the system dynamics.

Many forcing variables are represented as boundary conditions to land components of ESMs. For example, prescribed atmospheric CO₂ concentration, temperature, and precipitation that are used to drive global land models are external forcing variables. A global N deposition model product from 1860 to 2100 offers dynamic forcing variables for coupled C-N models [Lamarque *et al.*, 2010]. Simulating horizontal and vertical movements caused by soil erosion at regional and global scales will require a global net soil redistribution map [Chappell *et al.*, 2014] to provide boundary conditions for global land models.

Many of those climate, edaphic, hydrological, and vegetation variables are exogenous to soil C processes but simulated in coupled models. For example, a coupled C-N model simulates not only N influences on C processes such as photosynthesis, plant C allocation, and litter decomposition but also N dynamics as influenced by C cycle processes, such as plant N uptake, N fixation, microbial N immobilization, and denitrification. To realistically represent N influences on the C cycle, the models have to accurately simulate both N processes and responses of C cycling to N. Temperature, precipitation, and atmospheric CO₂ concentration are boundary conditions in global land models but evolve over time as prognostic variables in the coupled climate-C models. As the domain of the modeling problem expands from just the soil to entire ecosystems, to the land surface, and to the Earth system with coupled land-atmosphere-ocean dynamics, more and more exogenous variables to the SOC dynamics are included in models as prognostic variables to reflect real interactions among system components.

Overall, external variables are represented in different ESMs in very different ways. Tracing the way of each exogenous variable being represented in each ESM, either as parameter, boundary condition, or prognostic variable, is critical to understanding the nature of model uncertainty and to disentangling the sources of uncertainty.

4.2. Recent Advances

The relative contributions of external variables to uncertainty in land C modeling have been recently quantified. Ahlström *et al.* [2013] investigated the potential sensitivity of the global terrestrial ecosystem C balance to different climate forcing generated by four general circulation models (GCMs) under three different CO₂ concentration scenarios. Variations in climate variables (e.g., temperature, precipitation, and shortwave radiation) generated by different GCM explained the majority of the uncertainty in the future evolution of global terrestrial ecosystem C. Studies with Dynamic Global Vegetation Models suggest that the uncertainty in total terrestrial C storage caused by differences in climate variables among GCMs is comparable to the uncertainty caused by the responses of the C cycle components [Berthelot *et al.*, 2005; Schaphoff *et al.*, 2006; Scholze *et al.*, 2006; Ahlström *et al.*, 2012, 2013].

To disentangle complex representations of external variables in influencing simulated C dynamics in ESMs, Xia *et al.* [2013] developed a traceability framework to decompose the complex terrestrial C cycle into a few traceable components (Figure 5). The traceability analysis helps identify sources of uncertainty in modeled steady state ecosystem carbon storage due to (1) C input as affected by phenology, physiology, and C use efficiency [Xia *et al.*, 2015], (2) edaphic and vegetation characteristics as related to baseline C residence time, (3) climate scalars, and (4) environmental variables among models. The traceability framework has been applied to assess influences of external variables being represented as parameters, boundary conditions, and diagnostic variables in models.

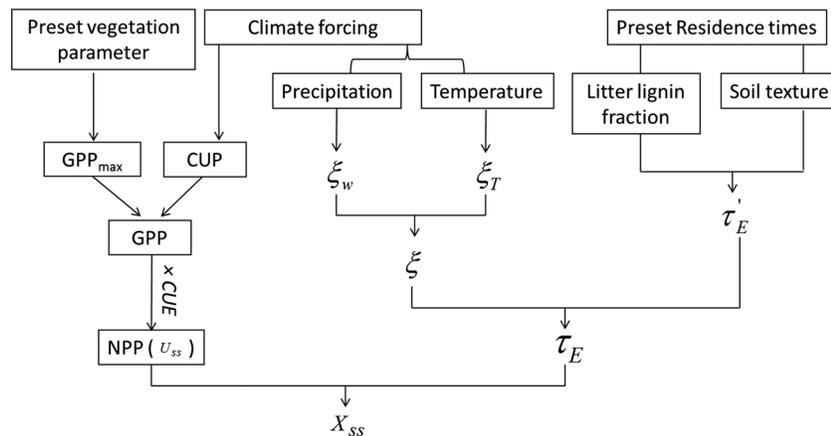


Figure 5. Schematic diagram of the traceability framework. The framework traces modeled ecosystem C storage capacity (X_{ss}) to a product of net primary productivity (NPP) and ecosystem residence time (τ_E). The latter τ_E can be further traced to (i) baseline C residence times (τ'_E), which are usually preset in a model according to vegetation characteristics and soil types, (ii) environmental scalars (ξ), including temperature and water scalars, and (iii) environmental forcing. NPP can be traced to C use efficiency (CUE), C uptake period, and the seasonal maximum of gross primary productivity (GPP) (adopted from Xia *et al.* [2013, 2015]).

Differences in values of external variables set as parameters can be a major source of uncertainty in modeling the C cycle. Driven with similar climate data, the CLM3.5 model simulated $\sim 31\%$ larger C storage capacity than the Australian Community Atmosphere Biosphere Land Exchange (CABLE) model [Rafique *et al.*, 2014]. According to the traceability analysis, the projected difference in C storage between the two models results from differences in either NPP or residence time or both. CLM3.5 simulated 37% higher NPP than CABLE due to higher rates of carboxylation in CLM3.5. On the other hand, residence time of ecosystem C was 11 years longer in CABLE than CLM3.5. The difference in residence time is mainly caused by longer baseline residence time of woody biomass (23 years in CABLE versus 14 years in CLM3.5) and higher proportion of NPP allocated to woody biomass (23% in CABLE versus 16% in CLM3.5).

A recent application of the traceability framework partitioned climate-induced soil C modeling uncertainties into soil decomposition rates, NPP, and vegetation turnover [Ahlström *et al.*, 2015]. A global dynamic vegetation-ecosystem model, LPJ-GUESS, was used with a detailed individual and patch-based representation of vegetation structure, demography, and resource competition [Smith *et al.*, 2001]. Changes in climate variables and CO_2 concentrations from 13 different climate or Earth system model simulations from CMIP5 under RCP8.5 radiative forcing were used as boundary conditions for simulations with LPJ-GUESS. The 13 climate change scenarios caused uncertainties in modeled global C storage through their influences on NPP, vegetation dynamics and turnover, and soil decomposition rates. To quantify relative contributions of climate-induced changes in those processes to modeled carbon storage uncertainty, an emulator was developed to describe the carbon flows and pools exactly as in simulations with LPJ-GUESS according to the traceability framework. Traceability analysis indicated that NPP, vegetation turnover, and soil decomposition rates explained 49%, 17%, and 33%, respectively, of uncertainties in modeled C storage.

When external variables to the soil C cycle evolve in the model as diagnostic variables, their impacts on C cycles are much more difficult to disentangle. Xia *et al.* [2013] applied the traceability framework (Figure 5) to analyze impacts of N feedback on the C cycle. Incorporation of N processes into the CABLE model decreased C storage in all biomes via decreased NPP or decreased residence times or both. The decreases in residence times resulted from N-induced changes in C allocation among plant pools and changes in transfers from plant to litter and soil pools.

4.3. Future Research Directions

To understand model-model differences, it is essential to trace what and how external variables are represented in models. Complex impacts of external variables on the modeled C cycle are due to at least three reasons: (1) different sets of external variables being incorporated into individual ESMs, (2) the same set of external variables being represented in different ways either as parameters, boundary conditions, or

prognostic variables in ESMs, and (3) different response functions to link external variables to C cycle processes. The diverse ways of representing of external variables substantially contribute to differences in modeled soil C dynamics for any model intercomparison projects.

It is also essential to expand the traceability framework to analyze transient dynamics of the C cycle under climate change. The traceability framework developed by Xia *et al.* [2013] can be only applied to steady state carbon cycle analysis. Carbon cycle modeling is primarily to study responses of ecosystems to climate change. Thus it is critical to trace how different sets of external variables and their diverse representations affect the transient dynamics of C cycle under climate change.

5. Concluding Remarks

Recent analyses of CMIP5 results have revealed enormous differences in SOC projections among ESMs. This paper attempts to identify causes of the model-model differences from the three components of modeling: structure, parameterization, and external forcing. The current generation of ESMs shares a similar model formulation to represent soil C processes (i.e., the donor pool-dominant and first-order C transfers among multiple pools). This formulation is consistent with fundamental properties of the terrestrial C cycle and captures the macroscopic patterns observed from litter decomposition and soil incubation studies. Synthesis of vast available data sets needs be done to examine whether the model formulation or merely its parameters should vary with microbial processes, soil depth, nutrient availability, and disturbances among many other processes and factors.

Incomplete use of observations in model parameterization is a major cause of model-model differences. Of the two synthetic parameters that determine soil C storage, contemporary C influx differed by threefold and residence times differed by fourfold among CMIP5 models. It is conceivable that optimized calibration of model parameters with common databases through data assimilation could substantially reduce systematic biases among models, especially if inputs and external forcings are also simultaneously constrained by common protocols. To achieve this, we need to improve the availability and use of global databases, develop C cycle data assimilation systems that can effectively assimilate both flux- and pool-based data sets into global C cycle models, and understand subgrid variability of model parameters.

Individual ESMs not only include different sets of external variables that are exogenous to soil C cycles but also represent them at least in three different ways (e.g., parameter, boundary condition, and prognostic variables) and using different response functions. The diverse representations of external variables contribute markedly to the differences in modeled soil C stock and dynamics. In the next few years, we should expand the list of output variables from ESMs so as to permit more comprehensive model evaluations, such as traceability analysis, to attribute the model differences to various causes. The long-term goal is to develop an evaluation-improvement system to allow fast feedback between performance evaluation and model development toward realistic representations of ecosystem C cycle responses to climate change.

Acknowledgments

The paper stemmed from a workshop "Representing soil carbon dynamics in global land models to improve future IPCC assessments" held at Breckenridge, CO, USA on 12–14 June 2014. The workshop was financially supported by the United States National Science Foundation Research Coordination (RCN) grant DEB 0840964 and Department of Energy DE SC0008270. Y.L. was financially supported by U.S. Department of Energy grants DE-SC0006982, DE-SC0008270, DE-SC0014062, DE-SC0004601, and DE-SC0010715 and U.S. National Science Foundation (NSF) grants DBI 0850290, EPS 0919466, DEB 0840964, and EF 1137293; N.C. by Project UID/AMB/04085/2013; CDJ by the Joint DECC/Defra Met Office Hadley Centre Climate Program (GA01101). Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government. Please contact the corresponding author at yluo@ou.edu for details of the data used in this work.

References

- Adair, E. C., W. J. Parton, S. J. Del Grosso, W. L. Silver, M. E. Harmon, S. A. Hall, I. C. Burke, and S. C. Hart (2008), Simple three-pool model accurately describes patterns of long-term litter decomposition in diverse climates, *Global Change Biol.*, *14*(11), 2636–2660.
- Ågren, G. I., and E. Bosatta (1996), Quality: A bridge between theory and experiment in soil organic matter studies, *Oikos*, *76*(3), 522–528.
- Ahlström, A., G. Schurgers, A. Arneth, and B. Smith (2012), Robustness and uncertainty in terrestrial ecosystem carbon response to CMIP5 climate change projections, *Environ. Res. Lett.*, *7*(4), 044008, doi:10.1088/1748-9326/7/4/044008.
- Ahlström, A., B. Smith, J. Lindstrom, M. Rummukainen, and C. B. Uvo (2013), GCM characteristics explain the majority of uncertainty in projected 21st century terrestrial ecosystem carbon balance, *Biogeosciences*, *10*(3), 1517–1528.
- Ahlström, A., J. Y. Xia, A. Arneth, Y. Q. Luo, and B. Smith (2015), Importance of vegetation dynamics for future terrestrial carbon cycling, *Environ. Res. Lett.*, *10*, 054019, doi:10.1088/1748-9326/10/5/054019/054019.
- Allison, S. D., M. D. Wallenstein, and M. A. Bradford (2010), Soil-carbon response to warming dependent on microbial physiology, *Nat. Geosci.*, *3*(5), 336–340.
- Arora, V. K., et al. (2013), Carbon-concentration and carbon-climate feedbacks in CMIP5 earth system models, *J. Clim.*, *26*(15), 5289–5314.
- Arrauy, D., et al. (2014), GlobalSoilMap: Toward a fine-resolution global grid of soil properties, *Adv. Agron.*, *125*, 93–134.
- Baisden, W. T., and S. Canessa (2013), Using 50 years of soil radiocarbon data to identify optimal approaches for estimating soil carbon residence times, *Nucl. Instrum. Methods Phys. Res., Sect. B*, *294*, 588–592.
- Berthelot, M., P. Friedlingstein, P. Ciais, J.-L. Dufresne, and P. Monfray (2005), How uncertainties in future climate change predictions translate into future terrestrial carbon fluxes, *Global Change Biol.*, *11*(6), 959–970.
- Bolker, B. M., S. W. Pacala, and W. J. Parton (1998), Linear analysis of soil decomposition: Insights from the century model, *Ecol. Appl.*, *8*(2), 425–439.

- Bradford, M. A., B. W. Watts, and C. A. Davies (2010), Thermal adaptation of heterotrophic soil respiration in laboratory microcosms, *Global Change Biol.*, *16*(5), 1576–1588.
- Brovkin, V., et al. (2013), Effect of anthropogenic land-use and land-cover changes on climate and land carbon storage in CMIP5 projections for the twenty-first century, *J. Clim.*, *26*(18), 6859–6881.
- Burke, I. C., C. M. Yonker, W. J. Parton, C. V. Cole, K. Flach, and D. S. Schimel (1989), Texture, climate, and cultivation effects on soil organic-matter content in U.S. grassland soils, *Soil Sci. Soc. Am. J.*, *53*(3), 800–805.
- Carvalhais, N., et al. (2008), Implications of the carbon cycle steady state assumption for biogeochemical modeling performance and inverse parameter retrieval, *Global Biogeochem. Cycles*, *22*, GB2007, doi:10.1029/2007GB003033.
- Carvalhais, N., M. Reichstein, P. Ciais, G. J. Collatz, M. D. Mahecha, L. Montagnani, D. Papale, S. Rambal, and J. Seixas (2010), Identification of vegetation and soil carbon pools out of equilibrium in a process model via eddy covariance and biometric constraints, *Global Change Biol.*, *16*(10), 2813–2829.
- Carvalhais, N., et al. (2014), Global covariation of carbon turnover times with climate in terrestrial ecosystems, *Nature*, *514*(7521), 213–217.
- Chappell, A., N. P. Webb, R. A. V. Rossel, and E. Bui (2014), Australian net (1950s–1990) soil organic carbon erosion: Implications for CO₂ emission and land-atmosphere modelling, *Biogeosciences*, *11*(18), 5235–5244.
- Ciais, P., et al. (2013), Carbon and other biogeochemical cycles, in *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by T. F. Stocker et al., Cambridge Univ. Press, Cambridge, U. K., and New York.
- Cramer, W., et al. (2001), Global response of terrestrial ecosystem structure and function to CO₂ and climate change: Results from six dynamic global vegetation models, *Global Change Biol.*, *7*(4), 357–373.
- Davidson, E. A., S. Samanta, S. S. Caramori, and K. Savage (2012), The Dual Arrhenius and Michaelis-Menten kinetics model for decomposition of soil organic matter at hourly to seasonal time scales, *Global Change Biol.*, *18*(1), 371–384.
- Davidson, E. A., K. E. Savage, and A. C. Finzi (2014), A big-microsite framework for soil carbon modeling, *Global Change Biol.*, *20*(12), 3610–3620.
- Doetterl, S., et al. (2015), Soil carbon storage controlled by interactions between geochemistry and climate, *Nat. Geosci.*, *8*, 780–783.
- Du, Z., Y. Nie, Y. He, G. Yu, H. Wang, and X. Zhou (2015), Complementarity of flux-and biometric-based data to constrain parameters in a terrestrial carbon model, *Tellus B*, *67*, 24102, doi:10.3402/tellusb.v67.24102.
- Dümig, A., R. Smittenberg, and I. Kögel-Knabner (2011), Concurrent evolution of organic and mineral components during initial soil development after retreat of the Damma glacier, Switzerland, *Geoderma*, *163*(1), 83–94.
- Edburg, S. L., J. A. Hicke, D. M. Lawrence, and P. E. Thornton (2011), Simulating coupled carbon and nitrogen dynamics following mountain pine beetle outbreaks in the western United States, *J. Geophys. Res.*, *116*, G04033, doi:10.1029/2011JG001786.
- Egli, M., M. Nater, A. Mirabella, S. Raimondi, M. Plötze, and L. Alioth (2008), Clay minerals, oxyhydroxide formation, element leaching and humus development in volcanic soils, *Geoderma*, *143*(1), 101–114.
- Exbrayat, J. F., A. J. Pitman, Q. Zhang, G. Abramowitz, and Y. P. Wang (2013), Examining soil carbon uncertainty in a global model: Response of microbial decomposition to temperature, moisture and nutrient limitation, *Biogeosciences*, *10*(11), 7095–7108.
- Falloon, P., C. D. Jones, M. Ades, and K. Paul (2011), Direct soil moisture controls of future global soil carbon changes: An important source of uncertainty, *Global Biogeochem. Cycles*, *25*, GB3010, doi:10.1029/2010GB003938.
- Feng, G., B. Sharratt, J. Vaughan, and B. Lamb (2009), A multiscale database of soil properties for regional environmental quality modeling in the western United States, *J. Soil Water Conserv.*, *64*(6), 363–373.
- Food and Agriculture Organization, I. ISRIC, ISSCAS, and JRC (2012), *Harmonized World Soil Database (Version 1.2)*, FAO and IIASA, Rome, Italy and Laxenburg, Austria.
- Friedlingstein, P., et al. (2006), Climate-carbon cycle feedback analysis: Results from the (CMIP)-M-4 model intercomparison, *J. Clim.*, *19*(14), 3337–3353.
- Fujita, Y., J. M. Witte, and P. M. van Bodegom (2014), Incorporating microbial ecology concepts into global soil mineralization models to improve predictions of carbon and nitrogen fluxes, *Global Biogeochem. Cycles*, *28*, 223–238, doi:10.1002/2013GB004595.
- Gaudinski, J. B., S. E. Trumbore, E. A. Davidson, and S. H. Zheng (2000), Soil carbon cycling in a temperate forest: Radiocarbon-based estimates of residence times, sequestration rates and partitioning of fluxes, *Biogeochemistry*, *51*(1), 33–69.
- Goetz, S. J., et al. (2012), Observations and assessment of forest carbon dynamics following disturbance in North America, *J. Geophys. Res.*, *117*, G02022, doi:10.1029/2011JG001733.
- Gregory, J. M., C. D. Jones, P. Cadule, and P. Friedlingstein (2009), Quantifying carbon cycle feedbacks, *J. Clim.*, *22*(19), 5232–5250.
- Grosse, G., et al. (2011), Vulnerability of high-latitude soil organic carbon in North America to disturbance, *J. Geophys. Res.*, *116*, G00K06, doi:10.1029/2010JG001507.
- Gulde, S., H. Chung, W. Amelung, C. Chang, and J. Six (2008), Soil carbon saturation controls labile and stable carbon pool dynamics, *Soil Sci. Soc. Am. J.*, *72*(3), 605–612.
- Hararuk, O., J. Y. Xia, and Y. Q. Luo (2014), Evaluation and improvement of a global land model against soil carbon data using a Bayesian Markov chain Monte Carlo method, *J. Geophys. Res. Biogeosci.*, *119*, 403–417, doi:10.1002/2013JG002535.
- Hararuk, O., M. J. Smith, and Y. Q. Luo (2015), Microbial models with data-driven parameters predict stronger soil carbon responses to climate change, *Global Change Biol.*, *21*, 2439–2453.
- Haxeltine, A., and I. C. Prentice (1996), A general model for the light-use efficiency of primary production, *Funct. Ecol.*, *10*(5), 551–561.
- He, Y., J. Yang, Q. Zhuang, A. D. McGuire, Q. Zhu, Y. Liu, and R. O. Teskey (2014), Uncertainty in the fate of soil organic carbon: A comparison of three conceptually different decomposition models at a larch plantation, *J. Geophys. Res. Biogeosci.*, *119*, 1892–1905, doi:10.1002/2014JG002701.
- Hengl, T., et al. (2014), SoilGrids1km—Global soil information based on automated mapping, *Plos One*, *9*(8), doi:10.1371/journal.pone.0125814.
- Hengl, T., G. B. Heuvelink, B. Kempen, J. G. Leenaars, M. G. Walsh, K. D. Shepherd, A. Sila, R. A. MacMillan, J. M. de Jesus, and L. Tamene (2015), Mapping soil properties of Africa at 250 m resolution: Random forests significantly improve current predictions, *Plos One*, *10*(6), doi:10.1371/journal.pone.0125814.
- Henin, S., and M. Dupuis (1945), Essai de bilan de la matière organique du sol, *Ann. Agron.*, *11*, 17–29.
- Hicke, J. A., et al. (2012), Effects of biotic disturbances on forest carbon cycling in the United States and Canada, *Global Change Biol.*, *18*(1), 7–34.
- Hobbie, S. E., J. P. Schimel, S. E. Trumbore, and J. R. Randerson (2000), Controls over carbon storage and turnover in high-latitude soils, *Global Change Biol.*, *6*, 196–210.
- Hurtt, G. C., et al. (2011), Harmonization of land-use scenarios for the period 1500–2100: 600 years of global gridded annual land-use transitions, wood harvest, and resulting secondary lands, *Clim. Change*, *109*(1–2), 117–161.

- Jackson, R. B., et al. (2000), Belowground consequences of vegetation change and their treatment in models, *Ecol. Appl.*, *10*(2), 470–483.
- Jenkinson, D. S. (1990), The turnover of organic-carbon and nitrogen in soil, *Philos. Trans. R. Soc., B*, *329*(1255), 361–368.
- Jenkinson, D. S., P. B. S. Hart, J. H. Rayner, and L. C. Parry (1987), Modelling the turnover of organic matter in long-term experiments at Rothamsted, *INTECOL Bull.*, *15*, 1–8.
- Jenny, H. (1961), *E. W. Hilgard and the Birth of Modern Soil Science*, Collana Della Rivista Agrochimica, Pisa.
- Jobbágy, E. G., and R. B. Jackson (2000), The vertical distribution of soil organic carbon and its relation to climate and vegetation, *Ecol. Appl.*, *10*, 423–436.
- Jones, C., and P. Falloon (2009), Sources of uncertainty in global modelling of future soil organic carbon storage, in *Uncertainties in Environmental Modelling and Consequences for Policy Making*, edited by P. C. Bayeye, M. Laba, and J. Mysiak, pp. 283–315, Springer, Netherlands.
- Jones, C., C. McConnell, K. Coleman, P. Cox, P. Falloon, D. Jenkinson, and D. Powlson (2005), Global climate change and soil carbon stocks: predictions from two contrasting models for the turnover of organic carbon in soil, *Global Change Biol.*, *11*(1), 154–166.
- Jones, C., et al. (2013), Twenty-first-century compatible CO₂ emissions and airborne fraction simulated by CMIP5 Earth system models under four representative concentration pathways, *J. Clim.*, *26*(13), 4398–4413.
- Jones, C. D., P. Cox, and C. Huntingford (2003), Uncertainty in climate-carbon-cycle projections associated with the sensitivity of soil respiration to temperature, *Tellus B*, *55*(2), 642–648.
- Jones, C. D., P. M. Cox, and C. Huntingford (2006), Climate-carbon cycle feedbacks under stabilization: Uncertainty and observational constraints, *Tellus B*, *58*(5), 603–613.
- Journet, E., Y. Balkanski, and S. P. Harrison (2014), A new data set of soil mineralogy for dust-cycle modeling, *Atmos. Chem. Phys.*, *14*(8), 3801–3816.
- Kaminski, T., et al. (2013), The BETHY/JSBACH Carbon Cycle Data Assimilation System: Experiences and challenges, *J. Geophys. Res. Biogeosci.*, *118*, 1414–1426, doi:10.1002/jgrg.20118.
- Keenan, T. F., E. A. Davidson, J. W. Munger, and A. D. Richardson (2013), Rate my data: Quantifying the value of ecological data for the development of models of the terrestrial carbon cycle, *Ecol. Appl.*, *23*(1), 273–286.
- Kleber, M., P. Sollins, and R. Sutton (2007), A conceptual model of organo-mineral interactions in soils: Self-assembly of organic molecular fragments into zonal structures on mineral surfaces, *Biogeochemistry*, *85*(1), 9–24.
- Kloster, S., N. M. Mahowald, J. T. Randerson, P. E. Thornton, F. M. Hoffman, S. Levis, P. J. Lawrence, J. J. Feddesma, K. W. Oleson, and D. M. Lawrence (2010), Fire dynamics during the 20th century simulated by the Community Land Model, *Biogeosciences*, *7*(6), 1877–1902.
- 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.
- Kuzyakov, Y. (2010), Priming effects: Interactions between living and dead organic matter, *Soil Biol. Biochem.*, *42*(9), 1363–1371.
- Lamarque, J. F., et al. (2010), Historical (1850–2000) gridded anthropogenic and biomass burning emissions of reactive gases and aerosols: Methodology and application, *Atmos. Chem. Phys.*, *10*(15), 7017–7039.
- Lawrence, C. R., J. C. Neff, and J. P. Schimel (2009), Does adding microbial mechanisms of decomposition improve soil organic matter models? A comparison of four models using data from a pulsed rewetting experiment, *Soil Biol. Biochem.*, *41*(9), 1923–1934.
- Lebauer, D. S., D. Wang, K. T. Richter, C. C. Davidson, and M. C. Dietze (2013), Facilitating feedbacks between field measurements and ecosystem models, *Ecol. Monogr.*, *83*(2), 133–154.
- Li, D. J., C. Schaedel, M. L. Haddix, E. A. Paul, R. Conant, J. W. Li, J. Z. Zhou, and Y. Q. Luo (2013), Differential responses of soil organic carbon fractions to warming: Results from an analysis with data assimilation, *Soil Biol. Biochem.*, *67*, 24–30.
- Li, J. W., Y. Q. Luo, S. Natali, E. A. G. Schuur, J. Y. Xia, E. Kowalczuk, and Y. P. Wang (2014), Modeling permafrost thaw and ecosystem carbon cycle under annual and seasonal warming at an Arctic tundra site in Alaska, *J. Geophys. Res. Biogeosci.*, *119*, 1129–1146, doi:10.1002/2013JG002569.
- Liang, J. Y., D. J. Li, Z. Shi, J. M. Tiedje, J. Z. Zhou, E. A. G. Schuur, K. T. Konstantinidis, and Y. Q. Luo (2015), Methods for estimating temperature sensitivity of soil organic matter based on incubation data: A comparative evaluation, *Soil Biol. Biochem.*, *80*, 127–135.
- Limpens, J., F. Berendse, C. Blodau, J. G. Canadell, C. Freeman, J. Holden, N. Roulet, H. Rydin, and G. Schaepman-Strub (2008), Peatlands and the carbon cycle: From local processes to global implications—A synthesis, *Biogeosciences*, *5*(5), 1475–1491.
- Liu, X. C., S. Q. Wan, B. Su, D. F. Hui, and Y. Q. Luo (2002), Response of soil CO₂ efflux to water manipulation in a tallgrass prairie ecosystem, *Plant Soil*, *240*, 213–223.
- Lloyd, J., and J. A. Taylor (1994), On the temperature-dependence of soil respiration, *Funct. Ecol.*, *8*(3), 315–323.
- Luo, Y., E. Weng, X. Wu, C. Gao, X. Zhou, and L. Zhang (2009), Parameter identifiability, constraint, and equifinality in data assimilation with ecosystem models, *Ecol. Appl.*, *19*(3), 571–574.
- Luo, Y. Q., and E. S. Weng (2011), Dynamic disequilibrium of the terrestrial carbon cycle under global change, *Trends Ecol. Evol.*, *26*(2), 96–104.
- Luo, Y. Q., S. Q. Wan, D. F. Hui, and L. L. Wallace (2001), Acclimatization of soil respiration to warming in a tall grass prairie, *Nature*, *413*(6856), 622–625.
- Luo, Y. Q., L. W. White, J. G. Canadell, E. H. DeLucia, D. S. Ellsworth, A. C. Finzi, J. Lichten, and W. H. Schlesinger (2003), Sustainability of terrestrial carbon sequestration: A case study in Duke Forest with inversion approach, *Global Biogeochem. Cycles*, *17*(1), 1021, doi:10.1029/2002GB001923.
- Luo, Y. Q., K. Ogle, C. Tucker, S. F. Fei, C. Gao, S. LaDeau, J. S. Clark, and D. S. Schimel (2011), Ecological forecasting and data assimilation in a data-rich era, *Ecol. Appl.*, *21*(5), 1429–1442.
- Luo, Y. Q., T. F. Keenan, and M. Smith (2015), Predictability of the terrestrial carbon cycle, *Global Change Biol.*, *21*(5), 1737–1751.
- Luo, Z. K., E. L. Wang, I. R. P. Fillery, L. M. Macdonald, N. Huth, and J. Baldock (2014), Modelling soil carbon and nitrogen dynamics using measurable and conceptual soil organic matter pools in APSIM, *Agric. Ecosyst. Environ.*, *186*, 94–104.
- Mathieu, J. A., C. Hatté, J. Balesdent, and É. Parent (2015), Deep soil carbon dynamics are driven more by soil type than by climate: A worldwide meta-analysis of radiocarbon profiles, *Global Change Biol.*, *21*, 4278–4292.
- McGuire, A. D., J. M. Melillo, D. W. Kicklighter, Y. D. Pan, X. M. Xiao, J. Helfrich, B. Moore, C. J. Vorosmarty, and A. L. Schloss (1997), Equilibrium responses of global net primary production and carbon storage to doubled atmospheric carbon dioxide: Sensitivity to changes in vegetation nitrogen concentration, *Global Biogeochem. Cycles*, *11*(2), 173–189.
- Medlyn, B. E., et al. (1999), Effects of elevated [CO₂] on photosynthesis in European forest species: A meta-analysis of model parameters, *Plant. Cell. Environ.*, *22*(12), 1475–1495.
- Melillo, J. M., A. D. McGuire, D. W. Kicklighter, B. Moore, C. J. Vorosmarty, and A. L. Schloss (1993), Global climate-change and terrestrial net primary production, *Nature*, *363*(6426), 234–240.
- Moorhead, D. L., and R. L. Sinsabaugh (2006), A theoretical model of litter decay and microbial interaction, *Ecol. Monogr.*, *76*(2), 151–174.

- Moyano, F. E., et al. (2012), The moisture response of soil heterotrophic respiration: Interaction with soil properties, *Biogeosciences*, 9(3), 1173–1182.
- Moyano, F. E., S. Manzoni, and C. Chenu (2013), Responses of soil heterotrophic respiration to moisture availability: An exploration of processes and models, *Soil Biol. Biochem.*, 59, 72–85.
- Niu, S. L., Y. Q. Luo, M. C. Dietze, T. F. Keenan, Z. Shi, J. W. Li, and F. S. Chapin (2014), The role of data assimilation in predictive ecology, *Ecosphere*, 5(5), 65, doi:10.1890/ES13-00273.1.
- Omuto, C., F. Nachtergaele, and R. V. Rojas (2013), *State of the Art Report on Global and Regional Soil Information: Where are We? Where to Go?*, Food and Agriculture Organization of the United Nations, Rome.
- Parton, W. J., D. S. Schimel, C. V. Cole, and D. S. Ojima (1987), Analysis of factors controlling soil organic-matter levels in great-plains grasslands, *Soil Sci. Soc. Am. J.*, 51(5), 1173–1179.
- Parton, W. J., J. W. B. Stewart, and C. V. Cole (1988), Dynamics of C, N, P and S in Grassland Soils—A Model, *Biogeochemistry*, 5(1), 109–131.
- Parton, W. J., et al. (1993), Observations and modeling of biomass and soil organic-matter dynamics for the grassland biome worldwide, *Global Biogeochem. Cycles*, 7(4), 785–809.
- Prentice, I. C., D. I. Kelley, P. N. Foster, P. Friedlingstein, S. P. Harrison, and P. J. Bartlein (2011), Modeling fire and the terrestrial carbon balance, *Global Biogeochem. Cycles*, 25, GB3005, doi:10.1029/2010GB003906.
- Rafique, R., J. Y. Xia, O. Hararuk, and Y. Q. Luo (2014), Structural analysis of three global land models on carbon cycle simulations using a traceability framework, *Biogeosci. Discuss.*, 11, 9979–10,014.
- Raupach, M. R., and H. Lu (2004), Representation of land-surface processes in aeolian transport models, *Environ. Modell. Software*, 19, 93–112.
- Rayner, P. J., M. Scholze, W. Knorr, T. Kaminski, R. Giering, and H. Widmann (2005), Two decades of terrestrial carbon fluxes from a carbon cycle data assimilation system (CCDAS), *Global Biogeochem. Cycles*, 19, GB2026, doi:10.1029/2004GB002254.
- Rosenbloom, N. A., J. W. Harden, J. C. Neff, and D. S. Schimel (2006), Geomorphic control of landscape carbon accumulation, *J. Geophys. Res.*, 111, G01004, doi:10.1029/2005JG000077.
- Saugier, B., J. Roy, and H. A. Mooney (2001), Estimations of global terrestrial productivity: Converging toward a single number, in *Terrestrial Global Productivity*, edited by J. Roy, B. Saugier, and H. A. Mooney, pp. 543–557, Academic Press, San Diego, Calif.
- Schädel, C., Y. Q. Luo, R. D. Evans, S. F. Fei, and S. M. Schaeffer (2013), Separating soil CO₂ efflux into C-pool-specific decay rates via inverse analysis of soil incubation data, *Oecologia*, 171(3), 721–732.
- Schädel, C., E. A. G. Schuur, R. Bracho, B. Elberling, C. Knoblauch, H. Lee, Y. Q. Luo, G. R. Shaver, and M. R. Turetsky (2014), Circumpolar assessment of permafrost C quality and its vulnerability over time using long-term incubation data, *Global Change Biol.*, 20(2), 641–652.
- Schaphoff, S., W. Lucht, D. Gerten, S. Sitch, W. Cramer, and I. Prentice (2006), Terrestrial biosphere carbon storage under alternative climate projections, *Clim. Change*, 74(1), 97–122.
- Schimel, D. S., B. H. Braswell, R. McKeown, D. S. Ojima, W. J. Parton, and W. Pulliam (1996), Climate and nitrogen controls on the geography and timescales of terrestrial biogeochemical cycling, *Global Biogeochem. Cycles*, 10(4), 677–692.
- Schimel, J. (2001), Biogeochemical models: implicit versus explicit microbiology, in *Global Biogeochemical Cycles in the Climate System*, edited by E. D. Schulze et al., pp. 177–183, Academic Press, San Diego, Calif.
- Schimel, J. P., and S. M. Schaeffer (2012), Microbial control over carbon cycling in soil, *Front. Microbiol.*, 3, doi:10.3389/fmicb.2012.00348.
- Schimel, J. P., and M. N. Weintraub (2003), The implications of exoenzyme activity on microbial carbon and nitrogen limitation in soil: A theoretical model, *Soil Biol. Biochem.*, 35(4), 549–563.
- Schmidt, M. W. I., et al. (2011), Persistence of soil organic matter as an ecosystem property, *Nature*, 478(7367), 49–56.
- Scholze, M., W. Knorr, N. W. Arnell, and I. C. Prentice (2006), A climate-change risk analysis for world ecosystems, *Proc. Natl. Acad. Sci. U.S.A.*, 103(35), 13,116–13,120.
- Shi, Z., X. Xu, O. Hararuk, L. Jiang, J. Y. Xia, J. Y. Liang, D. Li, and Y. Q. Luo (2015), Experimental warming altered rates of carbon processes, allocation, and carbon storage in a tallgrass prairie: A data assimilation approach, *Ecosphere*, 6(11), 210 doi:10.1890/ES14-00335.1.
- Sierra, C. A. (2012), Temperature sensitivity of organic matter decomposition in the Arrhenius equation: Some theoretical considerations, *Biogeochemistry*, 108(1–3), 1–15.
- Sierra, C. A., M. Muller, and S. E. Trumbore (2012), Models of soil organic matter decomposition: The SoilR package, version 1.0, *Geosci. Model Dev.*, 5(4), 1045–1060.
- Sierra, C. A., S. E. Trumbore, E. A. Davidson, S. Vicca, and I. Janssens (2015a), Sensitivity of decomposition rates of soil organic matter with respect to simultaneous changes in temperature and moisture, *J. Adv. Model. Earth Syst.*, 7(1), 335–356.
- Sierra, C. A., S. Malghani, and M. Müller (2015b), Model structure and parameter identification of soil organic matter models, *Soil Biol. Biochem.*, 90, 197–203.
- Sistla, S. A., and J. P. Schimel (2013), Seasonal patterns of microbial extracellular enzyme activities in an arctic tundra soil: Identifying direct and indirect effects of long-term summer warming, *Soil Biol. Biochem.*, 66, 119–129.
- Smith, B., I. C. Prentice, and M. T. Sykes (2001), Representation of vegetation dynamics in the modelling of terrestrial ecosystems: Comparing two contrasting approaches within European climate space, *Global Ecol. Biogeogr.*, 10, 621–638.
- Smith, J. U., P. Smith, R. Monaghan, and J. MacDonald (2002), When is a measured soil organic matter fraction equivalent to a model pool?, *Eur. J. Soil Sci.*, 53(3), 405–416.
- Smith, M. J., D. W. Purves, M. C. Vanderwel, V. Lyutsarev, and S. Emmott (2013), The climate dependence of the terrestrial carbon cycle, including parameter and structural uncertainties, *Biogeosciences*, 10(1), 583–606.
- Stewart, C. E., A. F. Plante, K. Paustian, R. T. Conant, and J. Six (2008), Soil carbon saturation: Linking concept and measurable carbon pools, *Soil Sci. Soc. Am. J.*, 72(2), 379–392.
- Tarnocai, C., J. G. Canadell, E. A. G. Schuur, P. Kuhry, G. Mazhitova, and S. Zimov (2009), Soil organic carbon pools in the northern circumpolar permafrost region, *Global Biogeochem. Cycles*, 23, GB2023, doi:10.1029/2008GB003327.
- Tenney, F. G., and S. A. Waksman (1929), Composition of natural organic materials and their decomposition in the soil: IV. The nature and rapidity of decomposition of the various organic complexes in different plant materials, under aerobic conditions, *Soil Sci.*, 28(1), 55–84.
- Thonicke, K., A. Spessa, I. C. Prentice, S. P. Harrison, L. Dong, and C. Carmona-Moreno (2010), The influence of vegetation, fire spread and fire behaviour on biomass burning and trace gas emissions: Results from a process-based model, *Biogeosciences*, 7(7), 2191–2191.
- Tian, H., et al. (2015), Global patterns and controls of soil organic carbon dynamics as simulated by multiple terrestrial biosphere models: Current status and future directions, *Global Biogeochem. Cycles*, 29, 775–792, doi:10.1002/2014GB005021.
- Todd-Brown, K. E. O., J. T. Randerson, W. M. Post, F. M. Hoffman, C. Tarnocai, E. A. G. Schuur, and S. D. Allison (2013), Causes of variation in soil carbon simulations from CMIP5 Earth system models and comparison with observations, *Biogeosciences*, 10(3), 1717–1736.
- Todd-Brown, K. E. O., et al. (2014), Changes in soil organic carbon storage predicted by Earth system models during the 21st century, *Biogeosciences*, 11(8), 2341–2356.

- Trumbore, S. (2009), Radiocarbon and soil carbon dynamics, *Annu. Rev. Earth Planet. Sci.*, *37*, 47–66.
- Waksman, S. A. (1952), *Soil Microbiology*, John Wiley, New York.
- Wang, G. B., Y. Zhou, X. Xu, H. H. Ruan, and J. S. Wang (2013a), Temperature sensitivity of soil organic carbon mineralization along an elevation gradient in the Wuyi Mountains, China, *Plos One*, *8*, doi:10.1371/journal.pone.0053914.
- Wang, G. S., W. M. Post, and M. A. Mayes (2013b), Development of microbial-enzyme-mediated decomposition model parameters through steady-state and dynamic analyses, *Ecol. Appl.*, *23*(1), 255–272.
- Wang, Q. K., S. L. Wang, T. X. He, L. Liu, and J. B. Wu (2014a), Response of organic carbon mineralization and microbial community to leaf litter and nutrient additions in subtropical forest soils, *Soil Biol. Biochem.*, *71*, 13–20.
- Wang, Y. P., R. Leuning, H. A. Cleugh, and P. A. Coppin (2001), Parameter estimation in surface exchange models using nonlinear inversion: How many parameters can we estimate and which measurements are most useful?, *Global Change Biol.*, *7*(5), 495–510.
- Wang, Y. P., B. C. Chen, W. R. Wieder, M. Leite, B. E. Medlyn, M. Rasmussen, M. J. Smith, F. B. Augusto, F. Hoffman, and Y. Q. Luo (2014b), Oscillatory behavior of two nonlinear microbial models of soil carbon decomposition, *Biogeosciences*, *11*(7), 1817–1831.
- Weng, E. S., and Y. Q. Luo (2011), Relative information contributions of model vs. data to short- and long-term forecasts of forest carbon dynamics, *Ecol. Appl.*, *21*(5), 1490–1505.
- Weng, E. S., Y. Q. Luo, C. Gao, and R. Oren (2011), Uncertainty analysis of forest carbon sink forecast with varying measurement errors: A data assimilation approach, *J. Plant. Ecol.*, *4*(3), 178–191.
- Weng, E. S., Y. Q. Luo, W. L. Wang, H. Wang, D. J. Hayes, A. D. McGuire, A. Hastings, and D. S. Schimel (2012), Ecosystem carbon storage capacity as affected by disturbance regimes: A general theoretical model, *J. Geophys. Res.*, *117*, G03014, doi:10.1029/2012JG002040.
- West, T. O., V. Bandaru, C. C. Brandt, A. E. Schuh, and S. M. Ogle (2011), Regional uptake and release of crop carbon in the United States, *Biogeosciences*, *8*, 2037–2046.
- Wieder, W., A. Grandy, C. Kallenbach, and G. Bonan (2014), Integrating microbial physiology and physiochemical principles in soils with the Microbial-Mineral Carbon Stabilization (MIMICS) model, *Biogeosci. Discuss.*, *11*(1), 1147–1185.
- Wieder, W. R., G. B. Bonan, and S. D. Allison (2013), Global soil carbon projections are improved by modelling microbial processes, *Nat. Clim. Change*, *3*(10), 909–912.
- Wieder, W. R., et al. (2015), Explicitly representing soil microbial processes in Earth system models, *Global Biogeochem. Cycles*, *29*, doi:10.1002/2015GB005188.
- 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, *Geosci. Model Dev.*, *5*(5), 1259–1271.
- Xia, J. Y., Y. Q. Luo, Y. P. Wang, and O. Hararuk (2013), Traceable components of terrestrial carbon storage capacity in biogeochemical models, *Global Change Biol.*, *19*(7), 2104–2116.
- Xia, J. Y., et al. (2015), Joint control of terrestrial gross primary productivity by plant phenology and physiology, *Proc. Natl. Acad. Sci. U.S.A.*, *112*(9), 2788–2793.
- Xiao, J., K. J. Davis, N. M. Urban, and K. Keller (2014), Uncertainty in model parameters and regional carbon fluxes: A model-data fusion approach, *Agric. For. Meteorol.*, *189*, 175–186.
- Xu, T., L. White, D. F. Hui, and Y. Q. Luo (2006), Probabilistic inversion of a terrestrial ecosystem model: Analysis of uncertainty in parameter estimation and model prediction, *Global Biogeochem. Cycles*, *20*, GB2007, doi:10.1029/2005GB002468.
- Xu, X., Y. Q. Luo, and J. Z. Zhou (2012), Carbon quality and the temperature sensitivity of soil organic carbon decomposition in a tallgrass prairie, *Soil Biol. Biochem.*, *50*, 142–148.
- Xu, X., Z. Shi, D. Li, A. Rey, H. Ruan, J. M. Craine, J. Liang, J. Zhou, and Y. Luo (2016), Soil properties control decomposition of soil organic carbon: Results from data-assimilation analysis, *Geoderma*, *262*, 235–242.
- Yan, Y., Y. Q. Luo, X. H. Zhou, and J. M. Chen (2014), Sources of variation in simulated ecosystem carbon storage capacity from the 5th Climate Model Intercomparison Project (CMIP5), *Tellus B*, *66*, 22568, doi:10.3402/tellusb.v66.22568.
- Zhang, D. Q., D. F. Hui, Y. Q. Luo, and G. Y. Zhou (2008), Rates of litter decomposition in terrestrial ecosystems: Global patterns and controlling factors, *J. Plant. Ecol.*, *1*(2), 85–93.
- Zhou, T., and Y. Q. Luo (2008), Spatial patterns of ecosystem carbon residence time and NPP-driven carbon uptake in the conterminous United States, *Global Biogeochem. Cycles*, *22*, GB3032, doi:10.1029/2007GB002939.
- Zhou, T., P. J. Shi, G. S. Jia, and Y. Q. Luo (2013), Nonsteady state carbon sequestration in forest ecosystems of China estimated by data assimilation, *J. Geophys. Res. Biogeosci.*, *118*, 1369–1384, doi:10.1002/jgrg.20114.
- Zhou, X. H., T. Zhou, and Y. Q. Luo (2012), Uncertainties in carbon residence time and NPP-driven carbon uptake in terrestrial ecosystems of the conterminous USA: A Bayesian approach, *Tellus B*, *64*, 17223, doi:10.3402/tellusb.v64i0.17223.