Matrix Approach to Land Carbon Cycle Modeling

Yiqi Luo[1,2,3], Yuanyuan Huang[4], Carlos A. Sierra[5,6], Jianyang Xia[7], Anders Ahlström[8], Yizhao Chen[9], Oleksandra Hararuk[10], Enqing Hou[11], Lifen Jiang[1], Cuijuan Liao[12], Xingjie Lu[13], Zheng Shi[14], Benjamin Smith[8,15], Feng Tao[12], and Ying-Ping Wang[4]

[1]Center for Ecosystem Science and Society, Northern Arizona University, Flagstaff, AZ, USA, [2]Department of Biological Sciences, Northern Arizona University, Flagstaff, AZ, USA, [3]Now at School of Integrative Plant Science, Cornell University, Ithaca, NY, USA, [4]CSIRO Oceans and Atmosphere, Aspendale, VIC, Australia, [5]Max Plank Institute of Biogeochemistry, Jena, Germany, [6]Department of Ecology, Swedish University of Agricultural Sciences, Uppsala, Sweden, [7]Center for Global Change and Ecological Forecasting, School of Ecological and Environmental Sciences, East China Normal University, Shanghai, China, [8]Department of Physical Geography and Ecosystem Science, Lund University, Lund, Sweden, [9]Joint Innovation Center for Modern Forestry Studies, College of Biology and the Environment, Nanjing Forestry University, Nanjing, China, [10]Northern Forestry Centre, Canadian Forest Service, Natural Resources Canada, Edmonton, AB, Canada, [11]Key Laboratory of Vegetation Restoration and Management of Degraded Ecosystems, South China Botanical Garden, Chinese Academy of Sciences, Guangzhou, China, [12]Department of Earth System Science, Ministry of Education Key Laboratory for Earth System Modeling, Institute for Global Change Studies, Tsinghua University, Beijing, China, [13]School of Atmospheric Sciences, Sun Yat-sen University, Guangzhou, China, [14]The Institute for Environmental Genomics, University of Oklahoma, Norman, OK, USA, [15]Hawkesbury Institute for the Environment, Western Sydney University, Richmond, VIC, Australia

Abstract Land ecosystems contribute to climate change mitigation by taking up approximately 30% of anthropogenically emitted carbon. However, estimates of the amount and distribution of carbon uptake across the world's ecosystems or biomes display great uncertainty. The latter hinders a full understanding of the mechanisms and drivers of land carbon uptake, and predictions of the future fate of the land carbon sink. The latter is needed as evidence to inform climate mitigation strategies such as afforestation schemes. To advance land carbon cycle modeling, we have developed a matrix approach. Land carbon cycle models use carbon balance equations to represent carbon exchanges among pools. Our approach organizes this set of equations into a single matrix equation without altering any processes of the original model. The matrix equation enables the development of a theoretical framework for understanding the general, transient behavior of the land carbon cycle. While carbon input and residence time are used to quantify carbon storage capacity at steady state, a third quantity, carbon storage potential, integrates fluxes with time to define dynamic disequilibrium of the carbon cycle under global change. The matrix approach can help address critical contemporary issues in modeling, including pinpointing sources of model uncertainty and accelerating spin-up of land carbon cycle models by tens of times. The accelerated spin-up liberates models from the computational burden that hinders comprehensive parameter sensitivity analysis and assimilation of observational data to improve model accuracy. Such computational efficiency offered by the matrix approach enables substantial improvement of model predictions using ever-increasing data availability. Overall, the matrix approach offers a step change forward for understanding and modeling the land carbon cycle.

Plain Language Summary Earth system models (ESMs) are the tools we have to predict future states of climate and ecosystems. However, land carbon cycle models, a critical component of ESMs, are highly diverse in both structures and predictions, hindering our ability to obtain consistent future projections. The latter is needed as part of the evidence base to inform climate change mitigation strategies. This paper describes a matrix approach that unifies land carbon cycle models in one matrix form. The matrix models offer consistency and simplicity in structure that make the models analytically tractable. In addition, the matrix approach provides a theoretical framework for understanding the general behavior of the land carbon cycle. More importantly, the matrix approach solves some key contemporary issues in land carbon cycle modeling, such as pinning down sources of model uncertainty and accelerating spin-up. The accelerated spin-up speeds up land carbon cycle simulations by tens of times, making it feasible to perform parameter sensitivity analysis and data assimilation to constrain models with big data. Overall, the matrix approach represents a step change forward for understanding and modeling the land carbon cycle.

1. Introduction

Modeling research in earth system science has blossomed in the past decades as models have become an essential tool to predict future states of the carbon cycle and climate. Hundreds of models have been developed to simulate the land carbon cycle from local to global scale (Manzoni & Porporato, 2009; Sierra, Ceballos-Núñez, et al., 2018). Individual models, such as the Community Land Model (CLM), also became increasingly sophisticated over time as additional processes and internal system feedbacks are incorporated (Lawrence et al., 2019). However, despite growing mechanistic details, model predictive skills have not improved much (Prentice et al., 2015). On the contrary, both model structures and predictions have become less tractable with increasing model complexity. Model intercomparison projects (MIPs) strive to unpack differences among models, the sensitivity of their predictions to drivers, and the processes responsible for such differences in model sensitivity and predictions. Results from nearly all MIPs suggest wide inter-model spread in responses of the terrestrial carbon cycle to climate change (Huntzinger et al., 2017). This large spread has motivated modelers to add more and more process details to address putative mechanistic "gaps" in the models. Paradoxically, the result has been even more divergent model structures, increased complexity, high computational costs, and low tractability (Huntzinger et al., 2017; Lawrence et al., 2018). Overall, modeling the land carbon cycle using current bottom-up modeling architectures faces several roadblocks, including (a) diverse and complex model structures that cannot be easily comprehended nor evaluated; (b) salient model uncertainty with its sources being difficult to identify, which impedes targeted model improvement; and (c) high computational cost that limits the use of rapidly accumulating big data to constrain model prediction. To overcome these roadblocks, it is imperative to develop new approaches to land carbon cycle modeling.

Over the past decades, we have developed a matrix approach in an effort to address the issues identified above via a complementary approach to bottom-up process modeling. The matrix approach helps gain simplicity in structure, clarity in diagnostics, and computational efficiency in spin-up. The latter is a common method to obtain steady states of carbon pools as the estimates of initial carbon pool sizes and is typically the most computationally intensive step in land carbon modeling. Reducing the computational load of spin-up opens the door for comprehensive evaluation and data assimilation to improve the predictive skill of the models. The matrix approach also offers a basis for theoretical analysis of general behavior of land carbon cycle under global change.

Matrix algebra was first used to study steady states and response times of the ocean's geochemical system, including calcium carbonate formation and turnover (Southam & Hay, 1976). Lasaga (1980) used the matrix equations to represent geochemical cycles and explore the stability of carbon and oxygen cycles. Sundquist (1985) performed eigenvalue analysis of matrix equations to examine the interrelationships among carbon cycle processes over a wide range of time scales. We first started organizing a land carbon cycle model in a matrix form to gain analytic spin-up before the model was used to examine the time courses of various processes in response to the elevated CO_2 treatment in an ecosystem experiment (Luo et al., 2001). From this exercise, we found that the matrix representation is simple to understand, convenient for mathematical analysis, and computationally efficient for data assimilation (Luo et al., 2003; White et al., 2005; Xu et al., 2006). Other carbon cycle models, such as InTEC, SiB-CASA, and SiB4, were also analytically spun up to steady states using matrix algebra (Chen et al., 2003; Haynes, Baker, Denning, Stockli, et al., 2019; Haynes, Baker, Denning, Wolf, et al., 2019; Schaefer et al., 2008).

More recently, we further developed the matrix representation as a generalizable approach to land carbon cycle modeling. Xia et al. (2012, 2013) converted the Australian Community Atmosphere Biosphere Land Exchange (CABLE) model to one matrix equation, which was used to develop the semi-analytic spin-up (SASU) and traceability analysis. Huang, Lu, et al. (2018) recoded a vertically resolved soil carbon module, which is a part of CLM4.5, in the matrix form for attribution analysis. Huang, Zhu, et al. (2018) converted a version of the Organizing Carbon and Hydrology in Dynamic Ecosystems (ORCHIDEE) model to the matrix form to facilitate a variance-based sensitivity analysis. Lu et al. (2020) fully implemented a matrix approach to represent the interconnected carbon and nitrogen cycle system of CLM5. Meanwhile, the matrix equation has been used for theoretical studies, for example, categorizing decomposition models (Sierra & Müller, 2015), and diagnosing the predictability and dynamic disequilibrium of the land carbon cycle (Luo et al., 2015, 2017).

This paper reviews our efforts in development and applications of the matrix approach. Specifically, we show how the matrix approach unifies land carbon cycle models and improves understanding of the general, transient behavior of the land carbon cycle. We illustrate how the matrix approach is effective in addressing some

contemporary issues in land carbon cycle modeling, such as pinning down sources of model uncertainty, accelerating spin-up, and constraining models with big data via data assimilation. While microbial models with nonlinear carbon transfers can also be represented in the matrix form (Sierra & Müller, 2015), this paper is focused on model schemes based on first-order kinetics, which is the most common approach in current land carbon cycle models. There are three reasons for this. First, first-order kinetics of carbon transfer have been supported in term of macroscopic patterns by almost all data sets we have synthesized from litter decomposition and soil incubation experiments (Cai et al., 2018; Luo, 2022; Schädel et al., 2014; Xu et al., 2016; Zhang et al., 2008), notwithstanding strong research community interest in the possibility of nonlinear kinetics emerging as a result of microbial community dynamics. Second, most Earth system models (ESMs) still use the classic first-order kinetics and that modeling community is a target audience for this paper. Third, it has been shown that any nonlinear model for which a solution is known can be expressed with an equivalent linear model that has the exact same solution (Metzler et al., 2018).

2. Unifying Land Carbon Cycle Models

Hundreds of models have been developed to simulate and predict carbon dynamics (Manzoni & Porporato, 2009). Individual models differ somewhat in their representations of carbon cycle processes, but all share fundamental features (Sierra & Müller, 2015). For example, all models consider carbon input through photosynthesis, partitioning of photosynthetic carbon among plant pools, and carbon transfer from plant to litter and soil carbon pools. Carbon cycling through the network of pools is determined by mass balance and strongly influenced by external forcing variables like temperature, moisture, and soil texture (Luo et al., 2003; Sierra & Müller, 2015). Thus, the land carbon cycle can be mathematically generalized as a nonautonomous, multi-compartmental system (Rasmussen et al., 2016; Sierra, Ceballos-Núñez, et al., 2018; Yang et al., 2011). Mathematically, a nonautonomous system is a dynamic system with parameters and inputs that vary with time. In contrast, an autonomous system is a system of ordinary differential equations with constant parameter values and inputs (i.e., a time-invariant system).

The system of the land carbon cycle can be represented by a set of ordinary differential equations in a matrix form. Equation B1 (Box 1) unifies land carbon cycle models by accommodating any number of pools (i.e., the number of elements in $X(t)$ in Equation B1) and folding the carbon cycle processes that operate on those pools into five components of Equation B1: carbon input $\mu(t)$, plant carbon allocation B, carbon transfers A, environmental modifier $\xi(t)$, and turnover rates K (Table 1). Note that vertically resolved soil carbon models have a sixth component, vertical transfer. For example, CLM5 with a vertically resolved soil carbon module includes 18 plant pools for each of the 17 vegetation types and 140 soil pools (i.e., 7 pools per layer over 20 layers) in $X(t)$ (Lawrence et al., 2019). Thus, CLM5 assigns 158 elements (i.e., 18 for plant + 140 for soil) to $X(t)$ for one grid cell if it is occupied by one vegetation type and 194 elements (i.e., $3 \times 18 + 140$) if the grid is occupied by three vegetation types (Lu et al., 2020). In contrast, the CABLE model has nine pools, therefore $X(t)$ has nine elements (Wang et al., 2010).

Each of the five (or six) components of Equation B1 can be described either by fixed values, functions, or nested models. For example, carbon input $\mu(t)$ can be simulated by a "bottom-up" approach via scaling up leaf-level biochemical processes in a nested model (Collatz et al., 1992; Farquhar et al., 1980; Rogers et al., 2017). Alternatively, $\mu(t)$ can be simulated by a "top-down" approach based on the linear relationship between photosynthesis and canopy absorbed radiation (light-use efficiency; Running et al., 2004), as in the CASA model (Potter et al., 1993). Plant carbon allocation B has been represented with diverse schemes, including fixed coefficients, functional relationships, resource limitation and optimization, for different models (Ceballos-Núñez et al., 2020; De Kauwe et al., 2014; Xia et al., 2017). The carbon transfer elements of matrix A are usually fixed as constants or vary with plant traits or soil texture as in the CENTURY model (Xia et al., 2013). The diagonal matrix K represents rate processes of mortality and litterfall for plant pools, or organic matter decomposition for litter and soil pools; these rate processes are often parameterized as constants (Huang, Lu, et al., 2018; Xia et al., 2012). Environmental modifiers ξ vary greatly among models with different shapes of response functions that characterize the effect of changes in the environment on rate processes (Lei et al., 2018; Sierra et al., 2015).

These different levels of complexity lead to a hierarchy of models under one overarching theory (Sierra, Ceballos-Núñez, et al., 2018) (Figure 1). For example, the elements of the matrix K are usually constant in ESMs

---

Box 1. Matrix Representation of Land Carbon Cycle Models

The land carbon cycle is usually described by a set of differential equations, which can be re-organized in a matrix form (Luo & Weng, 2011; Luo et al., 2003, 2015, 2016) as:

$$dX(t)/dt = B\mu(t) + A\xi(t)KX(t) \tag{B1}$$

where $X(t)$ is a vector of pool sizes, B is a vector of partitioning coefficients from carbon input to plant pools, $\mu(t)$ is carbon input (e.g., Net Primary Production, NPP), A is a matrix with -1 in the diagonal and transfer coefficients in the off-diagonal, K is a diagonal matrix of process rates (mortality for plant pools and decomposition for litter and soil pools), and $\xi(t)$ is a diagonal matrix of environmental modifiers to represent responses of carbon cycle processes to changes in temperature, moisture, or other factors. Parameters B, K, and A are constants in some models but may vary with time. With time-varying $\mu(t)$, $\xi(t)$, B, A, and K, Equation B1 represents a multi-compartmental, nonautonomous, dynamical system.

When B, A, K and $u(t)$ are functions of state variable X, Equation B1 can become

$$dX(t)/dt = B(X,t)\mu(X,t) + A(X,t)\xi(t)K(X,t)X(t) \tag{B2}$$

Equation B2 represents nonlinear kinetics and usually leads to complex patterns of carbon cycle dynamics (Wang et al., 2014, 2016) that are different from those defined by Equation B1.

It is conceptually possible that B, A, and K are functions of state variable X, but not much empirical evidence has been revealed to support this notion (Cai et al., 2018; K. Zhu et al., 2018; Schädel et al., 2013; Yang et al., 2011; Zhang et al., 2008). On the other hand, carbon input $\mu(t)$ is a function of foliage biomass during fast canopy expansion (Cui et al., 2019). This nonlinear kinetics is nullified once the canopy is developed. Nonlinearity in land carbon cycling induced by other mechanisms, such as seasonal fluctuation, acclimatory changes in photosynthetic capacity, and nonlinear responses to temperature and moisture, can be accounted for by Equation B1 via environmental modifiers $\xi(t)$ and parameter changes (Luo & Schuur, 2020). For simplicity, most analysis in this paper is based on Equation B1. Analysis of nonlinear carbon cycle models such as Equation B2 can be found in, for example, Metzler et al. (2018) and Sierra, Ceballos-Núñez, et al. (2018).

---

with static plant functional types (PFTs; Oleson et al., 2010), but are variable if a model has dynamic vegetation (Moorcroft et al., 2001; Sitch et al., 2003; Smith et al., 2001). To simulate dynamic vegetation, demographic models such as ED2, LPJ-GUESS, or FATES simulate plant physiology, competition, ecosystem assembly, and vegetation distribution (Fisher et al., 2015; Koven et al., 2020). Those processes ultimately influence the carbon cycle by changing values of elements in the matrix K as well as allocation coefficients in vector B, carbon influx $\mu(t)$, and, potentially, transfer coefficients A. Recent modeling efforts linking plant traits to tree demography, or microbial traits to carbon processes, make K a function of traits (Thomas et al., 2019). When the decomposition rates of soil organic matter are represented by Michaelis-Menten equations, as formulated in some microbial models, it generates a nonlinearity (Sierra & Müller, 2015; Wieder et al., 2015). Along this hierarchy of model complexity, turnover rates in matrix K change from constants to progressively complex functions of more and more processes (Figure 1).

Matrix models can be developed by reorganizing carbon balance equations in existing models, through reconstruction from model outputs, or from a flow diagram of element cycling. The CLM5 matrix model was developed by reorganizing hundreds of carbon and nitrogen equations from six modules into four matrix equations representing carbon and nitrogen cycles of vegetation and soil. This matrix version of CLM5 operates as a replica in parallel with the original version online or offline within the Community Earth System Model (CESM; Lu et al., 2020). Matrix models were reconstructed from numerical outputs of carbon stocks and fluxes among pools for LPJ-GUESS, ORCHIDEE, and ELM-ECA (Ahlström et al., 2015; Metzler et al., 2020; Wu et al., 2020). In addition, a new matrix model was developed directly from a phosphorus flow diagram (Hou et al., 2019). The time needed for developing a matrix representation varies greatly, depending on the complexity of the original model. For example, a team of three (E. Hou, X. Lu, and Y. Luo) worked together for less than 1 day to

Table 1
Examples of Models or Processes With Different Representations of the Five Terms in the Matrix Equation

		Number of pools	A, carbon transfer	$\xi(t)$, environmental effects	K, process rates	B, carbon allocation	$\mu(t)$, carbon input
Model	CLM5.0	≥158	Fixed coefficients with vertical transfer of both litter and soil carbon via e-folding over 20 layers	Q_{10} for temperature, log-linear for water, exponential decrease for depth.	Plant mortality and fire release by nested models, fixed values for litter and soil carbon decomposition	Fixed coefficients for nonwoody PFT Dynamic functions for woody PFT	Biochemical photosynthesis model with scaling by multi-layer canopy
	CABLE	14	Fixed coefficients without loops of carbon from slow and passive pools to active soil pool without layers	Q_{10} for temperature and Power function for water scalar	Fixed values for litter and soil carbon decomposition. Leaf litterfall is a function of phenology	Optional fixed or dynamic by resource limitation	Biochemical photosynthesis model with scaling by two-leaf canopy
	LPJ-GUESS	9	Fixed coefficients	Q_{10} for temperature scalar and linear function for water scalar	Plant mortality and fire release by nested models, fixed values for litter and soil carbon decomposition	Dynamic by allometric constraints	Light use efficiency model with different vegetation types
	MIMICS	7	Fixed coefficients. No vegetation pool	Q_{10} for temperature on decomposition rate	Soil carbon decomposition regulated by microbial enzymes	N/A	Litter input
Process	Trait-based modeling		Influenced by lignin:nitrogen ratio in leaf, leaf lignin fraction, wood lignin fraction, and microbial traits		Mortality influenced by hydraulic and other traits; decomposition by lignin fraction and microbial traits	Availability of light, nitrogen, water in root and wood	GPP; PFT-based carbon use efficiency
	Vegetation dynamics		Influenced by vegetation via litter quality	Varying scalars for different vegetation	Varying mortality and litter decomposition with vegetation	Varying with vegetation	Altering capacity of photosynthesis
	Soil microbial processes		Influenced by microbial taxa, functional attributes, and processes		Influenced of microbial taxa, functional attributes, and processes on decomposition	Microbial influence on plant carbon allocation	
	Nitrogen		Influenced via altering litter quality		Altering rates of decomposition	Altering root to shoot ratio	Altering capacity of photosynthesis

develop a running phosphorus matrix model from a flow diagram. In comparison, the development of the matrix model for CLM5-CN by Lu et al. (2020) took more than 1 year of careful work to guarantee its compatibility with the overall framework of CESM2. To date, dozens of existing models have been converted into the matrix form, including CLM3.5, CLM4, CLM4.5, CLM5, ORCHIDEE, CABLE, LPJ-GUESS, and nonlinear microbial models (Ahlström et al., 2015; Huang, Lu, et al., 2018; Huang, Zhu, et al., 2018; Lu et al., 2020; Sierra & Müller, 2015; Wu et al., 2020; Xia et al., 2012). Four coupled carbon-nitrogen cycle models, CABLE, CLM5, TECO, and GECO, were converted into the matrix form (Lu et al., 2020; Shi et al., 2016; Wang et al., 2022; Xia et al., 2012). A fully coupled carbon-nitrogen-phosphorus model, TECO-CNP, was developed in the matrix form to study effects of nutrient limitations on subtropical forests (Du et al., 2021). The only dynamic vegetation model yet to be converted to the matrix form is LPJ-GUESS, in which vegetation pools and fluxes were lumped together to estimate vegetation turnover (Ahlström et al., 2015). While this was sufficient for the purposes of the target study, it is feasible to represent PFTs and demographic dynamics comprehensively within the structure of the matrix equation. Indeed, demographic models are often expressed in a matrix form (Caswell, 2001; Tuljapurkar & Caswell, 1997).
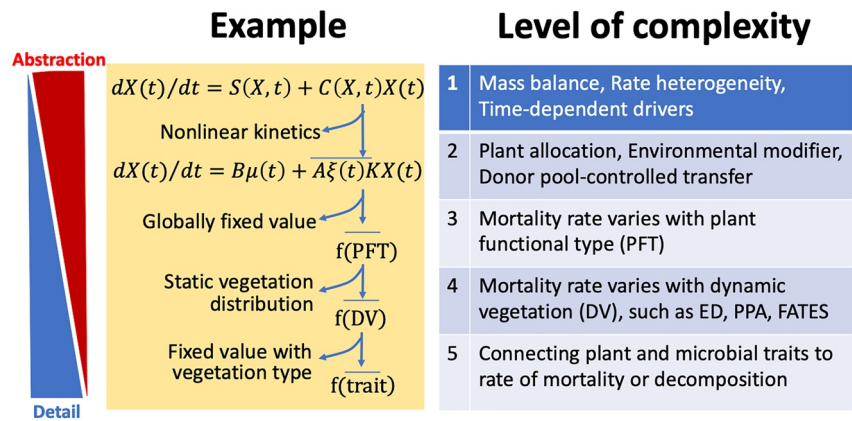
Figure 1. Hierarchical representation of land carbon models with different levels of complexity. The extremely general representation of the carbon cycle that is physically possible using mass balance constraints (Sierra & Müller, 2015) (Level 1). The inputs to the system are captured by the vector $S(X, t)$ that adds external C to all pools as a function of time and state variable X. The releases are the product of a compartmental matrix $C(X, t)$ that accounts for interactions with state variable similarly as in Equation B2 in Box 1. Any term in the level 1 equation can be expanded to accommodate more details. For example, $C(X, t)$ can be expressed (depicted by blue arrows) as either nonlinear kinetics or linear kinetics as related to process rates (K), donor-controlled transfer rates (A), and environmental modifiers ($\xi(t)$) (Level 2). These terms, together with carbon allocation (B) and time-dependent photosynthetic rates ($\mu(t)$), form a more specific representation of carbon cycle models as in Equation B1. Each of these terms can be developed further in a hierarchical way. For example, the process rates K can be globally fixed values or vary with plant functional types (Level 3). The latter can be either static or dynamic (Level 4). The dynamic vegetation can be linked to either fixed parameter values for each vegetation type or plant and microbial traits (Level 5). This flexibility in the system representation allows any known land carbon cycle model architecture to be generalized under one theoretical framework without compromising detail in process representation.

Once developed, matrix models allow for new processes to be incorporated, as any process that influences carbon cycle must be related to one or more of the five (or six) components of the matrix equation (Table 1). For example, vegetation dynamics, which usually result from competition among plants with different strategies, influence carbon input rate $\mu(t)$, plant allocation B, mortality and litter decomposition rates K, and carbon transfer A due to litter quality changes for different vegetation types (Fisher et al., 2015; Koven et al., 2020; Moorcroft et al., 2001; Smith et al., 2001). Nitrogen and phosphorus cycle processes primarily modify the rates of photosynthesis $\mu(t)$, allocation B, transfer A, and mortality and decomposition rates K (Q. Zhu et al., 2019; Wang et al., 2010), but do not directly affect the environmental modifier $\xi(t)$. Trait-based modeling has the potential to link plant and microbial traits to photosynthesis, allocation, mortality, decomposition, transfer, rooting depth, and plant and microbial responses to environmental variables (Fry et al., 2019; Thomas et al., 2019). Thus, it may influence all five components (six if vertical mixing is resolved) of the matrix equation. The newly incorporated processes can be evaluated for their impacts on modeled carbon cycle dynamics using a traceability analysis as discussed in Section 4.1 (e.g., Liao et al., in prep.).

3. Understanding General Behavior of Land Carbon Cycle Under Global Change

Despite extensive research using modeling, experimental, and observational approaches, key gaps in our theoretical understanding of the land carbon cycle have remained for the past six decades since Olson (1963) characterized steady-state dynamics at which the carbon storage capacity equals the product of carbon input and residence time. However, it is unlikely that the global carbon cycle has ever been at steady state. In contrast, a constantly changing environment keeps the carbon cycle in a dynamic disequilibrium. Since pre-industrial times, the most important factor behind this disequilibrium is that human activities result in substantial release of CO_2 into the atmosphere from land cultivation and fossil fuel burning (Luo & Weng, 2011). As a nonautonomous, multi-compartment system (Rasmussen et al., 2016; Sierra, Ceballos-Núñez, et al., 2018), how land ecosystem carbon balance varies under dynamic disequilibrium has not been theoretically well explored.

The matrix equation as a generic expression of carbon cycle model structure makes it possible to study the general behavior of the land carbon cycle. The matrix equation B1 in Box 1 can be rearranged to describe carbon storage capacity and carbon storage potential (Luo et al., 2017) as:

$$X(t) = (-A\xi(t)K)^{-1}B\mu(t) - (-A\xi(t)K)^{-1}X'(t) \tag{1a}$$

$$= \tau_E(t)\mu(t) - \tau_{ch}X'(t) \tag{1b}$$

$$= X_c(t) - X_p(t), \tag{1c}$$

where $X'(t)$ equals $dX(t)/dt$ in Equation B1 and represents net pool change, τ_{ch} is chasing time, and $\tau_E(t)$ is ecosystem residence time. Chasing time τ_{ch} measures how fast the net C pool change, $X'(t)$, is redistributed through the network of multiple carbon pools (Luo et al., 2017). Ecosystem residence time $\tau_E(t)$ equals:

$$\tau_E(t) = (-A\xi(t)K)^{-1}B. \tag{2}$$

$X_c(t)$ is the land carbon storage capacity. It can be expressed as:

$$X_c(t) = (-A\xi(t)K)^{-1}B\mu(t) = \tau_E(t)\mu(t). \tag{3}$$

$X_p(t)$ is the land carbon storage potential. It can be described as:

$$X_p(t) = (-A\xi(t)K)^{-1}X'(t) = \tau_{ch}(t)X'(t). \tag{4}$$

The carbon storage capacity is the maximal amount of carbon that a land ecosystem can store. Traditionally, the storage capacity has been considered as a static concept that results from the balance between carbon input and decomposition when an ecosystem is approximately at steady state (Olson, 1963) (Figure 2a). For example, long-term averaged carbon stock in the soil is lower in tropical than boreal forests, due to faster decomposition rates (i.e., shorter residence time) in tropical soils, notwithstanding higher carbon input from litterfall and plant mortality (Crowther et al., 2019). When this definition is applied to an ecosystem with multiple carbon compartments under non-steady state (i.e., a time-dependent, nonautonomous system), the carbon storage capacity is no longer a static constant; instead, it becomes time-dependent, as both carbon input and residence time vary with time. Regardless, it still sets the maximal amount of carbon that an ecosystem can store at time t. Thus, carbon storage capacity represents an instantaneously driving state that carbon storage $X(t)$ chases, varying over seasons of a year (Figure 2b) and with climate change over decades (Figure 2c). When the capacity $X_c(t)$ is larger than the
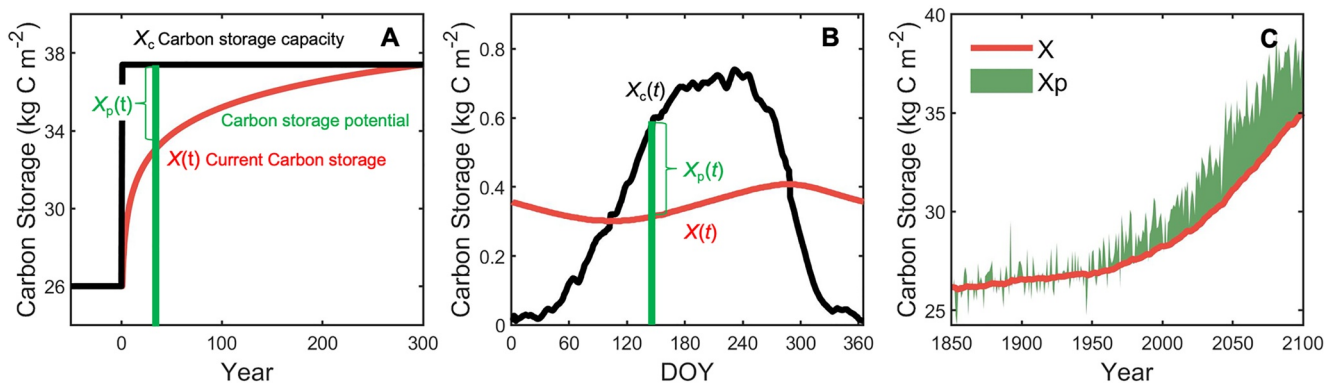

Figure 2. Transient carbon cycle dynamics as determined by carbon storage capacity and potential on different time scales. Panel (a) is for an autonomous system in which carbon storage capacity (X_c) is a constant while carbon storage potential ($X_p(t)$) and carbon storage itself ($X(t)$) vary with time. In the illustrated case, the capacity is assumed to abruptly increase by 40%, mainly due to instantaneous increase in carbon input as in a CO_2 experiment (Luo & Reynolds, 1999). Consequently, the potential immediately increases and then gradually declines as $X(t)$ increases toward equilibrium. Panel (b) illustrates time-dependent $X(t)$, its capacity $X_c(t)$ and potential $X_p(t)$ in a nonautonomous system over days of year (DOY; Rasmussen et al., 2016). Seasonal change in the capacity is due to change in carbon input, which is low in winter and high in summer. Panel (c) illustrates time-dependent $X(t)$, its capacity and potential in a nonautonomous system over years in response to climate change (Luo et al., 2017). Carbon input and residence time, which jointly determine the capacity, fluctuate over the year and change with climate change, resulting in strong fluctuation in the storage capacity while $X(t)$ of multiple pools smoothly changes over time (Rasmussen et al., 2016). In panels (b and c), the capacity is a driving state that $X(t)$ chases. The rate of chasing is proportional to the storage potential.

carbon storage $X(t)$ itself, carbon sequestration occurs and carbon stock increases; otherwise, the ecosystem loses carbon and carbon stock decreases.

The land carbon storage potential or its variants (e.g., carbon sink or sequestration potential) is a term that has been used in various contexts (Hudiburg et al., 2009; Zhou et al., 2015), but it has not been rigorously defined before. The matrix equation yields a precise definition for and offers a biogeochemical interpretation of the carbon storage potential (Equation 4). It is the amount of carbon that an ecosystem can take up to reach the carbon storage capacity and is a measure of the magnitude of disequilibrium (Luo & Weng, 2011). Thus, carbon storage potential represents the internal capability of an ecosystem to equalize its current carbon storage with the carbon storage capacity. Biogeochemically, it describes the redistribution of net carbon pool change, $X'(t)$, of individual pools at each time step through networked pools connected by carbon transfers from one pool to the others through all the pathways, each pool having its distinct turnover time (Luo et al., 2017). The time to fully equalize the current carbon storage with the capacity through the matrix of pools is measured by chasing time, which can be statistically approximated by the slope of regression between the carbon storage potential and net ecosystem production (NEP; Jiang et al., 2017; Luo et al., 2017).

Both the carbon storage capacity and potential are regulated by ecosystem-internal processes that govern carbon influx and residence time. Carbon influx represents the combined effect of all the processes at biochemical, leaf, and canopy scales that influence plant photosynthesis. Many other processes influence residence time. For example, an increase in plant mortality results in a decrease in residence time and consequently carbon storage. An increase in microbial carbon use efficiency, has the contrary effect of prolonging residence time and carbon sequestration. External forcing variables, such as cyclic changes in light and temperature, long-term directional changes in global climate, disturbance events, disturbance regime shifts, and state changes, influence the carbon storage capacity and potential through changes in carbon input, plant carbon allocation, rate processes and transfers among plant and soil pools (Luo et al., 2015; Luo & Weng, 2011).

The carbon storage potential mathematically integrates fluxes $X'(t)$ with time $\tau_{ch}(t)$ (see Equation 4). It quantifies how net ecosystem exchange (NEE; $X'(t)$), an instantaneous measure of carbon fluxes at a given time, is transformed through a network of multiple pools. The time scale of NEE transformation within the multi-pool network is measured by chasing time $\tau_{ch}(t)$ (Luo et al., 2017). Carbon chasing time in the multi-pool network cannot be easily measured in the field, as the data required to estimate turnover time (equaling to an inverse of elements in matrix K) of individual carbon pools and transfer coefficients among them (element in matrix A) are usually not available. Nonetheless, the carbon storage potential can be estimated from multiple datasets. For instance, radiocarbon ^{14}C has been used to quantify mean carbon ages of various plant, litter, and soil pools (Gaudinski et al., 2000; Shi et al., 2020). Soil respiration and its isotopic signature can be used to estimate the age of the release flux out of an ecosystem, which is a measure of turnover time (K^{-1}) (Sierra, Hoyt, et al., 2018; Sierra et al., 2012; Trumbore et al., 1996). Carbon contents in various pools—leaf, wood, root, litter, and soil—contain information on timescales of carbon cycling. Together, those datasets have been used to constrain multi-pool soil carbon models through data assimilation (Liang et al., 2018; Xu et al., 2006). Therefore, they could be used to estimate the carbon storage potential. Note that carbon sequestration between two time points is a result of changes in both the carbon storage capacity and potential (Jiang et al., 2021).

Overall, the matrix equation offers a new foundation for understanding the general behavior of the land carbon cycle and quantifying carbon uptake or loss by ecosystems at different timescales.

4. Addressing Contemporary Issues

Once models are expressed in the matrix form, we can address a range of perennial issues in land carbon cycle modeling, such as sources of model uncertainty, computational efficiency of spin-up for complex models, and uses of big data to constrain model predictions.

4.1. Pinning Down Sources of Model Uncertainty With Traceability Analysis

Despite extensive research and model development over the past 20 years, the model uncertainty (i.e., across-model spread) in ESM projections of the global land C sink has remained large from the Third Assessment Report of the
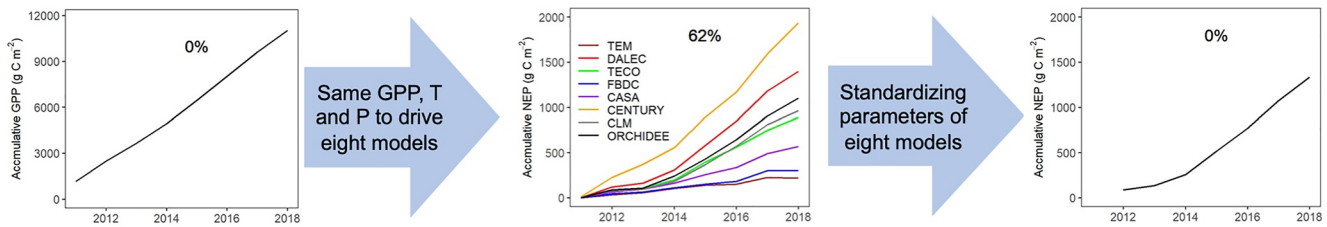
Figure 3. Spreading and shrinking model uncertainty. We developed a matrix-based model intercomparison project by expressing eight land carbon cycle models (i.e., TEM, CENTURY, DALEC2, TECO, FDBC, CASA, CLM4.5, and Organizing Carbon and Hydrology in Dynamic Ecosystems [ORCHIDEE]) in a unified matrix form. The eight models differ greatly in complexity, with the number of carbon pools ranging from 2 in TEM to 101 in ORCHIDEE. All eight models were driven by the same gross primary production (GPP) and environmental variables (temperature, T and precipitation, P) from a peatland in northern Minnesota, USA. Despite the same forcing data, the variations among the models are large for cumulative net ecosystem production (NEP). The model uncertainty shrinks to almost zero with nearly identical predictions by the eight models when we standardize these parameters in the matrix equations related to plant carbon partitioning coefficients B, transfer coefficients A, environmental scalar $\xi(t)$, vertical mixing coefficients, and decomposition coefficients K (Hou et al., submitted).

Intergovernmental Panel on Climate Change published in 2001 to the sixth report published in 2021 (Canadell et al., 2021; Ciais et al., 2014; Denman et al., 2007; Prentice et al., 2001). To address the model uncertainty issue effectively we must understand what causes it. The matrix approach offers a unique method to precisely pin down sources of model uncertainty.

Once models are expressed in the unified matrix form, the model-to-model variations can be easily traced to differences between models in the five (or six) terms in Equation B1. For example, the model-model differences can be shrunk to zero (i.e., identical predictions) among eight models if the terms in Equation B1 are standardized by adjusting response functions $\xi(t)$ and changing values of parameters B, A, and K to yield the same outputs for each of the terms (Hou et al., submitted) (Figure 3). This manipulation demonstrates that the matrix approach makes model uncertainty a tractable issue: through simple mathematical adjustments in the matrix equation terms, we can analytically and precisely track the model uncertainty down to its sources. To do so, we developed a traceability analysis based on the mathematical properties of the matrix equation. The traceability analysis first decomposes simulated carbon pool sizes at a given time t_0 $X(t = t_0)$ into carbon storage capacity $X_c(t = t_0)$ and carbon storage potential $X_p(t = t_0)$ according to Equation 1c. X_c can be further decomposed to carbon residence time $\tau_E(t = t_0)$ and net primary production $\mu(t = t_0)$ according to Equation 3. Meanwhile, $X_p(t = t_0)$ can be decomposed to chasing time $\tau_{ch}(t = t_0)$ and NEP $X'(t = t_0)$ according to Equation 1b. The components $(A\xi(t = t_0)K)^{-1}$ in $\tau_E(t = t_0)$ or $\tau_{ch}(t = t_0)$ can be further decomposed to A^{-1}, $\xi(t = t_0)^{-1}$, and K^{-1}, which are related to parameters or environmental variables.

The traceability analysis can place any modeling outputs in a 3-dimension space ordinated by carbon input, residence time, and carbon storage potential to understand how modeled land carbon storage differs across an ensemble of models, regardless of differences in model structure and external forcing variables (Jiang et al., 2017; Luo et al., 2017; S. Zhou et al., 2018; Wu et al., 2020). Once model-model differences are identified, sources of uncertainty in model predictions can be tracked down hierarchically to model structures, parameters, forcing variables as well as influences of different initial values in simulations after spin-up (Jiang et al., 2017; Luo et al., 2017; Xia et al., 2013). For example, driven with similar forcing data, CLM3.5 predicted ~31% larger carbon storage capacity than CABLE (Figure 4). Rafique et al. (2016) used the matrix-based traceability analysis to show that this difference is due to 37% higher NPP and 11 years shorter ecosystem residence time in CLM3.5 than these in CABLE (Figure 4). The latter results from shorter baseline carbon residence time for the woody biomass pool in CLM3.5 than CABLE (14 years in CLM3.5 vs. 23 years in CABLE), and a lower proportion of NPP allocated to woody biomass in CABLE (16%) than CLM3.5 (23%). The traceability analysis has also been successfully applied to understand the spread in carbon storage capacity among different biomes in CABLE (Xia et al., 2013) and the different responses of Harvard and Duke Forests to climate change (Jiang et al., 2017).

Once the main sources of model uncertainty are identified, we can target model improvement efforts toward the parameters, processes or feedbacks responsible for that uncertainty, using empirical data to inform more realistic model structure (Luo et al., 2012) and parameterization (Luo & Schuur, 2020). More accurate model parameterization can also be achieved by data assimilation, which is described in Section 4.3.
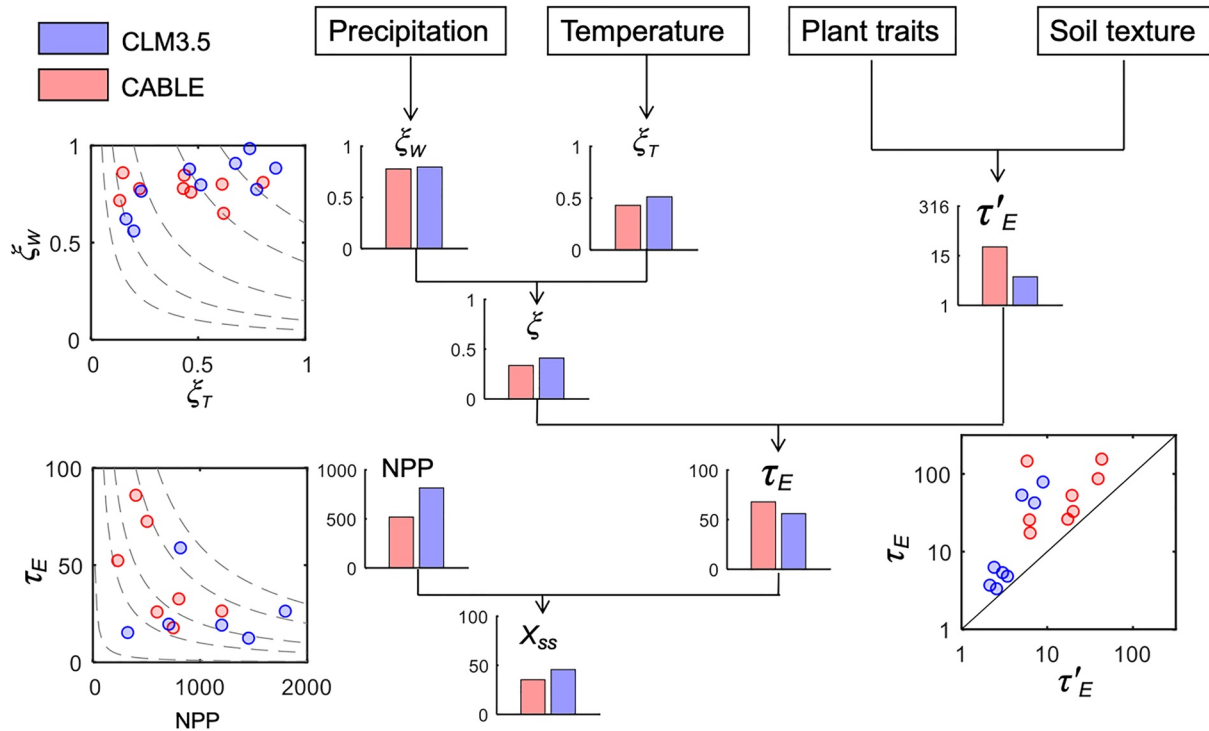
Figure 4. Clarity in diagnostics enabled by the matrix equation. A traceability analysis shows CLM3.5 has a higher carbon storage capacity at steady state (X_{ss}), due to higher NPP than CABLE. However, carbon residence time (τ_E) is higher in Community Atmosphere Biosphere Land Exchange (CABLE) than Community Land Model 3.5 (CLM 3.5) mainly due to difference in the baseline carbon residence time (τ'_E). The environmental scalar ξ, including its temperature (ξ_T) and water components (ξ_w), is similar between the two models. CABLE has higher τ_E than CLM 3.5 mainly because CABLE allocates more carbon to wood growth than CLM 3.5 (Rafique et al., 2016). Model differences in NPP can also be traced to processes as shown in other studies (Cui et al., 2019; Xia et al., 2015). Overall, the matrix approach offers clarity in diagnostics to track model differences to model structure and/or parameter values. Note that symbols in inserted figures represent different vegetation types.

4.2. Accelerating Spin-Up With Semi-Analytic Spin-Up

Spin-up in biogeochemical models is a common procedure to estimate initial values of carbon pool sizes. Spinning up a complex biogeochemical model is computationally very expensive. For example, it takes more than 35 days on Supercomputer Cheyenne to run CLM5 to steady state using the native dynamic method, or about 1 week using the accelerated decomposition (AD) method (Lawrence et al., 2019). The computationally expensive spin-up procedure takes resources away from some critical analyses, such as parameter sensitivity analysis, that could inform improvements in CLM5 or similar models. An efficient spin-up method can greatly liberate models from this computational burden.

Once a model is expressed in a matrix form, the steady-state pool sizes at a given time $t = t_0$ can be analytically obtained by letting Equation B1 equal to zero:

$$\frac{dX(t=t_0)}{dt} = B\mu(t=t_0) + A\xi(t=t_0)KX(t=t_0) = 0. \tag{5}$$

Equation 5 can then be solved to obtain steady-state pool sizes as:

$$X_{ss} = (A\xi(t=t_0)K)^{-1}B\mu(t=t_0). \tag{6}$$

In practice, we usually use multi-year forcing data to run a model repeatedly until slow pools reach quasi-equilibrium. When values of forcing variables change, $\mu(t)$ and $\xi(t)$ in Equation 6 take the new values, leading to different X_{ss}. Thus, we need to solve Equation 5 many times to obtain X_{ss} at quasi-equilibrium of slow pools when multi-year forcing data are used to run a model repeatedly (Xia et al., 2012). This is the general procedure of SASU.

SASU has been applied to the carbon-only and coupled carbon-nitrogen versions of CABLE and saved about 92.4% (i.e., 12.2 times faster) and 86.6% (i.e., 7.5 times faster) of the computational time, respectively, in compar-

ison with the default spin-up method (Xia et al., 2012). When the SASU method is implemented into CLM5, it is 49.6 times faster than the native dynamic method and 8.0 times faster than the AD method (Liao et al., in prep.). The AD method is the default spin-up method in CLM5 (Lawrence et al., 2019). It uses a set of acceleration factors to reach steady-state pool sizes (Thornton & Rosenbloom, 2005). Note that forward modeling with matrix equations can be computationally very slow. However, this issue can be overcome by using a sparse matrix algorithm (Lu et al., 2020). Nevertheless, the greatly enhanced computational efficiency of SASU compared to traditional spin-up enables computational resources to be redeployed toward parameter sensitivity analyses (Huang, Zhu, et al., 2018) and data assimilation with complex carbon cycle models (Hararuk et al., 2014, 2015; Shi et al., 2018).

4.3. Constraining Complex Models With Big Data by Matrix-Enabled Data Assimilation

The matrix approach makes it computationally feasible to assimilate multiple data sets to constrain the predictions of land carbon sequestration. Presently, land carbon dynamics simulated by ESMs deviate substantially from observation-based estimates. For example, global soil carbon storage simulated by 11 models participating in the Coupled Model Intercomparison Project phase 5 (CMIP5) fit poorly with the soil organic carbon stocks reported in the Harmonized World Soil Database (HWSD; Luo et al., 2015). Observational and experimental research produces ever-larger quantities of high-resolution, multi-dimensional data, offering great resources to constrain model prediction via data assimilation (Luo et al., 2011; Reichstein et al., 2019; Xia et al., 2020). In the past, flux-based data have been used to improve model predictive skills (Fox et al., 2018; MacBean et al., 2018; Rayner et al., 2005). While pool-based data are essential to constrain rate processes (Xu et al., 2006), their assimilation into ESMs is inhibited by unattainable computational expense.

The SASU method derived from the matrix approach makes it computationally feasible to assimilate both flux- and pool-based data to constrain predictions of land carbon sequestration from full dynamic models such as ESMs (Hararuk et al., 2014, 2015; Liang et al., 2018; Shi et al., 2018; Tao et al., 2020). When the matrix version of CLM3.5 was used in data assimilation to constrain its parameters, the model-data fit was improved to the extent that modeled soil carbon density accounted for 41% of variation in the HWSD database, in comparison to 27% without data assimilation (Hararuk et al., 2014). We have recently used the matrix version of the CLM5 soil carbon module in PROcess-guided deep learning and DAta-driven modeling (PRODA) to further improve predictions of soil carbon variability at regional and global scales. With PRODA, CLM5 can explain 62% of variation in soil organic carbon content in the conterminous USA (Tao & Luo, 2022; Tao et al., 2020) (Figure 5). Data assimilation can help not only improve model performance but also evaluate the relative contributions of model structure and parameterization, versus data, to model uncertainty (Shi et al., 2018). Such data assimilation work at regional or global scales has become possible thanks to the computational savings enabled by the accelerated spin-up facilitated by the matrix approach.
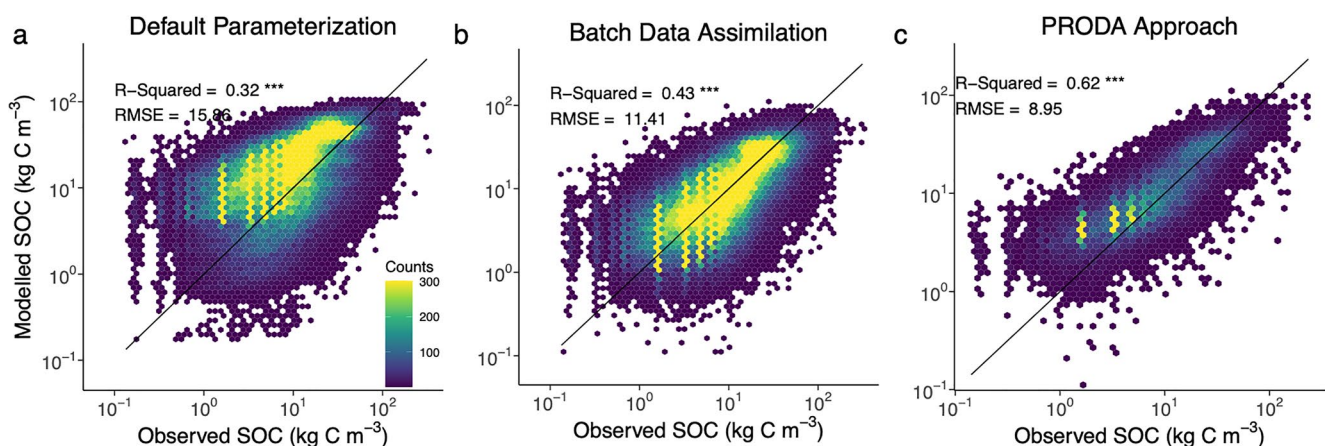

Figure 5. The agreement between observed and modeled SOC content for the conterminous USA with different approaches. CLM5 simulations of SOC with default parameters explain 32% of variation in 25,444 observed vertical profiles of SOC content at different depths over the conterminous USA (a). When CLM5 matrix model was used in data assimilation to constrain model parameters, 43% of variation in observations was explained (b). When CLM5 matrix model was trained by the PROcess-guided deep learning and DAta-driven modeling (PRODA) approach to constrain and predict spatially varying model parameters, explained variation rose to 62% (c) (Tao & Luo, 2022).

5. Concluding Remarks

We have shown that the matrix approach can unify land carbon cycle models. The unified matrix equation represents the dynamic disequilibrium of the land carbon cycle in terms of three overarching variables: carbon input, residence time, and carbon storage potential. Carbon input and residence time together quantify the carbon storage capacity at equilibrium, whereas the storage potential measures the magnitude of disequilibrium. The matrix equation folds internal processes and external variables into one framework to yield a system-level representation that facilitates understanding and analysis of the carbon cycle in all its complexity.

We have demonstrated the feasibility of converting different kinds of land carbon, nitrogen, and phosphorus cycle models to matrix equations. The result is simplicity in structure, high modularity in code, clarity in diagnostics, and computational efficiency in spin-up. Once expressed in the matrix form, all models regardless of complexity are organized in a standardized yet concise form that allows us to understand differences in structures and parameters among models, and to trace how these affect their outputs.

The matrix approach makes it possible to pinpoint sources of uncertainty in model predictions via traceability analysis. Pinpointing sources of model uncertainty across ensemble members of MIPs helps direct targeted efforts on model improvement. The fast spin-up enabled by the matrix approach makes it possible to capitalize on ever-increasing availability of empirical data by assimilating multiple datasets into complex models to improve prediction. The integration of big data with multiple models likely presents the best hope to accurately characterize the current state of the land carbon cycle and predict the future trajectory of terrestrial carbon sequestration.

Benefits of converting a traditional carbon cycle model to the matrix form might be limited if the original model is relatively simple and easily understood, spin-up is fast enough, and/or there is no need to calculate the new analytics offered by the matrix approach for understanding simulation outputs. Converting existing models, especially complex schemes used in current generation ESMs, to matrix models could be time-consuming and, thus, requires a considerable initial investment before benefits are gained. Although nonlinear microbial models can be expressed in the matrix form, it is not clear which advantages of the matrix approach would be gained. It is yet to be fully explored how best the matrix approach is applied to dynamic vegetation models. While the matrix approach empowers researchers to study the land carbon cycle, there are still many processes that are insufficiently understood or quantified, theoretically and empirically, and to be accurately represented in models. Ultimately, concerted collaboration among experimental, observational, and modeling research communities is a precondition for substantive progress. The matrix approach offers a toolbox of novel analytical and technical solutions that can accelerate this collaborative work towards an improved understanding of the land carbon cycle, its drivers, and its role within climate dynamics.

Data Availability Statement

No original data or code was used in this manuscript.

References

Ahlström, A., Xia, J., Arneth, A., Luo, Y., & Smith, B. (2015). Importance of vegetation dynamics for future terrestrial carbon cycling. *Environmental Research Letters*, *10*(5), 054019. https://doi.org/10.1088/1748-9326/10/5/054019

Cai, A., Liang, G., Zhang, X., Zhang, W., Li, L., Rui, Y., et al. (2018). Long-term straw decomposition in agro-ecosystems described by a unified three-exponentiation equation with thermal time. *Science of the Total Environment*, *636*, 699–708. https://doi.org/10.1016/j.scitotenv.2018.04.303

Canadell, J. G., Monteiro, P. M., Costa, M. H., Cotrim da Cunha, L., Cox, P. M., Eliseev, A., et al. (2021). Global carbon and other biogeochemical cycles and feedbacks. In V. Masson-Delmotte, P. Zhai, A. Pirani, S. L. Connors, C. Péan, S. Berger, et al. (Eds.), *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (pp. 673–816). Cambridge University Press. https://doi.org/10.1017/9781009157896.007

Caswell, H. (2001). *Matrix population models*. Sinauer Associates, Inc.

Ceballos-Núñez, V., Müller, M., & Sierra, C. A. (2020). Towards better representations of carbon allocation in vegetation: A conceptual framework and mathematical tool. *Theoretical Ecology*, *13*(3), 317–332. https://doi.org/10.1007/s12080-020-00455-w

Chen, J., Ju, W., Cihlar, J., Price, D., Liu, J., Chen, W., et al. (2003). Spatial distribution of carbon sources and sinks in Canada's forests. *Tellus*, *55B*, 622–641. https://doi.org/10.1034/j.1600-0889.2003.00036.x

Ciais, P., Sabine, C., Bala, G., Bopp, L., Brovkin, V., Canadell, J., et al. (2014). 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* (pp. 465–570). Cambridge University Press.

Collatz, G. J., Ribas-Carbo, M., & Berry, J. (1992). Coupled photosynthesis-stomatal conductance model for leaves of C_4 plants. *Functional Plant Biology*, *19*(5), 519–538. https://doi.org/10.1071/PP9920519

Crowther, T. W., Van den Hoogen, J., Wan, J., Mayes, M. A., Keiser, A., Mo, L., et al. (2019). The global soil community and its influence on biogeochemistry. *Science*, *365*(6455), eaav0550. https://doi.org/10.1126/science.aav0550

Cui, E., Huang, K., Arain, M. A., Fisher, J. B., Huntzinger, D. N., Ito, A., et al. (2019). Vegetation functional properties determine uncertainty of simulated ecosystem productivity: A traceability analysis in the East Asian monsoon region. *Global Biogeochemical Cycles*, *33*(6), 668–689. https://doi.org/10.1029/2018GB005909

De Kauwe, M. G., Medlyn, B. E., Zaehle, S., Walker, A. P., Dietze, M. C., Wang, Y. P., et al. (2014). Where does the carbon go? A model–data intercomparison of vegetation carbon allocation and turnover processes at two temperate forest free-air CO2 enrichment sites. *New Phytologist*, *203*(3), 883–899. https://doi.org/10.1111/nph.12847

Denman, K. L., Brasseur, G., Chidthaisong, A., Ciais, P., Cox, P. M., Dickinson, R. E., et al. (2007). Couplings between changes in the climate system and biogeochemistry. In *Climate Change 2007: The Physical Science Basis* (pp. 499–588). Cambridge University Press.

Du, Z. G., Wang, J. W., Zhou, G. Y., Bai, S. H., Zhou, L. Y., Fu, Y. L., et al. (2021). Differential effects of nitrogen vs. phosphorus limitation on terrestrial carbon storage in two subtropical forests: A Bayesian approach. *Science of The Total Environment*, *795*(2021) 148485. https://doi.org/10.1016/j.scitotenv.2021.148485

Farquhar, G. D., von Caemmerer, S. V., & Berry, J. A. (1980). A biochemical model of photosynthetic CO_2 assimilation in leaves of C_3 species. *Planta*, *149*(1), 78–90. https://doi.org/10.1007/BF00386231

Fisher, R. A., Muszala, S., Verteinstein, M., Lawrence, P., Xu, C., McDowell, N. G., et al. (2015). Taking off the training wheels: The properties of a dynamic vegetation model without climate envelopes, CLM4. 5. *Geoscientific Model Development*, *8*(11), 3593–3619. https://doi.org/10.5194/gmd-8-3593-2015

Fox, A. M., Hoar, T. J., Anderson, J. L., Arellano, A. F., Smith, W. K., Litvak, M. E., et al. (2018). Evaluation of a data assimilation system for land surface models using CLM4. 5. *Journal of Advances in Modeling Earth Systems*, *10*(10), 2471–2494. https://doi.org/10.1029/2018MS001362

Fry, E. L., De Long, J. R., Álvarez Garrido, L., Alvarez, N., Carrillo, Y., Castañeda-Gómez, L., et al. (2019). Using plant, microbe, and soil fauna traits to improve the predictive power of biogeochemical models. *Methods in Ecology and Evolution*, *10*(1), 146–157. https://doi.org/10.1111/2041-210X.13092

Gaudinski, J. B., Trumbore, S. E., Davidson, E. A., & Zheng, S. (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. https://doi.org/10.1023/A:1006301010014

Hararuk, O., Smith, M. J., & Luo, Y. (2015). Microbial models with data-driven parameters predict stronger soil carbon responses to climate change. *Global Change Biology*, *21*(6), 2439–2453. https://doi.org/10.1111/gcb.12827

Hararuk, O., Xia, J., & Luo, Y. (2014). Evaluation and improvement of a global land model against soil carbon data using a Bayesian Markov chain Monte Carlo method. *Journal of Geophysical Research: Biogeosciences*, *119*(3), 403–417. https://doi.org/10.1002/2013JG002535

Haynes, K. D., Baker, I. T., Denning, A. S., Stockli, R., Schaefer, K., Lokupitiya, E. Y., & Haynes, J. M. (2019). Representing ecosystems using dynamic prognostic phenology based on biological growth stages: Part 1. Implementation in the Simple Biosphere Model (SiB4). *Journal of Advances in Modeling Earth Systems*, *11*, 4423–4439. https://doi.org/10.1029/2018ms001540

Haynes, K. D., Baker, I. T., Denning, A. S., Wolf, S., Wohlfahrt, G., Kieley, G., et al. (2019). Representing ecosystems using dynamic prognostic phenology based on biological growth stages: Part 2. Carbon cycling. *Journal of Advances in Modeling Earth Systems*, *11*, 4440–4465. https://doi.org/10.1029/2018ms001541

Hou, E., Lu, X., Jiang, L., Wen, D., & Luo, Y. (2019). Quantifying soil phosphorus dynamics: A data assimilation approach. *Journal of Geophysical Research: Biogeosciences*, *124*(7), 2159–2173. https://doi.org/10.1029/2018jg004903

Hou, E., Luo, Y., Ma, S., Huang, Y., Zhou, Y., Kim, H.-S., et al. Spreading or shrinking of model uncertainty in simulating peatland carbon dynamics. Submitted to *Nature Communications*.

Huang, Y., Lu, X., Shi, Z., Lawrence, D., Koven, C. D., Xia, J., et al. (2018). Matrix approach to land carbon cycle modeling: A case study with the Community Land Model. *Global Change Biology*, *24*(3), 1394–1404. https://doi.org/10.1111/gcb.13948

Huang, Y., Zhu, D., Ciais, P., Guenet, B., Huang, Y., Goll, D. S., et al. (2018). Matrix-based sensitivity assessment of soil organic Carbon storage: A Case study from the ORCHIDEE-MICT Model. *Journal of Advances in Modeling Earth Systems*, *10*(8), 1790–1808. https://doi.org/10.1029/2017ms001237

Hudiburg, T., Law, B., Turner, D. P., Campbell, J., Donato, D., & Duane, M. (2009). Carbon dynamics of Oregon and Northern California forests and potential land-based carbon storage. *Ecological Applications*, *19*(1), 163–180. https://doi.org/10.1890/07-2006.1

Huntzinger, D. N., Michalak, A., Schwalm, C., Ciais, P., King, A., Fang, Y., et al. (2017). Uncertainty in the response of terrestrial carbon sink to environmental drivers undermines carbon-climate feedback predictions. *Scientific Reports*, *7*(1), 1–8. https://doi.org/10.1038/s41598-017-03818-2

Jiang, L., Liang, J., Lu, X., Hou, E., Hoffman, F. M., & Luo, Y. (2021). Country-level land carbon sink and its causing components by the middle of the twenty-first century. *Ecological Processes*, *10*(1), 1–13. https://doi.org/10.1186/s13717-021-00328-y

Jiang, L., Shi, Z., Xia, J., Liang, J., Lu, X., Wang, Y., & Luo, Y. (2017). Transient traceability analysis of Land Carbon storage dynamics: Procedures and its application to two forest ecosystems. *Journal of Advances in Modeling Earth Systems*, *9*(8), 2822–2835. https://doi.org/10.1002/2017ms001004

Koven, C. D., Knox, R. G., Fisher, R. A., Chambers, J. Q., Christoffersen, B. O., Davies, S. J., et al. (2020). Benchmarking and parameter sensitivity of physiological and vegetation dynamics using the functionally assembled terrestrial ecosystem simulator (FATES) at Barro Colorado Island, Panama. *Biogeosciences*, *17*(11), 3017–3044. https://doi.org/10.5194/bg-17-3017-2020

Lasaga, A. C. (1980). The kinetic treatment of geochemical cycles. *Geochimica et Cosmochimica Acta*, *44*, 815–828. https://doi.org/10.1016/0016-7037(80)90263-x

Lawrence, D. M., Fisher, R. A., Koven, C. D., Oleson, K. W., Swenson, S. C., Bonan, G., et al. (2019). The Community Land Model version 5: Description of new features, benchmarking, and impact of forcing uncertainty. *Journal of Advances in Modeling Earth Systems*, *11*(12), 4245–4287. https://doi.org/10.1029/2018ms001583

Lawrence, P. J., Lawrence, D. M., & Hurtt, G. C. (2018). Attributing the carbon cycle impacts of CMIP5 historical and future land use and land cover change in the Community Earth System Model (CESM1). *Journal of Geophysical Research: Biogeosciences*, *123*(5), 1732–1755. https://doi.org/10.1029/2017jg004348

Lei, L., Xia, J., Li, X., Huang, K., Zhang, A., Chen, S., et al. (2018). Water response of ecosystem respiration regulates future projection of net ecosystem productivity in a semiarid grassland. *Agricultural and Forest Meteorology*, *252*, 175–191. https://doi.org/10.1016/j.agrformet.2018.01.020

Liang, J., Xia, J., Shi, Z., Jiang, L., Ma, S., Lu, X., et al. (2018). Biotic responses buffer warming-induced soil organic carbon loss in Arctic tundra. *Global Change Biology*, *24*(10), 4946–4959. https://doi.org/10.1111/gcb.14325

Liao, C., Lu, X., Huang, Y., Lawrence, D., Kluzek, E., Oleson, K., et al. Accelerated spin-up of Community Land Model version 5 (CLM5) with coupled terrestrial carbon and nitrogen cycles. In Preperation.

Lu, X., Du, Z., Huang, Y., Lawrence, D., Kluzek, E., Collier, N., et al. (2020). Full Implementation of Matrix approach to biogeochemistry Module of CLM5. *Journal of Advances in Modeling Earth Systems*, *12*(11), e2020MS002105. https://doi.org/10.1029/2020ms002105

Luo, Y. (2022). Theoretical foundation of the land carbon cycle and matrix approach. In *Land Carbon Cycle Modeling: Matrix Approach, Data Assimilation, & Ecological Forecasting*. CPC Press, Taylor & Francis Group.

Luo, Y., Ahlström, A., Allison, S. D., Batjes, N. H., Brovkin, V., Carvalhais, N., et al. (2016). Toward more realistic projections of soil carbon dynamics by Earth system models. *Global Biogeochemical Cycles*, *30*(1), 40–56. https://doi.org/10.1002/2015gb005239

Luo, Y., Keenan, T. F., & Smith, M. (2015). Predictability of the terrestrial carbon cycle. *Global Change Biology*, *21*(5), 1737–1751. https://doi.org/10.1111/gcb.12766

Luo, Y., Ogle, K., Tucker, C., Fei, S., Gao, C., LaDeau, S., et al. (2011). Ecological forecasting and data assimilation in a data-rich era. *Ecological Applications*, *21*(5), 1429–1442. https://doi.org/10.1890/09-1275.1

Luo, Y., Randerson, J. T., Friedlingstein, P., Hibbard, K., Hoffman, F., Huntzinger, D., et al. (2012). A framework for benchmarking land models. *Biogeosciences*, *9*(10), 3857–3874. https://doi.org/10.5194/bg-9-3857-2012

Luo, Y., & Reynolds, J. F. (1999). Validity of extrapolating field CO_2 experiments to predict carbon sequestration in natural ecosystems. *Ecology*, *80*(5), 1568–1583. https://doi.org/10.1890/0012-9658(1999)080[1568:voefce]2.0.co;2

Luo, Y., & Schuur, E. A. (2020). Model parameterization to represent processes at unresolved scales and changing properties of evolving systems. *Global Change Biology*, *26*(3), 1109–1117. https://doi.org/10.1111/gcb.14939

Luo, Y., Shi, Z., Lu, X., Xia, J., Liang, J., Jiang, J., et al. (2017). Transient dynamics of terrestrial carbon storage: Mathematical foundation and its applications. *Biogeosciences*, *14*(1), 145. https://doi.org/10.5194/bg-14-145-2017

Luo, Y., & Weng, E. (2011). Dynamic disequilibrium of the terrestrial carbon cycle under global change. *Trends in Ecology & Evolution*, *26*(2), 96–104. https://doi.org/10.1016/j.tree.2010.11.003

Luo, Y., White, L. W., Canadell, J. G., DeLucia, E. H., Ellsworth, D. S., Finzi, A., et al. (2003). Sustainability of terrestrial carbon sequestration: A case study in Duke forest with inversion approach. *Global Biogeochemical Cycles*, *17*(1), 1021. https://doi.org/10.1029/2002gb001923

Luo, Y., Wu, L., Andrews, J. A., White, L., Matamala, R., Schäfer, K. V., & Schlesinger, W. H. (2001). Elevated CO_2 differentiates ecosystem carbon processes: Deconvolution analysis of Duke forest FACE data. *Ecological Monographs*, *71*(3), 357–376. https://doi.org/10.2307/3100064

MacBean, N., Maignan, F., Bacour, C., Lewis, P., Peylin, P., Guanter, L., et al. (2018). Strong constraint on modelled global carbon uptake using solar-induced chlorophyll fluorescence data. *Scientific Reports*, *8*(1), 1973. https://doi.org/10.1038/s41598-018-20024-w

Manzoni, S., & Porporato, A. (2009). Soil carbon and nitrogen mineralization: Theory and models across scales. *Soil Biology and Biochemistry*, *41*(7), 1355–1379. https://doi.org/10.1016/j.soilbio.2009.02.031

Metzler, H., Müller, M., & Sierra, C. A. (2018). Transit-time and age distributions for nonlinear time-dependent compartmental systems. *Proceedings of the National Academy of Sciences*, *115*(6), 1150–1155. https://doi.org/10.1073/pnas.1705296115

Metzler, H., Zhu, Q., Riley, W., Hoyt, A., Müller, M., & Sierra, C. A. (2020). Mathematical reconstruction of Land Carbon Models from their numerical output: Computing soil radiocarbon from C dynamics. *Journal of Advances in Modeling Earth Systems*, *12*(1), e2019MS001776. https://doi.org/10.1029/2019ms001776

Moorcroft, P. R., Hurtt, G. C., & Pacala, S. W. (2001). A method for scaling vegetation dynamics: The ecosystem demography model (ED). *Ecological Monographs*, *71*(4), 557–586. https://doi.org/10.1890/0012-9615(2001)071[0557:amfsvd]2.0.co;2

Oleson, K., Lawrence, D., Bonan, G., Flanner, M., & Kluzek, E. (2010). Technical description of version 4.0 of the Community Land Model (CLM) (No. NCAR/TN-478+STR). University Corporation for Atmospheric Research. https://doi.org/10.5065/D6FB50WZ

Olson, J. S. (1963). Energy storage and the balance of producers and decomposers in ecological systems. *Ecology*, *44*(2), 322–331. https://doi.org/10.2307/1932179

Potter, C. S., Randerson, J. T., Field, C. B., Matson, P. A., Vitousek, P. M., Mooney, H. A., & Klooster, S. A. (1993). Terrestrial ecosystem production: A process model based on global satellite and surface data. *Global Biogeochemical Cycles*, *7*(4), 811–841. https://doi.org/10.1029/93gb02725

Prentice, I. C., Farquhar, G., Fasham, M., Goulden, M., Heimann, M., Jaramillo, V., et al. (2001). The carbon cycle and atmospheric carbon dioxide. In J. T. Houghton, Y. Ding, D. J. Griggs, M. Noguer, P. J. V. D. Linden, X. Dai, et al. (Eds.), *Climate Change 2001: The Scientific Basis* (pp. 183–237). Cambridge University Press.

Prentice, I. C., Liang, X., Medlyn, B. E., & Wang, Y.-P. (2015). Reliable, robust and realistic: The three R's of next-generation land-surface modelling. *Atmospheric Chemistry and Physics*, *15*(10), 5987–6005. https://doi.org/10.5194/acp-15-5987-2015

Rafique, R., Xia, J., Hararuk, O., Asrar, G. R., Leng, G., Wang, Y., & Luo, Y. (2016). Divergent predictions of carbon storage between two global land models: Attribution of the causes through traceability analysis. *Earth System Dynamics (Online)*, *7*, 649–658. https://doi.org/10.5194/esd-7-649-2016

Rasmussen, M., Hastings, A., Smith, M. J., Agusto, F. B., Chen-Charpentier, B. M., Hoffman, F. M., et al. (2016). Transit times and mean ages for nonautonomous and autonomous compartmental systems. *Journal of Mathematical Biology*, *73*(6–7), 1379–1398. https://doi.org/10.1007/s00285-016-0990-8

Rayner, P. J., Scholze, M., Knorr, W., Kaminski, T., Giering, R., & Widmann, H. (2005). Two decades of terrestrial carbon fluxes from a carbon cycle data assimilation system (CCDAS). *Global Biogeochemical Cycles*, *19*(2), GB2026. https://doi.org/10.1029/2004gb002254

Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., & Carvalhais, N. (2019). Deep learning and process understanding for data-driven Earth system science. *Nature*, *566*(7743), 195. https://doi.org/10.1038/s41586-019-0912-1

Rogers, A., Medlyn, B. E., Dukes, J. S., Bonan, G., Von Caemmerer, S., Dietze, M. C., et al. (2017). A roadmap for improving the representation of photosynthesis in Earth system models. *New Phytologist*, *213*(1), 22–42. https://doi.org/10.1111/nph.14283

Running, S. W., Nemani, R. R., Heinsch, F. A., Zhao, M., Reeves, M., & Hashimoto, H. (2004). A continuous satellite-derived measure of global terrestrial primary production. *BioScience*, *54*(6), 547–560. https://doi.org/10.1641/0006-3568(2004)054[0547:acsmog]2.0.co;2

Schädel, C., Luo, Y., Evans, R. D., Fei, S., & Schaeffer, S. M. (2013). Separating soil CO_2 efflux into C-pool-specific decay rates via inverse analysis of soil incubation data. *Oecologia*, *171*(3), 721–732. https://doi.org/10.1007/s00442-012-2577-4

Schädel, C., Schuur, E. A., Bracho, R., Elberling, B., Knoblauch, C., Lee, H., et al. (2014). Circumpolar assessment of permafrost C quality and its vulnerability over time using long-term incubation data. *Global Change Biology*, *20*(2), 641–652. https://doi.org/10.1111/gcb.12417

Schaefer, K., Collatz, G. J., Tans, P., Denning, A. S., Baker, I., Berry, J., et al. (2008). Combined simple biosphere/Carnegie-Ames-Stanford approach terrestrial carbon cycle model. *Journal of Geophysical Research*, *113*, G03034. https://doi.org/10.1029/2007jg000603

Shi, Z., Allison, S. D., He, Y., Levine, P. A., Hoyt, A. M., Beem-Miller, J., et al. (2020). The age distribution of global soil carbon inferred from radiocarbon measurements. *Nature Geoscience*, *13*(8), 555–559. https://doi.org/10.1038/s41561-020-0596-z

Shi, Z., Crowell, S., Luo, Y., & Moore, B. (2018). Model structures amplify uncertainty in predicted soil carbon responses to climate change. *Nature Communications*, *9*(1), 2171. https://doi.org/10.1038/s41467-018-04526-9

Shi, Z., Yang, Y. H., Zhou, X. H., Weng, E. S., Finzi, A. C., & Luo, Y. Q. (2016). Inverse analysis of coupled carbon-nitrogen cycles against multiple datasets at ambient and elevated CO_2. *Journal of Plant Ecology*, *9*(3), 285–295. https://doi.org/10.1093/jpe/rtv059

Sierra, C. A., Ceballos-Núñez, V., Metzler, H., & Müller, M. (2018). Representing and understanding the carbon cycle using the theory of compartmental dynamical systems. *Journal of Advances in Modeling Earth Systems*, *10*(8), 1729–1734. https://doi.org/10.1029/2018ms001360

Sierra, C. A., Hoyt, A. M., He, Y., & Trumbore, S. E. (2018). Soil organic matter persistence as a stochastic process: Age and transit time distributions of carbon in soils. *Global Biogeochemical Cycles*, *32*(10), 1574–1588. https://doi.org/10.1029/2018gb005950

Sierra, C. A., & Müller, M. (2015). A general mathematical framework for representing soil organic matter dynamics. *Ecological Monographs*, *85*(4), 505–524. https://doi.org/10.1890/15-0361.1

Sierra, C. A., Trumbore, S., Davidson, E., Frey, S., Savage, K., & Hopkins, F. (2012). Predicting decadal trends and transient responses of radiocarbon storage and fluxes in a temperate forest soil. *Biogeosciences*, *9*(8), 3013–3028. https://doi.org/10.5194/bg-9-3013-2012

Sierra, C. A., Trumbore, S. E., Davidson, E. A., Vicca, S., & Janssens, I. (2015). Sensitivity of decomposition rates of soil organic matter with respect to simultaneous changes in temperature and moisture. *Journal of Advances in Modeling Earth Systems*, *7*(1), 335–356. https://doi.org/10.1002/2014ms000358

Sitch, S., Smith, B., Prentice, I. C., Arneth, A., Bondeau, A., Cramer, W., et al. (2003). Evaluation of ecosystem dynamics, plant geography and terrestrial carbon cycling in the LPJ dynamic global vegetation model. *Global Change Biology*, *9*(2), 161–185. https://doi.org/10.1046/j.1365-2486.2003.00569.x

Smith, B., Prentice, I. C., & Sykes, M. T. (2001). Representation of vegetation dynamics in the modelling of terrestrial ecosystems: Comparing two contrasting approaches within European climate space. *Global Ecology and Biogeography*, *10*, 621–637. https://doi.org/10.1046/j.1466-822x.2001.t01-1-00256.x

Southam, J. R., & Hay, W. W. (1976). Dynamical formulation of Broecker's model for marine cycles of biologically incorporated elements. *Mathematical Geology*, *8*, 511–527. https://doi.org/10.1007/bf01042991

Sundquist, E. T. (1985). Geological perspectives on carbon dioxide and the carbon cycle. *The Carbon Cycle and Atmospheric CO₂: Natural Variations Archean to Present*, *32*, 55–59.

Tao, F., & Luo, Y. (2022). PROcess-guided deep learning and DAta-driven modelling (PRODA). In Y. Luo & B. Smith (Eds.), *Land carbon cycle modeling: Matrix approach, data assimilation, and ecological forecasting*. Taylor and Francis.

Tao, F., Zhou, Z., Huang, Y., Li, Q., Lu, X., Ma, S., et al. (2020). Deep learning optimizes data-driven representation of soil organic carbon in Earth system model over the conterminous United States. *Frontiers in Big Data*, *3*, 17. https://doi.org/10.3389/fdata.2020.00017

Thomas, F. M., Yen, J. D., & Vesk, P. A. (2019). Using functional traits to predict species growth trajectories, and cross-validation to evaluate these models for ecological prediction. *Ecology and Evolution*, *9*(4), 1554–1566. https://doi.org/10.1002/ece3.4693

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

Trumbore, S. E., Chadwick, O. A., & Amundson, R. (1996). Rapid exchange between soil carbon and atmospheric carbon dioxide driven by temperature change. *Science*, *272*(5260), 393–396. https://doi.org/10.1126/science.272.5260.393

Tuljapurkar, S., & Caswell, H. (1997). *Structured-population models in marine, terrestrial, and freshwater systems* (Vol. 18). Springer Science & Business Media.

Wang, S., Luo, Y., & Niu, S. (2022). Reparameterization required after model structure changes from carbon only to carbon-nitrogen coupling. *Journal of Advances in Modeling Earth Systems*, *14*, e2021MS002798. https://doi.org/10.1029/2021ms002798

Wang, Y., Chen, B., Wieder, W. R., Leite, M., Medlyn, B. E., Rasmussen, M., et al. (2014). Oscillatory behavior of two nonlinear microbial models of soil carbon decomposition. *Biogeosciences*, *11*(7), 1817–1831. https://doi.org/10.5194/bg-11-1817-2014

Wang, Y., Jiang, J., Chen-Charpentier, B. M., Agusto, F. B., Hastings, A., Hoffman, F. M., et al. (2016). Responses of two nonlinear microbial models to warming and increased carbon input. *Biogeosciences*, *13*(4), 887–902. https://doi.org/10.5194/bg-13-887-2016

Wang, Y., Law, R., & Pak, B. (2010). A global model of carbon, nitrogen and phosphorus cycles for the terrestrial biosphere. *Biogeosciences*, *7*(7), 2261–2282. https://doi.org/10.5194/bg-7-2261-2010

White, L. W., Luo, Y., & Xu, T. (2005). Carbon sequestration: Inversion of FACE data and prediction. *Applied Mathematics and Computation*, *163*(2), 783–800. https://doi.org/10.1016/j.amc.2004.04.021

Wieder, W., Grandy, A., Kallenbach, C., Taylor, P., & Bonan, G. (2015). Representing life in the Earth system with soil microbial functional traits in the MIMICS model. *Geoscientific Model Development Discussions*, *8*(6), 1789–1808. https://doi.org/10.5194/gmd-8-1789-2015

Wu, D., Piao, S., Zhu, D., Wang, X., Ciais, P., Bastos, A., et al. (2020). Accelerated terrestrial ecosystem carbon turnover and its drivers. *Global Change Biology*, *26*(9), 5052–5062. https://doi.org/10.1111/gcb.15224

Xia, J., Luo, Y., Wang, Y. P., & Hararuk, O. (2013). Traceable components of terrestrial carbon storage capacity in biogeochemical models. *Global Change Biology*, *19*(7), 2104–2116. https://doi.org/10.1111/gcb.12172

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

Xia, J., Niu, S., Ciais, P., Janssens, I. A., Chen, J., Ammann, C., et al. (2015). Joint control of terrestrial gross primary productivity by plant phenology and physiology. *Proceedings of the National Academy of Sciences*, *112*(9), 2788–2793. https://doi.org/10.1073/pnas.1413090112

Xia, J., Wang, J., & Niu, S. (2020). Research challenges and opportunities for using big data in global change biology. *Global Change Biology*, *26*(11), 6040–6061. https://doi.org/10.1111/gcb.15317

Xia, J., Yuan, W., Wang, Y.-P., & Zhang, Q. (2017). Adaptive carbon allocation by plants enhances the terrestrial carbon sink. *Scientific Reports*, *7*(1), 1–11. https://doi.org/10.1038/s41598-017-03574-3

Xu, T., White, L., Hui, D., & Luo, Y. (2006). Probabilistic inversion of a terrestrial ecosystem model: Analysis of uncertainty in parameter estimation and model prediction. *Global Biogeochemical Cycles*, *20*(2). https://doi.org/10.1029/2005gb002468

Xu, X., Shi, Z., Li, D., Rey, A., Ruan, H., Craine, J. M., et al. (2016). Soil properties control decomposition of soil organic carbon: Results from data-assimilation analysis. *Geoderma*, *262*, 235–242. https://doi.org/10.1016/j.geoderma.2015.08.038

Yang, Y., Luo, Y., & Finzi, A. C. (2011). Carbon and nitrogen dynamics during forest stand development: A global synthesis. *New Phytologist*, *190*(4), 977–989. https://doi.org/10.1111/j.1469-8137.2011.03645.x

Zhang, D., Hui, D., Luo, Y., & Zhou, G. (2008). Rates of litter decomposition in terrestrial ecosystems: Global patterns and controlling factors. *Journal of Plant Ecology*, *1*(2), 85–93. https://doi.org/10.1093/jpe/rtn002

Zhou, S., Liang, J., Lu, X., Li, Q., Jiang, L., Zhang, Y., et al. (2018). Sources of uncertainty in modeled land carbon storage within and across three MIPs: Diagnosis with three new techniques. *Journal of Climate*, *31*(7), 2833–2851. https://doi.org/10.1175/jcli-d-17-0357.1

Zhou, T., Shi, P., Jia, G., Dai, Y., Zhao, X., Shangguan, W., et al. (2015). Age-dependent forest carbon sink: Estimation via inverse modeling. *Journal of Geophysical Research: Biogeosciences*, *120*(12), 2473–2492. https://doi.org/10.1002/2015jg002943

Zhu, K., Zhang, J., Niu, S., Chu, C., & Luo, Y. (2018). Limits to growth of forest biomass carbon sink under climate change. *Nature Communications*, *9*(1), 1–8. https://doi.org/10.1038/s41467-018-05132-5

Zhu, Q., Riley, W. J., Tang, J., Collier, N., Hoffman, F. M., Yang, X., & Bisht, G. (2019). Representing nitrogen, phosphorus, and carbon interactions in the E3SM land model: Development and global benchmarking. *Journal of Advances in Modeling Earth Systems*, *11*(7), 2238–2258. https://doi.org/10.1029/2018ms001571