Fall 2021

Merging Ecology and Earth System Modeling: Biotic and Abiotic Controls Drive Soil Carbon and Nitrogen Cycling in Microbial-Explicit Soil Biogeochemistry Models

Emily Kyker-Snowman
University of New Hampshire, Durham

Recommended Citation

Kyker-Snowman, Emily, "Merging Ecology and Earth System Modeling: Biotic and Abiotic Controls Drive Soil Carbon and Nitrogen Cycling in Microbial-Explicit Soil Biogeochemistry Models" (2021). *Doctoral Dissertations*. 2624.
https://scholars.unh.edu/dissertation/2624

MERGING ECOLOGY AND EARTH SYSTEM MODELING:
BIOTIC AND ABIOTIC CONTROLS DRIVE SOIL CARBON AND NITROGEN CYCLING IN MICROBIAL-EXPLICIT SOIL BIOGEOCHEMISTRY MODELS

BY

Emily Kyker-Snowman

Biology and Engineering, A.B., Dartmouth College, 2011
Environmental Engineering, B.E., Dartmouth College, 2012

DISSERTATION

Submitted to the University of New Hampshire
in Partial Fulfillment of
the Requirements for the Degree of

Doctor of Philosophy
in
Earth and Environmental Science

September, 2021

This dissertation has been examined and approved in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Earth and Environmental Science by:

Dissertation Director, Dr. A. Stuart Grandy, Associate Professor, Department of Natural Resources and the Environment, UNH

Dr. Gordon B. Bonan, Senior Scientist, National Center for Atmospheric Research, Climate and Global Dynamics, Boulder, CO

Dr. Serita D. Frey, Professor, Department of Natural Resources and the Environment, UNH

Dr. Stephen E. Frolking, Research Professor, Institute for the Study of Earth, Oceans and Space and Department of Earth Sciences, UNH

Dr. William R. Wieder, Project Scientist II, National Center for Atmospheric Research, Climate and Global Dynamics, Boulder, CO

On July 20th, 2021

DEDICATION

To my father, Thom Snowman, who has been editing my papers since they were about dolphins and banana pancakes.

ACKNOWLEDGEMENTS

I am deeply grateful to my advisor, Dr. A. Stuart Grandy, for his unfailing commitment to his students' wellbeing and the progress of their careers. Through his thoughtful consideration of individual mentoring needs, enthusiastic support of research activities both inside and out of his expertise, long meandering conversations on far-reaching scientific ideas, dedication to connecting his students to his widespread network of brilliant scientific minds, and recognition of the often stressful human experience of graduate school, he has defined "good mentorship" for me. I feel profoundly lucky to have spent my graduate years in the welcoming and supportive lab community he built. I am also grateful to my committee members, Drs. Will Wieder, Serita Frey, Gordon Bonan, and Steve Frolking, for their careful tutoring, frequent input, shared opportunities and ideas, and good-humored support over the years. I offer my thanks and appreciation to the members of the Grandy and Frey labs, who have been a family to me for the better part of a decade and continue to be close friends. Finally, I owe a great deal to the family and friends who have supported me, listened to me ramble about my research, and kept me in good spirits throughout my time at UNH, especially my father Thom Snowman and my sister Kelly Kyker-Snowman.

Financial support for this dissertation has been provided by the USDA National Institute of Food and Agriculture (Project No. 2015-35615-22747), the US Department of Energy (Grant Number DE-SC0016590), a NSF Graduate Research Fellowship under Grant No. DGE-1450271, and the New Hampshire Agricultural Experiment Station.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I

CHAPTER II

CHAPTER III

LIST OF FIGURES

CHAPTER 1

CHAPTER II

CHAPTER III

ABSTRACT

Merging Ecology and Earth System Modeling:
Biotic and Abiotic Controls Drive Soil Carbon and Nitrogen Cycling in Microbial-Explicit Soil Biogeochemistry Models

By

Emily Kyker-Snowman
University of New Hampshire

Ecological processes drive terrestrial biogeochemistry, yet the incorporation of ecology into the Earth system models that we use to understand and project global change remains. My dissertation focuses on expediting the incorporation of ecology into Earth system models, first by laying out a roadmap from initial assessment of ecological insights to eventual ESM incorporation, and then by demonstrating this roadmap using the example of microbially-controlled carbon and nitrogen cycling in soil. The paradigm around SOM formation and loss has shifted in recent decades away from a focus on the chemically recalcitrant leftovers of litter decomposition and towards a paradigm with microbial residues and mineral interactions at its heart. The MIcrobial-MIneral Carbon Stabilization model (MIMICS) was developed as a way of exploring this new paradigm and examining the relationships between environmental drivers, litter chemistry, microbial physiology, and physical and chemical stabilization mechanisms for SOM. In the first chapter of my dissertation, I document a systematic approach to improve ecological process representation in Earth system models, highlighting multiple points along the way where ecological observations and modeling iteratively strengthen one another. In the second chapter, I develop and validate a new version of MIMICS with coupled N cycling using a

large litter decomposition dataset. In the final chapter, I examine MIMICS-CN's representation of the drivers of SOM C:N ratios using a landscape-scale data synthesis and model-data comparison. Together, these chapters describe and demonstrate the process of improving biogeochemical models along the path to ESMs by introducing new process representations of ecological concepts.

INTRODUCTION

The Earth system is tremendously complex, and models ranging from simple process representations to full Earth system models (ESMs) are critical tools in aiding our understanding of global environmental change. Models allow scientists to explore interactions and scales that are impossible to study through direct measurement and to project insights into the workings of the Earth system into the future. Modern ESMs primarily originated in the atmospheric and physical science communities (Edwards, 2011), and the incorporation of ecological knowledge into ESMs by comparison has been relatively slow, despite the crucial role that ecology plays in controlling many ecosystem processes (Fer et al., 2021; Rogers et al., 2017). The empirical ecology and Earth system modeling communities have been slow to integrate due to a lack of a shared understanding of the steps involved in developing ecological theory and integrating it into large-scale models. In this dissertation, I strive to strengthen ecologist-modeler connections by developing and demonstrating a pathway for incorporating new ecological insights (in this case, insights into microbial control over soil carbon and nitrogen cycling) into large-scale models.

Soils within the Earth system

Soils are a vast global store of carbon (C) and nitrogen (N), holding roughly three times as much C as the atmosphere (Lal, 2004). Soils are also one of the most promising leverage points in the fight against climate change: policy efforts like the “4 per Mille Initiative” strive to combat climate change by altering agriculture and land management to sequester atmospheric C in soil and reduce environmental N pollution while maintaining high productivity (Minasny et al., 2017). However, meeting such lofty goals requires more reliable understanding and control over soil C and N cycling than currently exists. Soils remain one of the dominant sources of uncertainty in the C cycle component of Earth System Models (ESMs; Ito et al., 2019; Todd-Brown et al., 2013). The influence of climate and management drivers on soil C and N cycling has been studied in great depth and breadth around the world, yet measured soil C and N cycling often deviates from patterns predicted by established theory (Lehmann and Kleber, 2015). Further, agricultural management impacts on soil are highly site-dependent, and the advice given to farmers for building and maintaining healthy soils that promote C accumulation is constantly evolving (Dignac et al., 2017). Given the physical, chemical, and biological complexity of soil, one approach that may help scientists to guide policy efforts is the use of numerical models.

Historical soil models

Models are a means for scientists to organize their understanding of the natural world and ask if individual theories make sense when integrated into a larger whole. Models provide a useful testbed for questions related to scaling: mathematical relationships developed in the lab can be applied at much larger scales, and emergent model behavior can be compared against measured patterns. In the case of soils, models are also a necessity because pools of C and N in soil often change too slowly to observe in short-term experiments, and long-term experiments are too expensive to maintain at more than a handful of locations (Bailey et al., 2018).

The models that form the basis for the representation of soils in modern ESMs were developed in the 80s and 90s and will be referred to here as “conventional,” “traditional” or “historical” models of soil. Conventional soil models including CENTURY (Parton and Rasmussen, 1994) and RothC (Jenkinson and Rayner, 1977) assume first-order linear decay dynamics of litter and SOM, meaning that a fixed proportion of the model’s litter pool decomposes each time step, with some fraction entering soil organic matter (SOM) while the rest is lost to respiration according to a model-defined microbial carbon use efficiency (CUE). These models also generally describe soils as a hierarchical cascade of successively more recalcitrant pools with longer turnover times. This representation is based on the long-held assumption that chemical recalcitrance protects organic matter from decomposition and determines the longevity of organic matter in soils (Todd-Brown et al., 2012). This mathematical representation is popular in ESMs because of its relative simplicity and ability to adequately reproduce patterns from long-term experiments in soils (Blankinship et al., 2018).

Early contributors to models like CENTURY and RothC realized that soil models would be used to make policy decisions that would impact large numbers of people and invested considerable effort into ensuring those models were trustworthy through validation exercises (Powlson et al., 1996; Smith et al., 1997). As a result of their efforts, there is a robust literature evaluating hierarchical linear models of soil organic matter against data from input manipulation and fractionation studies (Powlson et al., 1996), as well as robust theoretical and mathematical analysis of such models (Manzoni and Porporato, 2009). Although CENTURY and RothC were both originally developed for a single site, subsequent studies explored their application across many other systems, at global scales, in managed and natural systems, and for simulating both standing soil C stocks and dynamic responses of soil pools to perturbations. This widespread application and evaluation of linear models of soil has contributed to their popularity and the longevity of their use within ESMs (Blankinship et al., 2018).

Despite the robust testing that models like CENTURY and RothC have undergone, they still frequently fail to accurately predict landscape-scale soil C stocks (Bradford et al., 2016). Research in the past few decades has revealed natural patterns like priming and C saturation that linear models cannot replicate, and new methods for characterizing the chemistry and microbiology of soils have upended the foundational assumptions behind traditional linear models (Blankinship et al., 2018). New research suggests that the longevity of organic matter in soils results more often from physical protection than chemical recalcitrance (Grandy and Neff, 2008; Heckman et al., 2013; Kallenbach et al., 2016; Rasmussen et al., 2018; Schimel and Schaeffer, 2012) and that the bulk of SOM is derived from microbial residues and necromass rather than plant matter (Gleixner, 2013; Kallenbach et al., 2016). Given this new understanding and the shortcomings of traditional models, a new generation of soil models has been developed that centers on microbial physiology and substrate accessibility as drivers of soil C and N cycling.

Next generation microbial-explicit models

Microbial-explicit models of soil include a wide variety of purely theoretical models that have examined everything from temperature acclimation of community-scale CUE (Allison, 2014) to the impact of microbial community structure on C cycling (Moorhead and Sinsabaugh, 2006). (Kaiser et al., 2014) used an individual-based model that represented metabolic linkages between microbial functional types (plant degraders, necromass degraders, or opportunists) in a grid and evaluated the succession of microbial community dynamics over the course of litter decay. Kaiser et al. provided evidence that intra-community recycling of N in necromass can alleviate N constraints and decompose plant litter without the need to decrease CUE. (Sistla et al., 2014) developed the SCAMPS model to represent C and N cycling in an arctic tundra under warming, and used a single microbial pool with a flexible C:N to simulate the acclimation of the microbial community to different nutrient conditions. (Averill and Waring, 2017) introduced the Carbon, Acidity, and Mineral Protection model (CAMP) to examine coupled C-N cycling during microbial processing of mineral-associated and particulate organic matter under an external constraint on microbial activity (acidity). Theoretical explorations of litter-microbe-soil interactions using models are diverse and abundant (Louis et al., 2016), athough large-scale application and evaluation of such models is rare.

A much smaller number of models has been tested against landscape-scale data gathered outside the lab. The Carbon, Organisms, Rhizosphere, and Protection in the Soil Environment model (CORPSE; (Sulman et al., 2014, 2017) couples belowground C and N through microbial biomass similar to MIMICS' representation, and has recently been coupled to the Fixation and Uptake of Nitrogen model (FUN, Sulman et al., 2017) to incorporate the exchange of plant C for soil N via mycorrhizal partners. The Millennial (Abramoff et al., 2017), TRIPLEX-GHG (Wang et al., 2017b), and ORCHIMIC (Huang et al., 2018b) models have all been developed to integrate explicit microbial decomposition into existing models designed for large-scale applications. Model-data comparisons with ORCHIMIC, TRIPLEX-GHG, and a number of other unnamed landscape-scale models (Hararuk et al., 2015; Li et al., 2014) have demonstrated that microbial-explicit decomposition formulations can reproduce global SOC stocks with greater accuracy than linear models like CENTURY.

The MIcrobial-MIneral Carbon Stabilization model

My dissertation focuses on the MIcrobial-MIneral Carbon Stabilization model (MIMICS; Wieder et al., 2014, 2015d, 2018), a relatively simple and computationally efficient microbial-explicit model that incorporates both physical and chemical mechanisms for stabilizing SOM. MIMICS has been validated previously with several continent- and global-scale datasets. In (Wieder et al., 2015d), MIMICS simulated C losses from litterbags in the Long-Term Intersite

Decomposition Experiment Team dataset (LIDET, a 10-year long litterbag experiment across 28 sites across North America; Parton et al., 2007) with a lower root mean square error (RMSE) than either DAYCENT or CLM-cn. MIMICS also simulated global distributions of total soil C in the Harmonized World Soil Database with higher spatial correlation and lower RMSE than CLM or any other model in the Coupled Model Intercomparison Project phase 5 archive (CMIP5) (Wieder et al., 2015d).

MIMICS was initially only a model of soil C cycling, but the second chapter of my dissertation covers the development of a version with coupled N cycling (Figure 1). MIMICS is a microbially-explicit model that incorporates 8 pools: two litter pools with different C:N stoichiometry, two microbial pools with different resource use strategies, two SOM pools with different stabilization mechanisms, and pools for inorganic N and "available" SOM (SOM available for microbial uptake). MIMICS partitions litter inputs to soil into metabolic and structural pools, where metabolic litter has a lower C:N and higher rates of microbial uptake. The model includes a pool of copiotrophic "r strategist" microbes that use C and N quickly and inefficiently, and a pool of oligotrophic "K strategist" microbes that have higher CUE and C:N (Fierer et al., 2007). Microbes in MIMICS take up available substrates via biomass-based Michaelis-Menten kinetics (German et al., 2012). This nonlinear mechanism allows for the "priming effect" whereby fresh, labile inputs to soil stimulate microbial activity and can result in a net loss of C from the soil system rather than a net increase in C storage. The combination of the two litter pools and two microbial pools allows MIMICS to explore tradeoffs between litter input chemistry and microbial use of those substrates. One of my alterations to the model in Chapter 2 was to introduce density-dependent microbial turnover (Georgiou et al., 2017) to dampen unrealistic oscillations in microbial biomass in response to perturbations and strike a balance between the input-insensitive equilibrium SOC of previous microbial models (Wang et al., 2014b) and the ability to replicate priming dynamics absent from linear models. Finally, MIMICS incorporates two mechanisms for organic matter stabilization in soil: chemical stabilization (i.e. recalcitrance) and physical stabilization (i.e. aggregation, sorption to mineral particles). Chemically stabilized SOM can be returned to the "available" pool via biomass-controlled Michaelis-Menten oxidation. Physically protected organic matter becomes available to microbes at a rate governed by soil clay content but independent of microbial biomass.

MIMICS and models like it may eventually replace traditional soil models in ESMs. However, just like traditional models, they must undergo rigorous testing to ensure that model projections are realistic and trustworthy. MIMICS integrates a number of new micro-scale theories that are difficult to verify at landscape scales, so validation exercises with MIMICS must explore both the underlying mechanisms and the emergent patterns in the model. Validating MIMICS against measured pools and fluxes will also help point to specific soil processes where the model is succeeding or failing. Whether or not MIMICS can reproduce measurements in soils with high fidelity, comparing modeled and measured results will guide future model development. My dissertation advances the development of MIMICS via rigorous testing of the model's equilibrium and dynamic behavior across litter, microbial, and SOM pools.

Dissertation overview

Each of the chapters of my dissertation examines a different component aspect of the journey to incorporating a new ecological concept into an Earth system model (describing the process, initial development of a simple model of a new ecological process, and evaluation of the new model using large-scale data), and Chapters 2 and 3 each use a different source of validation data to interrogate the model. In Chapter 1 (currently in review at Global Change Biology), I describe a complete workflow for incorporating a new ecological process into an ESM. The workflow has three stages: initial process evaluation to assess if a process is ready for modeling, simple model development and evaluation, and finally model development and testing within the context of an ESM.

In Chapter 2 (which is published in Geoscientific Model Development), I incorporated N into the C-only version of the model developed in (Wieder et al., 2014) and evaluated the new model's ability to replicate litter decomposition dynamics using the Long-term Inter-site Decomposition Experiment (LIDET; Parton et al., 2007). The model simulates C and N losses from litterbags in the LIDET study (6 litter types, 10 years of observations, 14 sites across North America) with reasonable accuracy (C: R^2=0.63; N: R^2=0.27). MIMICS-CN simulations of litterbag N dynamics are better than Community Land Model (CLM-cn) simulations of the same data and as good or better than DAYCENT simulations. Across the 14 simulated LIDET sites, MIMICS-CN produces reasonable equilibrium values for total soil C and N, microbial biomass C and N, respiration, inorganic N, and N mineralization rate. Chapter 2 serves as both the initial introduction of the coupled C-N version of the model and as a validation exercise for the representation of litter decomposition.

Chapter 3 (currently in preparation for submission to Biogeochemistry) examines the drivers of soil C:N stoichiometry in models and real soils. Here, I compare a novel global database of soil measurements (the Soils Data Harmonization or SoDaH database) against global soil simulations using the MIMICS-CN and CASA-CNP models. In parallel with statistical analysis of the SoDaH database, I compare the drivers of soil C:N in MIMICS-CN and CASA-CNP against the drivers discerned from the measured database. The comparison highlighted differences between models in which drivers were most important due to differences in underlying model structure. In addition, data-model comparisons emphasized the importance of clay content in driving soil stoichiometry, and analysis of the SoDaH database alone pointed to pH and specific soil mineralogy as potentially important controls on soil stoichiometry that are missing from models.

CHAPTER I: INCREASING THE SPATIAL AND TEMPORAL IMPACT OF ECOLOGICAL RESEARCH: A ROADMAP FOR INTEGRATING A NOVEL TERRESTRIAL PROCESS INTO AN EARTH SYSTEM MODEL

Under review at Global Change Biology

Authored by Emily Kyker-Snowman, Danica L. Lombardozzi, Gordon B. Bonan, Susan J. Cheng, Jeffrey S. Dukes, Serita D. Frey, Elin M. Jacobs, Risa McNellis, Joshua M. Rady, Nicholas G. Smith, R. Quinn Thomas, William R. Wieder, and A. Stuart Grandy

Abstract

Terrestrial ecosystems regulate Earth's climate through water, energy, and biogeochemical transformations. Despite a key role in regulating the Earth system, terrestrial ecology has historically been underrepresented in the Earth system models (ESMs) that are used to understand and project global environmental change. Ecology and Earth system modeling must be integrated for scientists to fully comprehend the role of ecological systems in driving and responding to global change. Ecological insights can improve ESM realism and reduce process uncertainty, while ESMs offer ecologists an opportunity to broadly test ecological theory and increase the impact of their work by scaling concepts through time and space. Despite this mutualism, meaningfully integrating the two remains a persistent challenge, in part because of logistical obstacles in translating processes into mathematical formulas and identifying ways to integrate new theories and code into large, complex model structures. To help overcome this interdisciplinary challenge, we present a framework consisting of a series of interconnected stages for integrating a new ecological process or insight into an ESM. First, we highlight the multiple ways that ecological observations and modeling iteratively strengthen one another, dispelling the illusion that the ecologist's role ends with initial provision of data. Second, we show that many valuable insights, products, and theoretical developments are produced through sustained interdisciplinary collaborations between empiricists and modelers, regardless of eventual inclusion of a process in an ESM. Finally, we provide concrete actions and resources to facilitate learning and collaboration at every stage of data-model integration. This framework will create synergies that will transform our understanding of ecology within the Earth system, ultimately improving our understanding of global environmental change and broadening the impact of ecological research.

1. The need to integrate ecology and Earth system models

Terrestrial ecosystems are an integral component of the Earth system. They govern the exchange of energy, water, and greenhouse gases between Earth's land surface and atmosphere and provide numerous services for society, including climate regulation and mitigation. For example, terrestrial ecosystems absorb approximately a third of anthropogenic carbon emissions (Friedlingstein et al., 2019), mitigating the impact of these emissions on climate change. They also play an essential role in regulating global water fluxes, from moderating soil water availability to influencing precipitation patterns and evaporative cooling. The physical properties of terrestrial ecosystems, including their surface reflectivity (i.e., albedo) and surface roughness, also help control the amount of energy absorbed and released by the land surface (Bonan, 2008, 2016). Human management of terrestrial ecosystems can change these biosphere-atmosphere interactions, for example by reducing carbon storage through deforestation and increasing greenhouse gas emissions through agricultural fertilization (Lade et al., 2019; Law et al., 2018). Given the importance of terrestrial ecosystems within the Earth system, modern ecological research papers frequently recommend updating existing ESMs to reflect new evidence or ideas about ecology that may have large-scale impacts on climate. This integration, however, has been slow (Fisher and Koven, 2020).

Historically, integration of ecological insights into ESMs has been hampered because of a disconnect between the scientists conducting empirical research and those engaging in modeling work (Fig. 1), a lack of cross-disciplinary training in modeling and empirical skills, and undervaluing of insights derived from modeling and data exercises completed along the way to incorporating an ecological process into an ESM. Although many scientists engage in both empirical and modeling work, the prevailing paradigm for integrating ecology into models tends to separate the tasks involved into the subdisciplines of empirical data collection and model development (Figs. 1, 2). Even when ecologists engage with model development, the models used in ecology often fall short of the global scale of ESMs. While these models generate valuable insights regardless of their ultimate contribution to ESMs, large-scale integrative understanding of global change impacts requires the use of ESMs because of the many interactions within and among the components of the Earth system. For clarity in terminology, we define "Earth system models" as models which represent the interactions among land, atmosphere, ocean, and cryosphere processes and follow the principles of energy and matter conservation. While we focus specifically on including ecology in the terrestrial component of ESMs, our recommendations can apply to similar challenges in other disciplines (e.g., marine ecology and modeling ocean-atmosphere interactions). The land component of ESMs can and should continue to incorporate ecological processes to improve model realism and to better understand the role of ecological processes within the larger Earth system.

Scientists in both empirical and modeling communities are aware of the need for and benefits of collaborating around ESMs. ESM developers understand that ecology plays an important role in controlling terrestrial ecosystems and that ecological insights can generate models that more faithfully represent real systems, both conceptually and in terms of model uncertainty. Empiricists, on the other hand, understand the potential large-scale impact of their work and that ESMs can help to realize this impact (Fig. 3). For example, ESMs are useful for expanding the temporal and spatial scale of ecological research beyond the constraints of a particular set of sites or experiments. Additionally, models can be used to explore interactions and feedbacks between ecological and climate factors that might be prohibitively complex to measure directly. Models are an important means for ecologists to explore new concepts and generate insights about complex systems that can lead to testable hypotheses. Finally, models are a means to understand the impact of specific management and policy decisions and help stakeholders to make science-informed decisions.

Despite the mutual benefits that empirical and modeling communities receive from collaborating, obstacles remain to better integrating these communities (Leuzinger and Thomas, 2011; Reed et al., 2015). While most empiricists are adept at developing ecological theory for their specific species or system, translating that theory into a generalized mathematical formula can be challenging without decades of research gathering long-term data over broad scales. Next, empiricists face the formidable task of integrating this mathematical formulation into an ESM. ESMs can exceed millions of lines of code (Danabasoglu et al., 2020), and hunting for the right place to insert new code without breaking the rest of the model can be daunting. Working within the particular computing language or framework of an ESM can also be intimidating without extensive training in computational science and applied mathematics, which university ecology programs typically do not offer. Additionally, the overwhelming complexity and ambiguity of large models can make it difficult, without training, to assess the reliability of model results. Given these obstacles, an empirically-focused ecologist might question whether it is a good use of their time to put in the training and work involved with modeling ecological processes in the Earth system.

Modelers working to integrate ecological processes into ESMs, many of whom have formal ecological training, also face challenges in this partnership. Ecological and biological processes are inherently more complex and challenging to quantitatively define than the physical and/or chemical processes that drive most atmospheric or ocean models. As an example, the physiology of stomata does not conform to the principles of fluid dynamics that underpin the atmospheric and ocean components of ESMs. Quantitative ecology is a robust field, but the math of ecology is often defined in units of genes or whole organisms using statistical relationships, rather than the units of matter and energy and process representations that ESMs use, and translating between the two is persistently difficult.

Even when ecology can be quantified in a way that can be incorporated into an ESM, ecological data can be time- and resource-intensive to gather, and model development can be limited by the availability of all the necessary data to drive, tune, or test a new process. Including all ecological processes that impact water, energy, or biogeochemical cycles can lead to models that are overly complex and lack adequate foundations in measured data. Modelers are sometimes reluctant to add a new process without convincing evidence that its impact outweighs the uncertainty it adds to the model. Most ESMs strive to balance ecological realism with excessive complexity, which can lead empiricists to be frustrated with the disconnect between model parameters, processes, and reality. Meanwhile, modelers may grow frustrated and overwhelmed by the abundance of ecological data that "should" but cannot easily be incorporated into models. Resolving the realism-complexity dilemma requires modelers to understand the principles and constraints of researching ecological processes, while empiricists should be more involved in model development and aware of the unique data needed to translate ecological concepts for ESMs.

We address these challenges by providing a clearly defined map of the stages involved in the incorporation of a new ecological idea into an ESM. We seek to pull back the curtain on the complex, multi-scale workflow of coupled model-data-theory development (Fig. 1, 2, 3) and lower the barriers to interdisciplinary collaboration by articulating various phases and considerations along the way (Fig. 4). Below, we discuss the history of incorporating ecology into ESMs to provide context for the characteristics of modern ESMs. We then present our suggested workflow for integrating ecological processes into ESMs (Fig. 4). In this workflow, we describe the iterative procedure of data collection and model development for understanding ecological processes and models at different scales (Fig. 3). We highlight three stages through this workflow and the valuable outcomes at each stage, regardless of whether the endpoint of incorporating an ecological process into an ESM is reached. Finally, we include a list of resources to guide scientists through all the stages of this workflow. These guidelines and the suggested workflow will facilitate stronger connections between empirical and modeling communities, improving ESMs through realistic process representation and increasing the impact of ecological research.

2. History and context for current decision-making in ESM development

For many ecologists, Earth system modeling may seem a distant discipline, but in fact, ecology is already an important part of ESMs. The origin of ESMs is nearly 100 years old. In the early 20th century, an early model of weather forecasting (Richardson, 1922) required knowledge of land surface temperature, surface-absorbed radiation, and exchanges of heat, moisture, and momentum with the atmosphere. As a result, the model acknowledged the role of energy and moisture fluxes from plant canopies, and included rough representations of stomatal conductance and leaf fluxes in its calculations. In the 1960s, modelers expanded their work to the global scale with different labs and centers developing atmospheric general circulation models, which would form the foundation of some of our present-day ESMs (Edwards, 2011). As model development continued, terrestrial vegetation and human modification of the land became recognized as necessary aspects of climate science (Schneider, S. H., & Dickinson, 1974), and prominent studies identified surface albedo, evapotranspiration, and deforestation as important climate regulators (Charney et al., 1975; Dickinson et al., 1986; Sagan et al., 1979; Shukla and Mintz, 1982).

In the 1980s, attention turned to representing more than the atmosphere in global models. Models of the land surface, such as the Biosphere-Atmosphere Transfer Scheme (BATS; Dickinson et al., 1986) and Simple Biosphere model (SiB; Sellers et al., 1986), were developed for coupling with atmosphere models. These models initially focused on the biogeophysical processes of energy, moisture, and momentum fluxes and the associated hydrologic cycle. These models represented vegetation in more detail, including traits such as stomatal conductance, canopy height, leaf area index, and rooting depth. Photosynthesis was also recognized as an essential process to model, initially as a diagnostic (Dickinson et al., 1981) and later as a predictor (Sellers et al., 1996) of carbon and water fluxes (Bonan, 1995; Denning et al., 1996). Building upon a history of ecosystem biogeochemical models first conceived during the International Biological Program (IBP) in the 1960s and 1970s, the carbon cycle was subsequently added to ESMs so that atmospheric CO_2 concentration automatically changed over time rather than being manually specified (Cox et al., 2000; Fung et al., 2005). Bioclimatic rules and simplified equations for competition for space were also added to allow vegetation composition and biogeography to change in relation to the simulated climate (Bonan et al., 2003; Foley et al., 1996; Sitch et al., 2003).

The current generation of ESMs now also includes models with nitrogen and phosphorus cycles, wildfires, biogenic volatile organic compound emissions, mineral dust emissions, methane, wetlands, agricultural management, and land use/land cover change (Bonan, 2016). That many ecological and biogeochemical processes are now included in ESMs is a defining feature in the evolution of climate models, which initially focused on the physical system, to today's more comprehensive ESMs that emphasize the interdisciplinary aspects of climate science (Bonan and Doney, 2018). For example, representations of the nitrogen and phosphorus cycles were added to some ESMs because of their role in regulating the carbon cycle (Thornton et al., 2009; Wang et al., 2010; Yang et al., 2014; Zaehle and Friend, 2010). Similarly, more soil biogeochemical models are including direct representations of microbial populations because of their controls on nutrient and carbon cycling (Huang, Y. et al., 2021; Kyker-Snowman et al., 2020; Wang et al., 2017b; Wieder et al., 2015d, 2018). However, many important processes are still absent from ESMs; for example, herbivores are recognized in ecology as important ecosystem drivers, but are not widely included in ESMs.

Conversations about including ecology in models have become increasingly common in the modeling community, particularly as modelers seek to better match model projections with observations. ESMs continue to be modified to include ecology that impacts model calculations of surface fluxes of energy, moisture, carbon, and momentum. What conditions need to be met for a process to be considered for integration into an ESM? The ecological properties and processes that have made their way into ESMs reflect choices by the modeling community about where to focus its efforts, as well as the practical limitations of the modeling work itself. In general, new ecological processes enter an ESM if:

- **The process can (or is hypothesized to) influence climate on large spatiotemporal scales.** Given the effort needed to code and test the addition of an ecological process into an ESM, the impact of this addition needs to be visible on large spatial scales or on long time frames. For example, explicit representations of vegetation were added to ESMs because they had a clear impact on and improved the performance of climate models through regulating water fluxes on long (e.g., decadal) timescales (Dickinson, 1984; Dickinson and Henderson-Sellers, 1988; Sato et al., 1989; Sellers et al., 1986).
- **The process can be reasonably incorporated into existing model infrastructure.** New ESM developments build on earlier ones, which means there needs to be a clear plan for how to insert the code for the new process into the existing model code. In addition, this linking should be able to occur without major restructuring to the model's existing structure. For example, in order to integrate nitrogen cycling into an ESM, code needed to be developed to link nitrogen fluxes to the physics of the land surface and calculations of carbon fluxes (Bonan and Levis, 2010; Thornton et al., 2007).

- **Process understanding and data are available to model the process globally.**
  The equations representing the process need to be solvable on a three-dimensional global grid (latitude, longitude, height) as well as on short time scales representing the model's timestep for calculations (e.g., 30 minutes). Ideally, any input data required by the new ecological process should be available globally as a gridded product or be connected to existing variables simulated by the ESM. For example, the TRY database provides data that has been used to create global maps of plant traits that are used as the foundation for plant functional types (Kattge et al., 2011).
- **The mathematics of the process are tractable within the limits of current computing resources.**
  Computing resources have significantly expanded, allowing more ecological processes to enter models. However, there are still limits to numerical processing power. Processes must be reducible to a mathematical form that does not dramatically increase computing costs of the entire ESM, given that existing ESMs already push the capacity of the world's most powerful supercomputers (Washington et al., 2009).
- **There is a community of researchers dedicated to developing, testing, and maintaining the process in the model.** Writing the code for a new ecological process is only one part of the process for integrating a new component into an ESM. Once code is written, it needs to be tested with different components of the ESM and under different simulation conditions before the process can be considered as an official addition to the ESM. In addition, the continued longevity of the process in the model requires there to be one or more researchers continuing to maintain and update the modeled process as new data about the process and new changes to the ESM are made. As such, a community of

researchers with the resources to both advocate for the inclusion of the process and support its inclusion in the model long-term is needed.

With the origin of ESMs in the atmospheric and physics communities, it is perhaps not surprising that the incorporation of ecology into ESMs started in these communities. The modeling community has initiated several grassroots efforts to bring more ecologists into ESM work. These efforts range from creating conference workshops and research coordination networks (e.g., Cheng, 2018; Leuzinger and Thomas, 2011; Rogers et al., 2014) to leading tutorials and short courses to provide training for empiricists and modelers to bridge these subdisciplines (e.g., the CTSM tutorial at NCAR; FluxCourse; Bracco et al., 2015). However, these efforts are limited in the number of people they can reach. Larger, systematic changes in education and training, funding structures, and engagement across communities are critical to shifting the current siloed paradigm. We propose a new practical roadmap for empiricist-modeler collaboration that breaks down traditional disciplinary boundaries and fosters iterative, shared conceptual development.

3. Introducing the practical roadmap for integrating ecology and ESMs

New efforts to close the gap between ecological empiricists and Earth system modelers are growing, but the two communities could still be better integrated. To do so, each community needs to understand the approaches used by the other and work together both to develop the technical advancements needed to expedite data-model integration (e.g., Fer et al., 2021) and to address the social dimensions of collaboration. Focusing only on technical or mathematical aspects of data-model integration can perpetuate barriers through the use of discipline-specific language and dismissal of non-technical obstacles to participation (Bernard and Cooperdock, 2018; Duffy et al., 2021; Morales et al., 2020), which can lead to members feeling excluded and keep disciplines siloed (Marín-Spiotta et al., 2020; Mattheis et al., 2019). In general, effective cross-disciplinary collaboration depends on several key principles that facilitate team dynamics (O'Rourke et al., 2014) and need to be built into the start of a collaboration; namely: respect and trust among all team members, clear communication, common goals, and effective project leadership (Nancarrow et al., 2013). Research shows that clear team communication is essential for optimizing project outcomes (Anderson-Cook et al., 2019; Kuziemsky et al., 2009), as it is the foundation for identifying shared objectives and building interpersonal relationships that are necessary for teams to remain cohesive during times of conflict (Cooley, 1994). Breaking down barriers to interdisciplinary collaboration requires researchers to adopt practices that not only improve their collaboration, but also dismantle the inequitable and exclusionary dimensions of their disciplines (Bala Chaudhary and Berhe, 2020; Duffy et al., 2021; Emery et al., 2021). Additionally, computing tools and frameworks evolve rapidly, and solutions that focus on facilitating collaboration will outlast any particular technological tool. To achieve better integration and collaboration among empirical and modeling communities, we outline a few necessary foundational principles of collaboration and educational change (Fig. 2). We also propose a workflow that highlights one possible pathway to improve collaboration between fields to improve the work of each (Fig. 4).

In addition to strengthening empiricist-modeler team dynamics, we emphasize the need to rethink ecological education to incorporate process modeling concepts and normalize regular collaboration between empirical and modeling subdisciplines. At many institutions, the ecology curriculum emphasizes field techniques and statistical analysis, but fewer options may exist for courses on ecological process-based modeling. While a given department may offer one or a few courses, often these are not required in ecological education, and programming skills development is limited to high-level statistics programs and languages like R and python that do not entirely prepare students for the computer science that powers modern ESMs. Conversely, educational requirements in other disciplines, such as atmospheric sciences, frequently include both field and modeling techniques and in-depth quantitative and programming skills in which computational science and applied mathematics are essential tools of the science. Ecologists wanting to learn modeling techniques often find themselves taking classes outside their discipline, attempting to separate content from technique and applying techniques to a different field, which is a challenging task. This can pose a large enough burden on the student that many do not follow through, finding it easier to continue with familiar skills. A detailed plan for modifying the way ecology programs teach quantitative skills is beyond the scope of this paper, but others have begun the difficult work of rethinking educational paradigms to address this problem (Hampton et al., 2017).

ESM communities also need to identify opportunities for redesigning their training so they can learn more about ecological concepts and data collection frameworks. Ecological data is complex and filled with caveats, and modelers often encounter data after it has been processed and organized and thus may be unfamiliar with the nuances of data collection and analysis. Modeler training in ecological concepts could take place at the student level, with classwork focused on the impacts of living organisms on biogeochemical, water and energy cycles, or at later career stages via field site visits, shared seminars, interdisciplinary conference sessions, etc. One powerful approach is for a modeler to take a day trip with an ecologist to engage in fieldwork. While we recognize that the outdoors are not a comfortable space for many people and this can be a barrier to participation (Giles et al., 2020; Morales et al., 2020), direct experience with how an ecologist gathers data can be an invaluable insight into the the limitations and interpretation of data in a modeled context. Virtual site visits using recorded video are another alternative for those unable to visit in person.

Beyond these foundational shifts, we propose a new workflow for modeler-empiricist collaboration with three specific stages (Fig. 4). This workflow is meant as one (but not the only) route for any empiricist or modeler to understand the stages involved in integrating a new process or idea into an ESM. We strive to break down traditional disciplinary barriers between modelers and empiricists and highlight the iterative collaboration and shared skill sets that are necessary at each stage. The first stage in this workflow ("Assess process & potential impact") includes a list of questions that anyone (regardless of programming ability) can ask to assess the readiness of a process for incorporation into an ESM. The second stage ("Test process alone") involves the quantification and scaling of the new ecological concept using simple models and large-scale parameter determination. Finally, the last stage of the flowchart ("Test process with ESM") discusses the multiple steps involved in making modifications to an ESM, evaluating the impact of the new process on model-wide behavior, and projecting the large-scale impact of the new process within the Earth system. Importantly, each stage of this workflow generates valuable scientific products (e.g. hypotheses, new or improved theory, regional or ecosystem-scale models), regardless of whether the endpoint of "inclusion in an ESM" is reached. We recognize that tackling any part of this workflow is challenging for aspiring and seasoned modelers alike, and we encourage researchers to see it through. We include specific illustrative examples for each stage of the workflow (Boxes 1-3) and one that illustrates stepping through the entire workflow (Box 4), as well as resources for accomplishing each step (Table 1).

Workflow part 1: Identifying and understanding a new process

The first stage of the proposed workflow assesses the readiness of a new process for inclusion in an ESM based on how well the process can be quantified and understood in an ecosystem context. Many empiricists recognize the importance of their work for understanding global change and highlight the need to incorporate new processes into models. However, highlighting this need has minimal impact on ESMs unless coupled to an understanding of the stages of model development and the unique types of data necessary to progress through those stages. As such, the first part of the workflow provides three guiding questions empiricists should ask to assess whether a new process is ready for inclusion in an ESM, each of which will be discussed in more detail in the following paragraphs (Fig. 4, "Assess process & potential impact"). These questions can help identify data gaps and point to valuable targets for future experiments to facilitate downstream ESM integration. Importantly, these questions can be addressed by any empiricist without requiring formal modeling skills. While connecting with modelers is not required at this point, it can be helpful in co-designing future experiments to make process integration more streamlined (Fig. 2).

The first guiding question aims to evaluate the level of theoretical/empirical understanding of the targeted process: *Do you expect your process to respond consistently to environmental drivers, enabling scaling across space and time?* Consistent, quantified patterns are the heart of process modeling. Detailed understanding of a process or mechanism at a single location can help to identify whether the process is likely to scale. In order to develop a broad theoretical representation of a process, it can help to determine whether data are available across multiple sites and ecosystem types and at various timescales. For example, if a specific tropical soil owes its high carbon storage capacity to a unique volcanic mineral (Torn et al., 1997), it would be wise to evaluate the carbon storage capacity of soils without this mineral before generalizing observed patterns to a global scale. While it is not necessary at this stage to gather enough data to create a fully quantified global representation of a process, information gained in this step may help identify data gaps and guide the design of additional empirical experiments needed for large-scale modeling, such as repeating experiments across underexplored regions or a wider range of environmental conditions. This step also helps to identify conceptual areas where a large amount of data may be available but consistent relationships with environmental factors and process rates have not yet been identified. For instance, soil microbial biodiversity is being rapidly catalogued through metagenomics, but these data do not yet provide critical information for representing process rates at large scales (Fierer et al., 2021).

The second question in this stage of the workflow requires ecologists to get familiarized with ESMs and the way processes are represented: *Is your process already in or related to an existing process in an ESM?* Investigating this question will help identify existing model frameworks that can be used as scaffolding for building simple models and ultimately incorporating the process into an ESM. ESMs represent similar environmental processes using a variety of different approaches and equations, so it might help to start by identifying one or more ESMs that you may be interested in and reading model documentation to determine how related processes are represented and whether the model will fit your needs. For example, if you want to improve the representation of foliar nitrogen acquisition, it is vital that the model you choose already has a terrestrial nitrogen cycle. This is also an ideal time to discuss collaborations with ESM developers. We encourage ESM developers at this stage to welcome ecologists interested in working with ESMs by taking the time to explain modeling concepts in jargon-free language and providing resources to work through technical challenges.

If the selected ESM already contains a model of the process, the empiricist can consider how it can be improved or revised using new data or theoretical understanding. Many times a process is represented implicitly (e.g. soil microbial activity is often represented using a cascading decomposition scheme (Wieder et al., 2015b, 2018)). Illustrating that explicit representation of the process will fundamentally change model behavior will help to determine whether an explicit representation is needed. In addition, if the current representation of the process connects multiple cycles (e.g. carbon and nitrogen, water and energy), exploring existing model structures can help empiricists understand all the connections between their process and various cycles that must be elucidated and quantified when updating the ESM. Like hooking up speakers to a television or finding the right dongle to plug in your phone, the new process will only work within the ESM if all the appropriate ins and outs are connected. If the process is not currently in a model, it is worth investigating why not (perhaps connecting with an ESM modeler) and whether it might be implicitly included through other model process representations. For example, plant hydraulic stress is not always explicitly included in ESMs

(Kennedy et al., 2019), but may be implicitly included by existing connections between soil moisture and stomatal conductance.

The third and final question helps to identify ecological concepts that may be more appropriate to a different type of modeling because they are unlikely to alter climate simulations within an ESM: *Is the process likely to influence climate on scales of time and space consistent with other ESM processes?* Put another way, is the process likely to change the results of global climate simulations using ESMs? Generally, ecology in ESMs impacts climate prediction in two major ways: through biogeochemical (carbon and nutrient cycling) and biogeophysical (evapotranspiration and energy fluxes) processes. Coupling these processes provides a means for assessing feedbacks between ecosystems and climate that distinguish ESMs from stand-alone ecosystem models.

Simple estimates can be made to assess whether a process, when applied to large regions or the entire globe, has the potential to meaningfully influence climate. For example, the general process of insect herbivory, which responds to temperature (e.g., Deutsch et al., 2018; Edburg et al., 2011) and could meaningfully affect carbon fluxes through changing plant biomass, might influence climate (Box 1). On the other hand, temperature affects the distribution and abundance of mosquito species (Hunt et al., 2017), but if mosquitoes are not known to have a meaningful impact on climate, inclusion of mosquito species distributions would not change the outcome of ESM simulations, and may be better suited to a different type of model. In addition, new, climate-influencing processes must occur or change at a rate that is meaningful at ESM timescales. For example, changes in environmental conditions may alter the rates of soil microbial metabolic processes over the course of minutes or even seconds, but these rapid fluctuations are too fast to capture in the timestep of a typical ESM. On the other end of the spectrum, bedrock weathering is a process that releases nutrients for plants and may impact plant biomass (Morford et al., 2011), but it happens so slowly that it is unlikely to shift simulated plant productivity in an ESM over decade to century timescales.

Apart from facilitating ESM incorporation, these questions produce valuable intellectual products on their own: greater understanding of how a process fits into the terrestrial system, identification of knowledge gaps and a clear path towards future empirical work, and determining whether an ESM is the appropriate modeling tool for the process of interest. Reflecting on these questions can help ecologists define "future directions" for their work with greater specificity than "inclusion in a model," and also generate valuable insights into the scale of an ecological process and its connections to water, energy, or biogeochemical cycles. In a classroom setting, these questions can be an effective way to practice "thinking like a modeler" without requiring any involvement with programming. Regardless of whether the answer to all of these questions for a given ecological concept is "yes", they are beneficial for ecologists to ask.

Box 1:

Herbivores like insects and grazers have large impacts on plant biomass and productivity, yet they are still absent from ESMs. How do the conceptual questions in Part 1 of the workflow guide next steps in deciding whether to incorporate herbivores in ESMs? Although herbivores are broadly not yet included in ESMs (Question 2) and are known to have important impacts on plant biomass with feedbacks to climate (Question 3), ESMs also require that any new process behave consistently across space and time (Question 1) in a way that can be captured quantitatively. To move forward with incorporating herbivores into ESMs, the known impact of herbivores on plant biomass must be reduced down to quantifiable patterns that are *consistent across space and time. For example, do herbivores reduce plant biomass by a fixed proportion, or by a proportion that depends on climate factors already present in ESMs like temperature and precipitation? Does the impact of herbivores vary in a predictable way across continents and ecoregions? If the answer is yes, then perhaps a simple model can be developed (Workflow part 2) or existing simple models can be considered for ESM incorporation (Workflow part 3).*

Workflow part 2: Beginning to work with simple models

After assessing the theoretical understanding of a process and its likely importance for terrestrial ecosystems and climate, the next workflow steps involve the iterative development, implementation, and evaluation of simple models outside of the ESM, in addition to the collection and/or assembly of data necessary to apply the simple model at large scales (Fig. 4, "Test process alone"). The aim of these activities is to generate knowledge, highlight uncertainties, and refine understanding of the process(es) in question. At its core, this stage involves identifying formulas to represent our theoretical understanding of ecological systems. This stage is a key precursor to working with ESMs because once a process is integrated into an ESM, it becomes harder to discern the cause of disagreement with observations, and uncertainty increases. For example, photosynthesis can be evaluated with leaf gas exchange data in highly controlled chambers. Gross primary productivity, on the other hand, is evaluated using eddy covariance flux towers. Errors can arise in the model's scaling from leaf to canopy, soil moisture, nitrogen availability, leaf area index, and aspects of the model other than the photosynthesis parameterization (Rogers et al., 2017). The "test process alone" stage is essential to identify the adequacy of a process model before compensating errors occur within the ESM. Although not a strict requirement, this phase of the workflow is best accomplished with equal, collaborative contributions from both empiricists and modelers (Fig. 2) including someone familiar with ESMs who can craft a bridge for future process incorporation.

Simple models are created at this stage by translating knowledge from conceptual models of organisms and ecosystems to mathematical representations of matter and energy. The development of simple models can start by creating a simple statistical model or using a pre-existing model. For example, R has a photosynthesis package (Duursma, 2015) that can be used as a starting point for modifications to photosynthesis like temperature acclimation (e.g., Smith et al., 2016) or ozone damage (e.g., Lombardozzi et al., 2012). Simple models can also be developed using any coding language (both R and Python are free and open source), or even start by using a spreadsheet program like Excel, and can range in complexity from a single equation to a complex web of variables and parameters. Unlike the first phase of the workflow, testing theory with data at this phase requires some comfort with programming and data management (for resources, see Table 1). These activities can be easily integrated into ecological coursework, and a variety of resources have been developed to facilitate this (e.g., Carey et al., 2020). Additionally, cross-disciplinary collaboration is beneficial at this stage, as it helps to formalize conceptual models, clarify assumptions, evaluate ideas within the scientific community about a process, connect various components of ecosystems and the Earth system, and test the broader applicability of theories over space and time.

In addition to simple model development, this phase of the workflow involves assembling the data necessary to estimate parameters and drive simple models at large scales. (Note: In a model, a “parameter” is the value of a variable in an equation. The word “parameterization” may seem like a derivative of “parameter”, but is in fact a separate concept referring to representing a complex microscale process as an approximate bulk process. For example, photosynthesis is a parameterization of subcellular-level processes, and may use parameter values within the calculation (Bonan, 2019)). Necessary data fall into several distinct categories: data for parameter estimation during model development, driver data to feed into the model (e.g., climate or soil characteristics), and data for benchmarking the model following simulations (i.e., observational data to compare against model output).

At this stage, it is worth making a "shopping list" of the data necessary for a given modeling exercise and evaluating the availability of values at the relevant scale (Fig. 3). These data may come initially from a single site or lab experiments, but to eventually scale model results globally, data gathered across multiple regions and experiments become useful. ESMs use a variety of large-scale datasets for parameter estimation and evaluation, and it can be helpful to seek out datasets already in use before attempting to assemble a new dataset from scratch. Large-scale data can come from meta-analytical techniques and syntheses (e.g., Ainsworth and Long, 2005; Field and Gillett, 2010; Lombardozzi et al., 2013), pre-existing large synthesized datasets (e.g., SoDaH (Wieder et al., 2021), TRY (Kattge et al., 2011)), satellite data (e.g., Li and Xiao, 2019), or model-derived products (e.g., Fluxnet-MTE (Jung et al., 2020)). Direct measurements are generally preferable for parameter estimation and model evaluation but are not always feasible to collect. As a result, parameter estimation and model evaluation often use data products (i.e., data that have been modified by models) to achieve the spatial and temporal scales required by the ESM. Data products can be closely connected to the original data (i.e., data averages) or less closely connected (i.e., output of another mechanistic model that uses data as an input). Understanding the uncertainty of a data product is critical for determining the value of its use in parameter estimation and model evaluation (Dagon et al., 2020; Dietze, 2017). Simple models often get stuck here on the way to ESM incorporation because of gaps in data requirements to run models at global scales (e.g., lack of maps of soil edaphic properties or other input data that may be critical for further model development).

The creation and improvement of simplified mathematical models and large-scale synthesized datasets makes several valuable contributions to understanding and refining ecological theories, regardless of the eventual implementation in ESMs. Simple models help formalize, and make explicit, the underlying assumptions in the theories they represent and can illustrate weaknesses in existing theory. As such, they can be used to generate testable hypotheses that can be interrogated with existing data or new experiments. Estimating parameters for simple models with available observations helps identify data and knowledge gaps that can be addressed with further study. Compared to larger ESMs, simple models have greater traceability, allowing scientists to explore and understand model complexity, their associated uncertainties, and emergent properties that can be evaluated with independent observations. These simpler models also have the advantage of being easier to use, with greater flexibility and lower computation costs than running a full ESM, and can potentially be implemented in ESMs in a modularized manner that allows for testing multiple ecological theories (e.g., Fisher and Koven, 2020). Finally, these models help to clarify theory and develop concepts through independent community efforts to use them and improve their process representation.

Box 2:

After establishing that a new process is appropriate to consider including in an ESM (Part 1), what comes next? Current models of soil microbial activity highlight Part 2 of the *workflow: simple quantified models evaluated at a variety of scales but not yet incorporated into ESMs. As an example, the MIcrobial-MIneral Carbon Stabilization (MIMICS) model was motivated by theories highlighting interactions among soil microbes and minerals that are responsible for soil organic matter decomposition and persistence. A simple process model was initially developed in R using measurements from laboratory experiments and rates of leaf litter mass loss. This model was tested first at a single site (Wieder et al., 2014), and subsequent evaluation across continental and global scale gradients illustrated reasonable agreement with litter decay rates and soil carbon stocks (Wieder et al., 2015d) and a higher vulnerability of Arctic soil C stocks, compared to models that implicitly represent microbial activity (Wieder et al., 2019). MIMICS continues to undergo further development (e.g. to include coupled C-N biogeochemistry (Kyker-Snowman et al., 2020) and vertical resolution (Wang et al., 2021)), refinement (Zhang et al., 2020), and evaluation (Basile et al., 2020; Koven et al., 2017; Shi et al., 2018; Sulman et al., 2018). All of these activities rely on conducting simulations across multiple study sites and at global scales, which is a valuable precursor to considering incorporating MIMICS into an ESM.*

Workflow part 3: Integrating processes into ESMs

Developing and evaluating a simple model ultimately paves the way for integrating a process into an ESM, as illustrated in the final stage of the workflow (Fig. 4, “Test process with ESM”). The first step is deciding which ESM to use. Many ESMs exist and vary substantially in their ecological process representations (Fisher and Koven, 2020), and adding a new process requires an understanding of how processes of interest are currently represented in a given ESM (as in Stage 1) and a simple model that can be integrated within the framework of that ESM (developed in Stage 2). Additionally, some ESMs have proprietary or restricted access (e.g., GFDL-ESM, IPSL-CM5 (Dufresne et al., 2013; Dunne et al., 2020)) and require collaboration and/or approval by model developers, while others are open-source and community driven (e.g., CESM, E3SM (Danabasoglu et al., 2020; Golaz et al., 2019)). While not always required, incorporating new processes will be most efficient when building relationships with model developers who can help with technical aspects of code development. For example, developers with experience in running and testing the model can provide code structure guidance and highlight possible interactions or feedbacks among processes that might not be obvious to a novice model developer. ESM communities can be insular and siloed at times, and ESM developers at this stage can help build more integrated empirical-modeling collaborations by seeking out and remaining open to working with ecologists (see Table 1 for several opportunities).

Once access to model code is available, integrating the new process representation can begin. The first step is finding the location to integrate the new process. While this will vary depending on the ESM, code modules will often have descriptive names and the location of variables within the code can be searched using linux- and editor-based search tools (e.g., grep). It is also helpful to find a similar variable or process in the code (with similar inputs and outputs) that can be used as an example for how to structure the new process code. Having an example to mirror can be particularly useful in identifying other modules where the variables may be required (e.g., sometimes setting the initial value for variables happens in a different module). Additionally, it can be helpful to outline or diagram a work plan in advance, noting the modules and variables that will need to be added, modified, and connected.

Modifications should build on each other, starting with a simple change: for example, add a single variable, and then test that the code will compile and run for a short period of time. Sequentially add more complexity, connecting the new variable or process to existing model structure. Using this layered approach will help to identify any structural bugs early in the development process. Although the ultimate goal is to have a sophisticated representation that includes spatially-varying processes, simpler versions of the model can -- and should -- be tested to determine the sensitivity of the system to the new process. These simpler model iterations are excellent training tools for graduate students and postdoctoral trainees as they become more familiar with the model. Once the basic framework for the new process is in place, it can be tested to identify the magnitude of change in relevant processes, as well as any interactions with other ecosystem processes. Often, these proof-of-concept simulations can turn into publications that highlight the potential importance of the process at site or global scales and identify gaps in data that can help to improve the process representation.

Throughout the development, testing, and evaluation process, the simplest relevant version or component of the ESM available should be used. For example, if the new process does not rely on carbon cycling, it may be possible to leave out this portion of the model in your testing, allowing the model to run faster and reducing the complexity of model interactions. Often with ecological processes, the development process uses only the terrestrial component of an ESM driven by a gridded atmospheric data product (e.g., reanalysis), since fully coupled ESM runs are far more computationally expensive than smaller terrestrial-only runs. Additionally, running in the coarsest available resolution and for the smallest spatial domain possible (e.g., a single site) will expedite model testing. Once code is tested, running it globally (and eventually coupled to an atmospheric model) is necessary to ensure the simulation operates appropriately over the global domain.

An approach called "modular development" can also be useful for testing and evaluating different ecological theories, and can be employed when implementing new processes in ESMs (Fisher and Koven (2020); see also Clark et al. (2015)). This involves adding an alternate representation of a process that is already simulated in a model (not removing the process) and letting the user specify which theory the model will use in a given simulation. For example, testing multiple representations of stomatal conductance (Franks et al., 2018), soil carbon and nitrogen cycling (Wieder et al., 2015a, 2018), and hydrology (Clark et al., 2008, 2011) have been helpful in testing different theories and highlighting when and where certain process representations perform best. This allows for refinement of existing theory and process representation, advancing the state of current knowledge.

Once the new process is incorporated, the model must be tested and evaluated. A first step is to determine whether the new process fundamentally changes model behavior relative to a simulation without this process. Does it affect other simulated processes, and by how much? Many processes do not exist in isolation within a model and thus cannot be modified for only one purpose. Better models of photosynthesis, for example, may be desired to improve the carbon cycle, but also impact energy and water fluxes to the atmosphere through stomatal conductance (Bonan et al., 2011). A second step is to evaluate model behavior against observations. Model evaluation is most effective if multiple processes are assessed, and is most useful when compared to evaluation of a baseline model simulation where the new process is not simulated. This step is similar to simple model evaluation in the second stage of this workflow, but this evaluation process should be repeated once the simple model is embedded within an

ESM. One simple form of evaluation is to run a simulation at a single location where relevant observational or experimental manipulation data have been collected, such as a field site or a flux tower (Cheng et al., 2019; Medlyn et al., 2015). These data can be used to assess whether the new model behavior fundamentally changes model performance (De Kauwe et al., 2013, 2014; Smith et al., 2015; Thomas et al., 2013; Zaehle et al., 2014). It is also important to evaluate global responses. While global data can be more challenging to access, several resources are currently available. Perhaps the most useful is the International Land Model Benchmarking (ILAMB; Collier et al., 2018) project, which has developed internationally accepted benchmarking standards for ESM performance. This project has compiled global datasets for a range of variables and can help to identify where model performance is enhanced or degraded. Remotely sensed data products can also help with model evaluation at regional to global scales.

One of the greatest challenges in ESM development is ensuring parsimony while capturing the full range of biological complexity. This is particularly challenging for community models with contributors from multiple fields and institutions, which commonly suffer from "feature fatigue". Human instinct is to continue to add features to a solution, even when removing features may be more beneficial or efficient (Adams et al., 2021). While adding processes can improve model realism, care must be taken to avoid sacrificing model reliability, which can be degraded with the addition of uncertain parameters (Prentice et al., 2015). Eco-evolutionary optimality theory is one recent tool that can be used to improve model realism while limiting the number of new parameters (Box 3; Scott and Smith, 2021; Wang et al., 2017a). Unlike statistical approaches where environmental responses are hard-coded with parameters, a theoretical approach allows process responses to emerge with fewer parameters

(Prentice et al., 2015). These responses can then be tested with data that might, in a more statistical approach, be needed to estimate parameters.

The workflow so far has presented guidelines for incorporating a new process into an ESM, which requires substantial work in developing and incorporating new code into a model and then evaluating the responses of terrestrial processes. Often, the ecological workflow ends here with the assessment of the global-scale impact of a process and how it may change ecological functioning through time. Beyond this, an exciting next step is to understand whether this new process has climate feedbacks by comparing land-only and coupled model simulations. Land models can be coupled to other ESM components (atmosphere, ocean, ice, etc.) to investigate global feedbacks in water, energy or biogeochemical cycles. Connecting land and atmosphere components allows investigation of unexpected feedbacks with the atmosphere that may be different from land-only simulations.

Box 3:

One example of how models have maintained parsimony (Part 3 of the workflow) is photosynthetic acclimation (Smith and Dukes, 2013). Initially, empirical models were developed to simulate temperature acclimation of photosynthetic biochemical capacity in ESMs based on observed responses (e.g., Kattge et al., 2009; Kattge and Knorr, 2007) and then incorporated in ESMs (Friend, 2010; Lombardozzi et al., 2015b; Mercado et al., 2018; Smith and Dukes, 2013; Ziehn et al., 2011). However, more recently, eco-evolutionary optimality theory has been invoked to simulate photosynthetic biochemical capacity in a way that incorporates the processes without added parameters (configuration variables internal to a model that rely on observational data), thus increasing model realism without altering model reliability (Scott and

Smith, 2021; Smith and Keenan, 2020; Wang et al., 2017a). Eco-evolutionary optimality theory approaches rely on the assumption that natural selection will remove non-competitive traits from an environment, thus providing testable, theoretical trait responses to the environment over short and long time scales, and offer potential promising avenues for adding biological processes to ESMs with little to no added parameters (Franklin et al., 2020). Eco-evolutionary optimality approaches are available to simulate processes at the leaf (Jiang et al., 2020; Prentice et al., 2014; Smith et al., 2019; Smith and Keenan, 2020; Wang et al., 2017a, 2020), plant (Dybzinski et al., 2015; Farrior et al., 2013; Weng et al., 2015) and ecosystem (Baskaran et al., 2017; Franklin et al., 2020) scales.

Box 4:

The following example illustrates the entire workflow, from initial conceptual development to simple modeling to working with ESMs. As part of her research, co-author Lombardozzi measured how leaf-level gas exchange changed in response to ground-level ozone. Upon analyzing her data, she found that leaf-level carbon (photosynthesis) and water (transpiration) fluxes decreased at different rates. Since these are both important greenhouse gases and affect fundamental plant processes (photosynthesis and stomatal conductance, which scale through time and space regardless of biome), she thought that ozone damage could have a global impact and therefore should be included in large-scale models. Although Lombardozzi had no modeling or coding experience, she emailed several people working on the Community Land Model (CLM) to see if they might want to collaborate. She did some research about the photosynthesis and stomatal conductance models used in CLM and talked with modeling colleagues to decide how to best include this type of damage. After completing online Linux and

Fortran tutorials, Lombardozzi started using a simple photosynthesis-stomatal conductance model provided by her colleagues. She applied linear regressions calculated from her experiment to the rates of maximum carboxylation (Vcmax) to simulate ozone damage to photosynthetic enzymes. She was able to show that including ozone damage improved simulated photosynthesis and stomatal conductance at the leaf scale (Lombardozzi et al., 2012).

Did these changes matter globally? Lombardozzi worked with model developers to find out, using the simple model to update code in the CLM to account for ozone damage. Using data from her experiment and a constant ozone concentration, she showed that ozone did have large consequences for carbon and water cycling globally (Lombardozzi et al., 2013). While this experiment highlighted the sensitivity of global processes to ozone damage, it did not provide a realistic assessment of how ozone changes carbon and water cycling. Lombardozzi therefore synthesized existing published literature to determine how photosynthesis and stomatal conductance change in relation to ozone exposure, and identified a complete lack of data for tropical forests (Lombardozzi et al., 2013). Despite missing data for large biomes, these data were then used to update the CLM code to capture responses across different plant functional categories (e.g., broadleaf trees, needleleaf trees, herbaceous vegetation), and when combined with realistic ozone data, simulated that ozone decreases global photosynthesis by 10.8% and transpiration by 2.2%, with larger impacts in Eastern US, Europe, and Southeast Asia (Lombardozzi et al., 2015a).

4. Creating community change across scales

Empirical and modeling communities already work together and influence one another in many ways, yet integrating ecological processes into ESMs remains a persistently slow process with myriad challenges limiting efficient collaboration. Historically, ESMs have been developed by atmospheric and physical science communities while ecology has only been integrated relatively recently, and the disciplinary requirements in trainee education have not provided enough of a shared foundation to build strong conceptual bridges between ESMs and empirical ecologists. These communities must collectively address persistent obstacles including confusing technical language, lack of resources for skills development, and the need for better connections and integration across scientific communities. We provide resources to help expand terrestrial ecological process representation in ESMs (Table 1). With the advent of these and other tools, empiricists will be better poised to take advantage of technical workflows that can help streamline data-model integration (e.g., Fer et al., 2021).

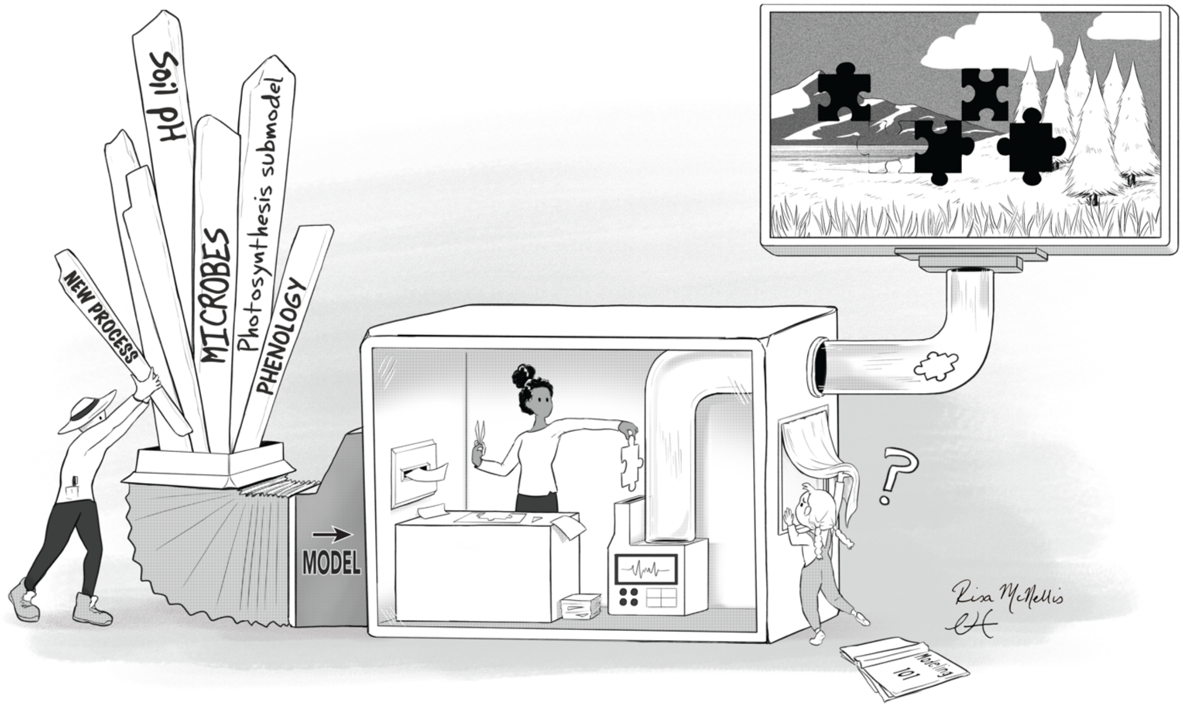
The interdisciplinary work of developing an Earth system model is not only technical, but also social. As such, in addition to the workflow presented above, we offer specific suggestions for restructuring ecological education and interactions within collaborations (see Section III), both of which are key to ensuring that the workflow does not break down. For bridge-building between communities to be inclusive, the modeling and empirical communities need to examine their community practices, values, and norms. This work includes understanding the demographics of *who* is (and is not) represented in the research communities (Bernard and Cooperdock, 2018), *what* processes our communities are willing to consider (or dismiss) as valuable contributions to ESMs (e.g., microbes, moths, management), *where* data are collected and why some regions or ecosystems are over/under sampled (Martin et al., 2012; Metcalfe et al., 2018), *when* we overlook potential collaborators or fail to provide them with platforms for sharing their work, such as at conferences (Ford et al., 2019), and *why* we make the decisions that we do about where to focus efforts.

Improved collaboration between empirical and modeling communities will positively benefit each community. Adding modeling to empirical work can increase its impact while simultaneously advancing ecological theory, modeling capabilities, and model realism. To get started or go further with this work, we have assembled a list of resources for skills development at each stage of the workflow (Table 1). To maintain contemporary resources, please visit the regularly updated website (https://ecoesm.github.io/). Despite the many complex challenges involved in integrating terrestrial ecology and Earth system modeling, there has never been a better time to attempt such difficult work. Finding and communicating with scientists across the globe is getting easier every year, computing resources are rapidly evolving, and the internet provides an ever-growing assortment of free tools for developing new quantitative and programming skills. In addition, funding sources are increasingly recognizing the value of data-model integration (e.g. the NASA Modeling, Analysis, and Prediction program (https://map.nasa.gov/) or the USDA NIFA Data Science for Food and Agricultural Systems program (https://nifa.usda.gov/program/dsfas)) and grassroots efforts are creating a framework for these collaborations using workshops and tutorials. Our insights into the history of ecology in ESMs, workflow for developing and incorporating ecological processes into ESMs, and specific resource suggestions will advance this exciting progress and provide a scaffold for building fruitful bridges between empirical and modeling communities.

Table 1. Table of textbooks and free resources for developing cross-disciplinary skill sets in empirical and modeling work and learning to traverse the stages of integrating new processes into an Earth System model. For a regularly updated list of resources, visit https://ecoesm.github.io/.

Skill/ Category	**Item**	**Description**	**Link**
Programming	NCAR Python tutorials	Basic introduction to the Python language from the National Center for Atmospheric Research	https://ncar.github.io/python-tutorial/
Programming	PEcAn project tutorials	Introduction to working with the Predictive Ecosystem Analyzer	https://pecanproject.github.io/tutorials.html
Programming	The Unix Shell	The basics of file systems and the shell	http://swcarpentry.github.io/shell-novice/
Programming	Udacity	Free courses on basic programming competency with github, linux, R, python, and many others	https://www.udacity.com/
Programming	Software Carpentry	Free courses on basic programming competency with github, linux, R, python, and many others	https://software-carpentry.org/lessons/index.html
Programming	R tutorial	Basic introduction to working with R	https://education.rstudio.com/learn/beginner/
Simple modeling	InsightMaker	Tools for developing quantitative stock-and-flow diagrams of processes	https://insightmaker.com/
Simple modeling	Teaching Resources	Lessons and other resources developed for teaching basic principles of ecological modeling	https://matthesecolab.com/teaching/ http://www.maryheskel.com/teaching.html https://onlinelibrary.wiley.com/doi/full/10.1002/ece3.6757
Simple modeling	Modeling the Environment	Textbook on environmental modeling by Andrew Ford	
Simple modeling	EDDIE	Modeling/forecasting teaching modules developed for NEON sites	https://serc.carleton.edu/eddie/macrosystems/index.html
Simple modeling	Excel modeling tutorial	Tutorial on building simple models in Excel	http://www.mbaexcel.com/excel/how-to-build-an-excel-model-step-by-step/
Earth system modeling	Climate Change and Terrestrial Ecosystem Modeling	Textbook on global-scale ecosystem modeling by Gordon Bonan	https://www.cgd.ucar.edu/staff/bonan/ecomod/index.html https://www.cgd.ucar.edu/staff/bonan/ecoclim/index.html
Earth system modeling	CESM tutorial	Workshop on working with the Community Earth System Model	https://www.cesm.ucar.edu/events/tutorials/
Earth system modeling	Earth System Modeling Framework	Introduction to working with Earth System Models	https://earthsystemmodeling.org/tutorials/
Earth system modeling	CESM-Lab	Cloud version of CLM	https://github.com/NCAR/CESM-Lab-Tutorial

Figure 1. Historically, the process of integrating ecology in Earth System models (ESMs) has often separated tasks along disciplinary lines, with empirical ecologists feeding data into a mysterious “modeling” process and modelers modifying and using data without a thorough understanding of data collection procedures and caveats. The newest generation of scientists has the opportunity to pull back the curtain by developing cross-disciplinary skill sets and building stronger, more collaborative bridges between empirical and modeling communities, with the goal of accelerating the integration of ecological concepts into ESMs.



Dorothy pulled back the curtain to find that the model wasn’t magic after all…

Figure 2. The prevalent existing paradigm in ecology-Earth System model (ESM) integration separates tasks along disciplinary lines, with empirical scientists giving data and generalized patterns to modelers who then develop quantitative models and work with ESMs. We recommend a shift away from this historical paradigm towards a more collaborative one in which empiricists and modelers are involved in co-producing knowledge (with differing degrees of contribution) at every stage of data collection, theory development, and model integration. We also emphasize the two-way exchange of ideas, insights, and data between empirical and modeling driven activities.

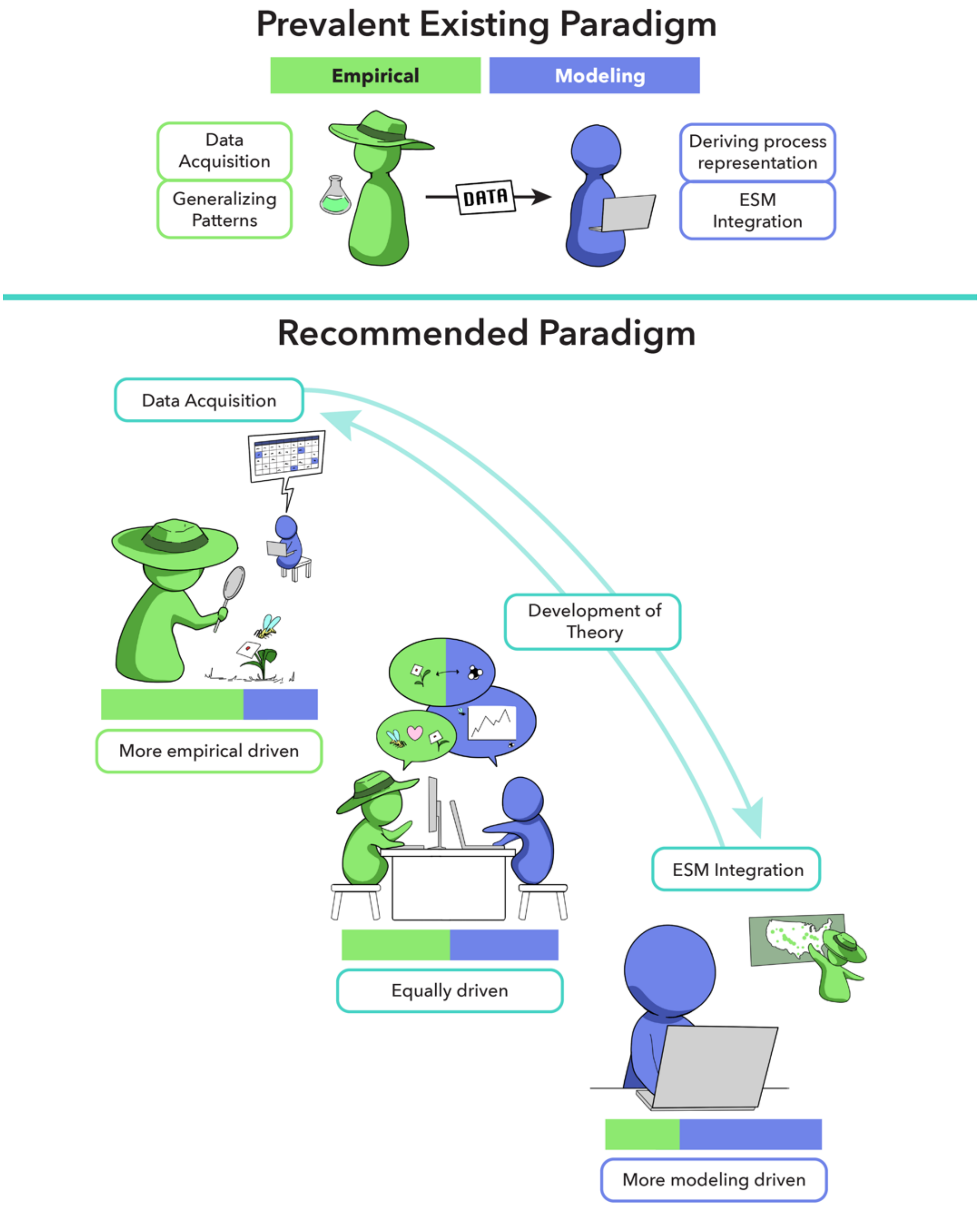


Figure 3. In the hierarchy of model development, simple models of individual processes, classes of organisms, and inorganic components (site/local scale) are often pieced together to form larger models of ecosystems and regions (ecosystem scale) and ultimately combined to form Earth system models (ESMs; global scale). Data gathered at each of these scales can be used to inform model development at the same scale.

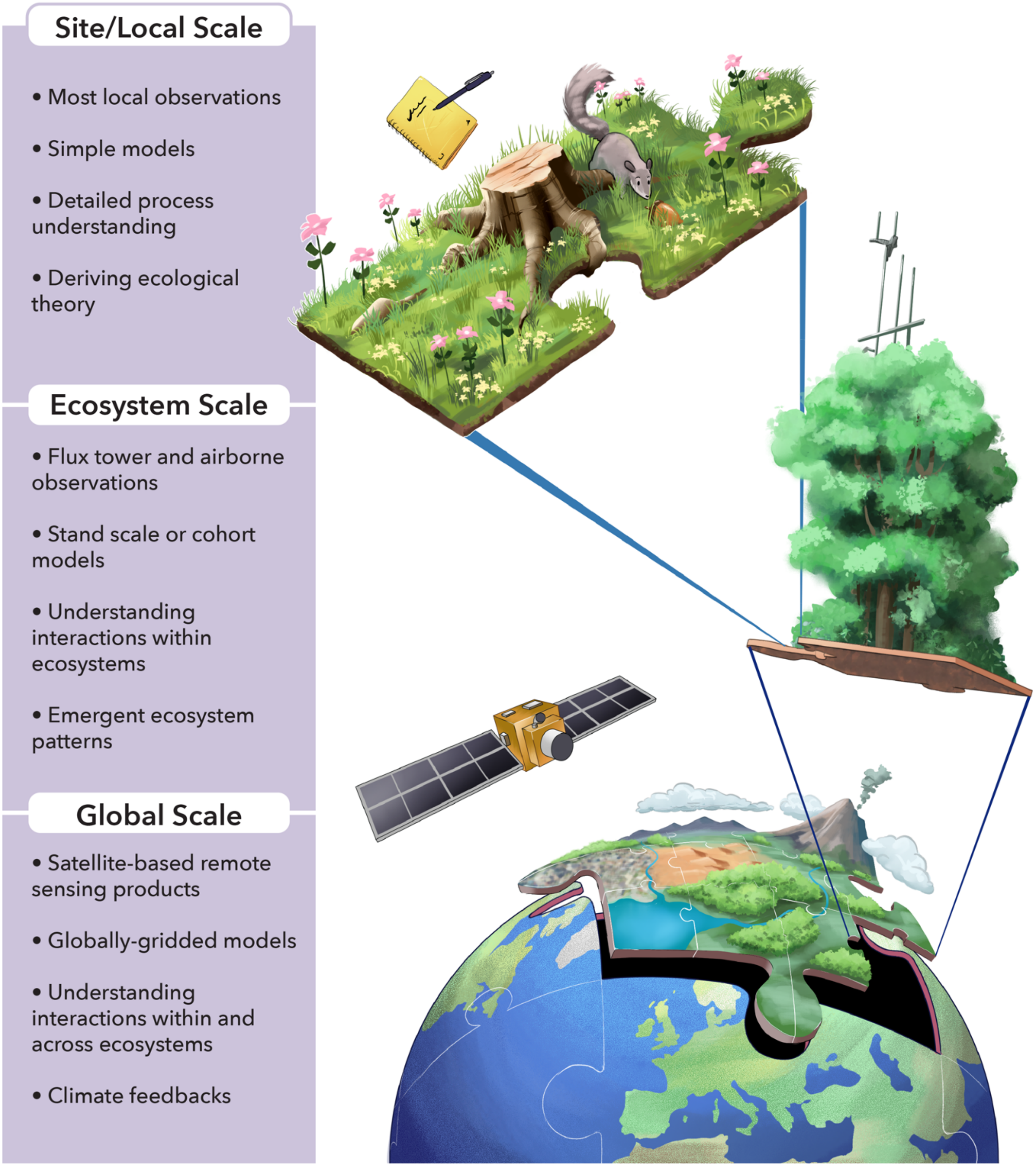
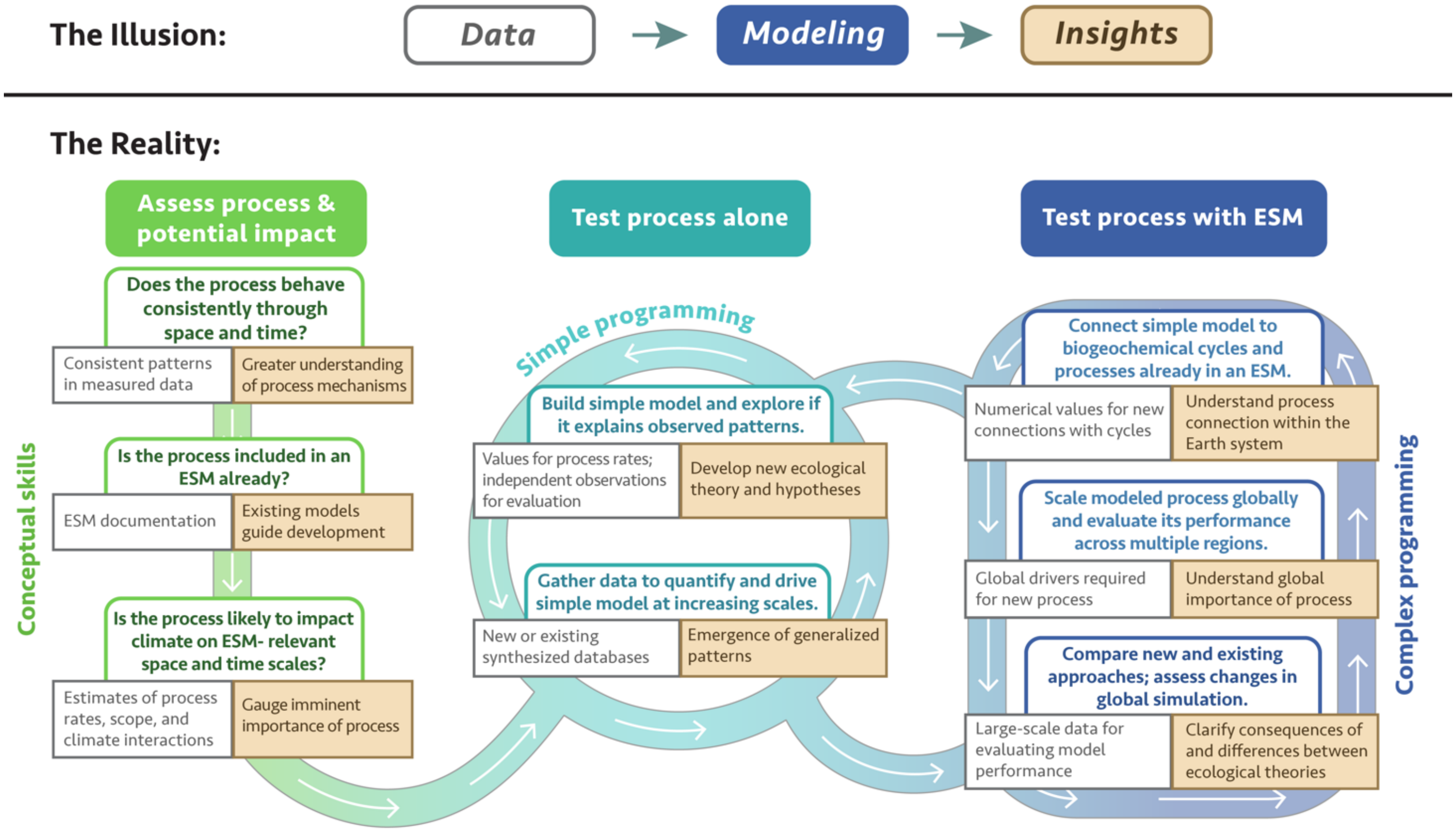


Figure 4. Although scientists sometimes think “The Illusion” (top panel) is the way that ecological concepts are integrated into Earth system models (ESMs), the reality is more like a complex metabolic cycle or eddy-filled stream, with different data inputs (gray boxes) and valuable insights (tan boxes) throughout the process. We identify three key phases in integrating a new process into an ESM, namely: “Assess process & potential impact”, which emphasizes conceptual skills (green boxes), “Test process alone”, which involves simple programming (teal), and “Test process with ESM”, which involves more complex programing (blue). Within each phase, we offer specific questions to guide empiricists and modelers along the way.



CHAPTER II: STOICHIOMETRICALLY COUPLED CARBON AND NITROGEN CYCLING IN THE MICROBIAL-MINERAL CARBON STABILIZATION MODEL VERSION 1.0 (MIMICS-CN V1.0).

Published in *Geoscientific Model Development* 13, no. 9 (2020): 4413-4434.

Authored by Emily Kyker-Snowman, William R. Wieder, Serita D. Frey, and A. Stuart Grandy.

Abstract

Explicit consideration of microbial physiology in soil biogeochemical models that represent coupled carbon-nitrogen dynamics presents opportunities to deepen understanding of ecosystem responses to environmental change. The MIcrobial-MIneral Carbon Stabilization (MIMICS) model explicitly represents microbial physiology and physicochemical stabilization of soil carbon (C) on regional and global scales. Here we present a new version of MIMICS with coupled C and nitrogen (N) cycling through litter, microbial, and soil organic matter (SOM) pools. The model was parameterized and validated against C and N data from the Long-Term Inter-site Decomposition Experiment Team (LIDET; 6 litter types, 10 years of observations, 13 sites across North America). The model simulates C and N losses from litterbags in the LIDET study with reasonable accuracy (C: R^2=0.63, N: R^2=0.29), which is comparable with simulations from the DAYCENT model that implicitly represents microbial activity (C: R^2=0.67, N: R^2=0.30). Subsequently, we evaluated equilibrium values of stocks (total soil C and N, microbial biomass C and N, inorganic N) and microbial process rates (soil heterotrophic respiration, N mineralization) simulated by MIMICS-CN across the 13 simulated LIDET sites against published observations from other continent-wide datasets. We found that MIMICS-CN produces equilibrium values in line with measured values, showing that the model generates plausible estimates of ecosystem soil biogeochemical dynamics across continental-scale gradients. MIMICS-CN provides a platform for coupling C and N projections in a microbial-explicit model but experiments still need to identify the physiological and stoichiometric characteristics of soil microbes, especially under environmental change scenarios.

1 Introduction

Soils contain the largest actively cycling terrestrial carbon (C) stocks on earth and also serve as the dominant source of nutrients, like nitrogen (N), that are critical for maintaining ecosystem productivity (Gruber and Galloway, 2008; Jobbágy and Jackson, 2000). Soil C cycle projections and their response to global change factors remain highly uncertain (Bradford et al., 2016; Todd-Brown et al., 2013), but recent empirical insights into microbial processing of soil C provide opportunities to update models and reduce this uncertainty (Cotrufo et al., 2013; Kallenbach et al., 2016; Lehmann and Kleber, 2015; Schmidt et al., 2011; Six et al., 2006). Several models have been developed recently with explicit representation of nonlinear microbial C processing dynamics, including the MIcrobial-MIneral Carbon Stabilization (MIMICS) model (Sulman et al., 2018; Wieder et al., 2014, 2015d) and others (Abramoff et al., 2017; Allison, 2014; Fatichi et al., 2019; Hararuk et al., 2015; Robertson et al., 2018; Sulman et al., 2014; Wang et al., 2013a, 2014a, 2017b). While these models serve different purposes, some can be as good as or better than models without explicit microbial pools at simulating global soil C stocks and the response of soil C to environmental perturbations (Wieder et al., 2013, 2015d), and they also predict very different long-term responses of soil C to global change (Wieder et al., 2013, 2018). Microbial-explicit models have thus furthered our understanding of C cycling in the terrestrial system, but they also provide new opportunities to explore couplings between C and nutrient cycles, especially N.

Terrestrial models that couple C and N cycles reveal important ecosystem feedbacks that are absent from C-only models. For example, across ecosystems, experimental manipulations consistently indicate that N availability limits plant productivity (LeBauer and Treseder, 2008). C-only model configurations in models typically predict that CO_2 fertilization will result in a large increase in both plant productivity and the land C sink in coming decades, but nutrient limitation may constrain the magnitude of this terrestrial ecosystem C uptake (Wieder et al., 2015c; Zaehle et al., 2015; Zaehle and Dalmonech, 2011). As terrestrial models increasingly represent coupled C-N biogeochemistry, accurate model estimates of N release from soil organic matter (SOM) will become important to reducing uncertainty in the CO_2 fertilization response of the terrestrial C cycle.

Currently, most biogeochemical models that couple C and N cycles have an implicit representation of microbial activity. These conventional models represent SOM decomposition with the assumption that chemical recalcitrance of organic matter dictates the turnover of litter and SOM pools (Luo et al., 2016). Carbon and N fluxes represented in these models are directly proportional to donor pool sizes, without any explicit representation of the microbes that mediate these fluxes (Schimel, 2001, 2013). Linear decay constants and transfer coefficients determine the flow of C and N through a decomposition cascade, and rates of N immobilization and mineralization emerge from the interaction of fixed respiration fractions and the stoichiometry of donor and receiver SOM pools. The lack of plant-microbe-soil feedbacks in these models may limit their predictive capacity, especially in the face of environmental change. For example, in these models increased plant inputs to soil only build soil C and N stocks, and plants have no way to stimulate the microbial community to mine existing SOM for N without model modifications (Guenet et al., 2016; Wutzler and Reichstein, 2013). This “N mining” or “priming” effect, where increased plant inputs result in increased microbial activity and decomposition rates, has been demonstrated in experimental studies (Cheng and Kuzyakov, 2005; Dijkstra et al., 2013; Phillips et al., 2012) and may be a critical pathway for plants to obtain more N and support increased plant productivity under elevated CO_2 (Thomas et al., 2015; Zaehle et al., 2014).

Microbes are critical mediators of soil C-N couplings and the release of plant-available N. As such, models that explicitly consider microbial activity provide an opportunity to explore potential microbial control over soil C-N biogeochemical cycling and improve simulations of patterns in ecosystem C and N. Towards this end, multiple models have been introduced that explicitly consider the role of microbial activity in ecosystem C-N interactions (Averill and Waring, 2017; Fatichi et al., 2019; Huang et al., 2018a; Schimel and Weintraub, 2003; Sistla et al., 2014; Sulman et al., 2014, 2017, 2018, 2019; Wang et al., 2017b, 2014b, 2013b). To date, the majority of these microbial-explicit C-N models have been developed to explore soil biogeochemical interactions and microbial community dynamics, while only one has been validated for N dynamics across a continental-scale gradient (Fatichi et al., 2019).

Although there is great value in exploring diverse approaches to explicitly representing microbes in purely theoretical or site-specific applications, implementing these conceptual developments within larger-scale models requires convincing evidence that adding them improves model performance against large-scale data. Recent soil model comparisons report divergent responses to simulated global change experiments among microbial-explicit model formulations, highlighting the large uncertainty in their underlying process-level representation and parameterization (Sulman et al., 2018; Wieder et al., 2018). The addition of explicit microbial pools may improve the predictive ability of landscape-scale models in the long run, but microbial models must be validated against landscape-scale datasets of a variety of pools and process rates before they can reasonably be expected to improve model performance and reduce uncertainty.

We developed a coupled C-N version of MIMICS (MIMICS-CN) to fill the need at the intersection of microbial-explicit models, coupled C-N models, models that work well enough to be considered for use in ESMs, and models that can be validated against currently available large-scale data. The C-only iteration of MIMICS considers trade-offs involved with microbial functional traits as well as both physicochemical (i.e. mineral associations) and chemical (i.e. recalcitrance) mechanisms of C stabilization in soil. (Wieder et al., 2014, 2015d) and (Sulman et al., 2018) evaluated this C only version of MIMICS across site, continental, and global scales. Here we expand on this work, introducing MIMICS-CN, which incorporates stoichiometrically coupled C and N cycling of all microbial, litter and SOM pools and stoichiometric constraints on microbial growth. Our core objectives were to: 1) Formulate a framework and parameterization for coupled C and N cycling in MIMICS; 2) Validate MIMICS-CN against a continental-scale litter decomposition dataset (LIDET) and compare MIMICS-CN to a microbially-implicit, linear model (DAYCENT); and 3) Evaluate equilibrium soil and microbial stocks and fluxes (and their parameter sensitivities) that are simulated by MIMICS-CN with data synthesized across published landscape-scale data. Our overarching goal was to create a microbial-explicit coupled C-N model of soil that balances ecological realism with the practical considerations of large-scale simulation, and to demonstrate the abilities of this model through parameterization, validation and evaluation exercises using both dynamic and equilibrium data.

2 Methods

2.1 Model formulation

MIMICS-CN builds upon the previous C-only version of MIMICS, described in (Wieder et al., 2014, 2015d), using the same pool structure for N as for C plus an additional pool for dissolved inorganic nitrogen (DIN; Fig. 1). In-depth discussion of the reasoning behind the development of the C-only version of the model is available in these previous publications, but the general intent behind the development of MIMICS was to incorporate a simplified representation of the important aspects of microbial communities (biomass-dependent control of process rates, diversity in life history strategies and physiological parameters) into a soil model that stabilizes organic matter through both physical (mineral-associated, protected from microbial decomposition) and chemical (recalcitrance-based, vulnerable to microbial decomposition) means. The C-only version of the model represents C flows through seven pools (Fig. 1): two litter pools, two microbial pools, and three SOM pools. Litter inputs to the model are partitioned into structural litter (LIT_s) and metabolic litter (LIT_m) pools based on estimates of litter quality for different biomes (Brovkin et al., 2012).

Temperature-sensitive forward Michaelis-Menten kinetics determine the flux of litter and SOM through microbial biomass pools that determine rates of organic matter decomposition, SOM formation, soil respiration and nitrogen mineralization fluxes. The microbial functional groups are intended to broadly capture tradeoffs in microbial growth rates and growth efficiency, with rapidly-growing microbial decomposers (low efficiency, r-strategist (MIC_r)) and slower-growing microbial decomposers (higher efficiency K-strategist (MIC_K; Wieder et al., 2015d)). In MIMICS-CN we extend these microbial physiological traits to include microbial stoichiometry and assume that the higher metabolic capacity of MIC_r also require more nitrogen and, thus a lower microbial biomass C:N ratio. Fluxes of C into microbial pools result in respiration losses according to a defined carbon use efficiency (CUE) that varies by microbial functional group and substrate quality (e.g. structural or metabolic litter). Microbial pool sizes are moderated by inputs, CUE, and biomass-specific turnover rates. We implemented density-dependent microbial turnover (sensu Georgiou et al., 2017; see Appendix A) for this iteration of the model to make microbial pools behave realistically in response to small changes in C inputs (Wang et al., 2014b, 2016). The density-dependent turnover of microbial biomass dampens the oscillatory response of microbial biomass to perturbations.

Microbial biomass turns over into physicochemically-stabilized (SOM_p), chemically-stabilized (SOM_c), and a pool that is 'available' for microbial decomposition (SOM_a). We consider the SOM_p pool to mostly consist of low C:N organic matter that is primarily composed of microbial products that are adsorbed onto mineral surfaces (e.g. mineral-associated organic matter, MAOM; Grandy and Neff, 2008). By contrast, the low-quality SOM_c pool consists of decomposed or partially decomposed litter that has more structural C compounds, such as lignin, and a higher C:N ratio (e.g. particulate organic matter, POM). Finally, the SOM_a is the only SOM pool that is available for microbial decomposition; it contains a mixture of fresh microbial residues, products that are desorbed from the SOM_p pool (e.g. Jilling et al., 2018), as well as depolymerized organic matter from the SOM_c pool. We do not specifically consider soil aggregates, but we recognize that in some soils they are an important component of accruing and maintaining persistent organic matter.

The current representation of N cycling in MIMICS-CN is based on the threshold element ratio idea described in (Sinsabaugh et al., 2009) and (Mooshammer et al., 2014) whereby organisms maintain biomass stoichiometry by spilling excess C or N on either side of a threshold ratio. We modified the C-only iteration of MIMICS to include N by adding a parallel set of pools and fluxes for N, as well as a pool for inorganic N (Fig. 1). The C cycle drives decomposition with fluxes from litter and SOM pools to microbes based on biomass-C-based forward Michaelis-Menten kinetics. Parallel N fluxes are determined by the C:N ratio of the donor pools, which is a fixed parameter for the metabolic litter pool, varies with litter input chemistry for the structural litter pool, and depends on inputs for SOM pools. We use a fixed C:N of 15 for metabolic litter inputs, while the C:N of structural litter was allowed to vary to ensure conservation of total N inputs from litterfall (Table 1).

The coupling between C and N cycles in MIMICS-CN occurs in the microbial biomass: at each hourly time step, the total C and N in incoming fluxes available to microbes is summed and adjusted based on the C use efficiency (CUE; varies with microbial functional group and substrate) and N use efficiency (NUE; set to 0.85 for all fluxes entering microbial biomass pools in this model iteration). If the C:N of substrates being assimilated by microbial functional groups is greater or less than the C:N of the microbial biomass (defined as 6 and 10 for r- and K-strategists, respectively; Table 1), the microbes will spill excess C or N to maintain their biomass stoichiometry through overflow respiration or excess N mineralization. In MIMICS-CN the C:N ratio of SOM pools is flexible and determined by the inputs from microbial residues and direct inputs from litterfall fluxes (f_i; Fig. 1). All N fluxes into microbial pools leak a small quantity of N into a dissolved inorganic N pool (DIN) based on the model-defined NUE. At each time step, each microbial functional group can access a fraction of the inorganic N pool proportional to their fraction of total microbial biomass. Plant N uptake and ecosystem losses (both hydraulic and gaseous) of inorganic N are handled implicitly at this stage, with a fixed fraction (20%) of DIN leaving the soil component model every time step.

2.2 Model parameterization and validation: Cross-site litter decomposition

We parameterized and validated MIMICS-CN using C and N dynamics observed across multiple sites participating in the 10-year Long-Term Intersite Decomposition Experiment Team (LIDET) experiment (Adair et al., 2008; Harmon et al., 2009; Parton et al., 2007). The LIDET study selected standardized plant litter types with a range of litter quality (lignin and N concentration), placed litterbags containing 100 g of each litter type at sites across a continental scale gradient of climatic conditions, and measured changes in the C and N in litterbags on an approximately annual basis for 10 years. Although the original dataset included 27 sites across North America, we utilized data from 14 sites ranging from Alaska to Puerto Rico based on the data available at those sites to drive MIMICS (see (Wieder et al., 2015d) for site information). We focus our analysis on six leaf litters that were simulated across all sites that have been used previously to evaluate litter decomposition dynamics in terrestrial models (Bonan et al., 2013; Parton et al., 2007; Wieder et al., 2015d). Root litter types included in the original LIDET experiment were not included. The LIDET dataset is a robust appraisal of the impacts of climate and litter chemistry on litter decomposition and has been used as a dataset for comparing models of soil and litter decomposition in the past (Bonan et al., 2013). MIMICS has been used previously to simulate C losses in the LIDET study (Wieder et al., 2014, 2015d).

We parameterized MIMICS-CN using observations from Harvard Forest in Petersham, MA, USA. Observations included both litterbag C loss and N data from the LIDET study as well as measurements of soil C and N stocks and microbial C and N from other studies at Harvard Forest (Colman and Schimel, 2013). Multiple combinations of parameters produced equally good fits to litter decomposition data; thus ancillary data on soil and microbial C stocks were used to inform the parameter values presented here (Table 1). These ancillary data were not reported in LIDET and were not measured on identical plots to those used for the LIDET study (Harvard Forest encompasses multiple experiments and ecotypes), but these general targets were useful in distinguishing among model parameterizations. Our general targets for stocks at Harvard Forest included soil C and N (0-5 cm mineral soils, coniferous stand): 61 mg C cm^{-3} and 2.9 mg N cm^{-3}; soil C:N: 21; and microbial biomass: 0.61 mg C cm^{-3} (estimated as 1% of soil C based on (Xu et al., 2013)).

After parameterizing the model to match observations at Harvard Forest, the model was validated using data from the remaining LIDET sites. To represent litterbags in MIMICS-CN, we first spun up the underlying model to simulate steady-state soil C and N pools and fluxes across sites in the LIDET study using site-level measurements of mean annual temperature, clay content, and litter input quantity, and litter chemistry (Wieder et al., 2015d). Then, we added a pulse of metabolic and structural litter based on the type of litter in the simulated litterbag. We tracked the C and N across all model pools for 10 years and calculated the C and N in litterbags as the difference between total model C and N in the simulations and total model C and N at steady state. In both the simulated and real litterbags, microbes immobilized N from the soil DIN pool, resulting in litterbag N contents for some time points in excess of the initial values. For each site, the model was sampled at time points equivalent to the real data collection dates in LIDET (approximately annually). Observed and modeled values of C and N in litterbags were compared by calculating R^2, root mean square error (RMSE) and bias.

To contextualize our results and better understand how our model functions compared to a widely used microbial-implicit model, we compared MIMICS-CN simulations of LIDET data against DAYCENT (Bonan et al., 2013) simulations of the same data. (Bonan et al., 2013) used the full complement of 27 LIDET sites in their analysis, but here we subset those results for the

13 sites used in the MIMICS-CN validation. We calculated R^2, RMSE and bias in the same way for each model and compared results across models, grouping results by biome.

2.3 Model evaluation: Equilibrium C and N cycling

Building on the LIDET simulations, we independently synthesized observations to evaluate the patterns of C and N pools and fluxes across a variety of sites. Although direct, site-specific comparisons of modeled and observed values like microbial biomass would have been ideal, MIMICS-CN represents many variables that were not measured in the LIDET study and have not been synthesized across these Long-Term Ecological Research sites. Instead, we compared the range and distribution of pools (soil organic C and N, microbial biomass C and N, and total inorganic N) and fluxes (heterotrophic respiration and N mineralization) using the modeled LIDET simulations and published syntheses of observations from other sites (Cleveland and Liptzin, 2007; Colman and Schimel, 2013; Xu et al., 2013; Zak et al., 1994). To more directly compare measurements with model results, stock measurements were converted to units of % of soil mass and fluxes (heterotrophic respiration and net N mineralization rates) were converted to units of $\mu g\ cm^{-3}\ hr^{-1}$. MIMICS reports pool values in units of $g\ cm^{-2}$ (0-30 cm); to compare MIMICS against observations we converted MIMICS values to % by mass assuming a bulk density of 1.5 $g\ cm^{-2}$. Soil depth simulated by MIMICS (30 cm) is deeper than most of the observations in the compiled dataset, but the purpose of this exercise was to evaluate whether MIMICS produces realistic values for soil biogeochemical stocks and fluxes across continental-scale ecoclimatological and edaphic gradients, rather than making a direct site-specific comparison. The distribution of values produced by MIMICS across the LIDET sites was superimposed on the distributions of observed values to illustrate data-model agreement and to visualize the median and range of measurements across studies.

Finally, we documented relationships between model input variables (mean annual temperature, productivity, clay content, and litter quality) and the distribution of SOM pools that were simulated at the LIDET sites. Our aim with these analyses was to illustrate the underlying assumptions in the model and how they influence the size and distribution of C across SOM pools. Specifically, we wanted to explore how assumptions made in the model structure and parameterization of MIMICS determine the quantity and distribution of SOM pools, and how they change among sites with variation in climatic, biological, and edaphic properties. To do this we looked at the absolute and relative contributions of each SOM pool simulated by MIMICS across the LIDET sites and conducted linear regressions to determine how environmental factors control their distributions. We also conducted linear regressions between soil C:N and both litter chemistry and environmental factors to assess the drivers of soil C:N in the model.

3 Results

3.1 Model parameterization and validation: Cross-site litter decomposition

We parameterized MIMICS-CN to replicate litter C decay rates and N dynamics of six litter types observed in the LIDET study at the Harvard Forest LTER site (Fig. 2). In its current parameterization, MIMICS slightly overestimates litter C loss at later stages of decay, but most time points are within uncertainty estimates of the observations (Fig. 2a). Similarly, for N, MIMICS-CN overestimates N accumulation in early stages of decay and underestimates N remaining at later stages, but most time points follow a reasonable trajectory given observations. MIMICS-CN also captures the effects of litter quality on both rates of litter decay (Fig. 2a) and litterbag N accumulation (Fig. 2b). The parameters we used to fit MIMICS-CN to Harvard Forest data also produce reasonable estimates of soil N stocks (2.0 vs. 2.9 mg N cm^{-3} for model and observations, respectively) and microbial biomass (0.65 vs 0.61 mg C cm^{-3}), although estimates of soil C (21 vs 61 mg C cm^{-3}) and soil C:N (11 vs. 21) are both lower than observations.

Parameter values used for this and subsequent simulations across all LIDET sites are shown in Table 1. Relative to the previous C-only version of the model (Wieder et al., 2014, 2015d), kinetic parameters and microbial turnover values were adjusted to account for density-dependent turnover (Georgiou et al., 2017). In addition, the fraction of structural litter that bypasses microbial biomass to enter the chemically-protected pool (f_i) was increased from 5% to 30% as a means to produce reasonable values for total soil C:N. Finally, we adjusted the partitioning of microbial turnover to stable soil pools in order to more closely match distributions at Harvard Forest.

Applying this parameterization across all six litter types at 13 LIDET sites, MIMICS-CN simulates C losses and N dynamics from litterbags with an R^2 of 0.63 and 0.29, respectively (Fig.

3). MIMICS-CN captures effects of litter quality on decay rates, with faster rates of C loss and more rapid N mineralization simulated with more N rich *Drypetes glauca* litter, and slower rates of C loss and greater N immobilization simulated by low quality *Triticum aestivum* litter (Fig 3a, c). MIMICS-CN is best at capturing C loss rates in high- and intermediate-quality litters (*Drypetes glauca*, *Pinus elliottii, Thuja plicata*, and *Acer saccharinum*) but tends to underestimate litter C loss rates from the lowest-quality litter (*T. aestivum*). For N immobilization and loss, the model performs well especially for high-quality litters but underestimates N accumulation slightly in the lowest-quality litter. The model also captures broad climate effects on litter C loss, with slower decay rates in tundra and boreal forests sites and faster decay in tropical and deciduous forests (Fig 3b).

MIMICS-CN and DAYCENT simulations of LIDET decomposition data are compared in Table 2. Across a broad range of biomes, MIMICS-CN and DAYCENT both show good agreement with LIDET observations. Across sites MIMICS-CN has similar R^2 and RMSE values but lower bias compared to DAYCENT for mass loss (MIMICS-CN: R^2=0.63, RMSE=16.0, bias=-0.12; DAYCENT: R^2 = 0.67, RMSE=14.4, bias=4.73), and percent N remaining (MIMICS-CN: R^2=0.29, RMSE=0.34, bias=0.03; DAYCENT: R^2=0.30, RMSE=0.40, bias=0.08). Broadly, MIMICS-CN outperformed DAYCENT in the warmest biomes while DAYCENT excelled for colder sites for both C and N (Table 2), but the differences in model fit to data were slight and would be difficult to attribute to any particular differences in model structure. DAYCENT simulates decomposition based on initial litter chemistry and showed no site-specific effects on the maximum N immobilized or the relationship between C and N during decomposition for a given litter type (Fig. S1 and S2). By contrast, the amount of N that can be immobilized by a litterbag in MIMICS-CN is driven by the availability of N and the stocks and flows of N in the simulated steady-state soil, and MIMICS-CN showed site-specific variability in the shape of N immobilization and loss curves (Fig. 3 and 4).

Litter quality determines the timing of N immobilization vs. mineralization in observations. This produces a functional relationship between initial litter chemistry, C loss, and N immobilization / mineralization that is fairly consistent across sites (colored dots; Fig. 4). MIMICS-CN broadly captured litter quality effects on the timing and magnitude of N immobilization and mineralization dynamics across all biomes (red triangles; Fig 4). For example, litter with high initial chemical quality consistently mineralize N throughout all stages of litter decay, and MIMIC-CN adequately captures this functional C-N relationship (Fig 4a,b). By contrast, litters with lower initial chemical quality immobilize N during early stages of litter decay, but subsequently mineralize N as decomposition proceeds. MIMICS-CN broadly captures these patterns, but without as much variation as the observations (Fig 4c-f). The lowest-quality litter (*Triticum aestivum*) immobilizes N until only 40% of C remains in litterbags. Although MIMICS-CN potentially underestimates total N immobilization *Triticum aestivum* litter, it does capture the point at which net N mineralization begins (Fig. 4f).

3.2 Model evaluation: Equilibrium C and N cycling

Across all sites and litter types in the LIDET simulations, the ranges of underlying pool sizes and process rates in MIMICS-CN were compared against published ranges from similarly diverse sets of sites (Cleveland and Liptzin, 2007; Colman and Schimel, 2013; Xu et al., 2013; Zak et al., 1994). MIMICS-CN simulations produced reasonable equilibrium values for most pools and fluxes (Table 3 and Fig. 5). In general, the range of values across the 13 sites simulated by MIMICS was smaller than the ranges across the thousands of sites included in the compiled dataset of observations. For example, total soil C ranged from 7.0-50 mg C cm^{-3} in MIMICS simulations but ranged from 2.7-610 mg C cm^{-3} in observations. Despite this discrepancy, the median values of the simulations and observations were generally within reason (Fig. 5). The distributions of measured and modeled values for microbial biomass C and N as a percent of total soil C and N overlapped, providing evidence that the model reasonably represents microbial stoichiometry, microbial activity as a function of biomass, and microbial biomass as a function of SOM. For soil C:N, the model tended to produce low values with a relatively narrow range, relative observed values.

Finally, we explored the environmental controls on the distribution of SOM across physicochemically-protected, chemically-protected, and available pools in MIMICS-CN by examining the correlations between pool sizes and salient input variables (mean annual temperature, productivity, clay content, and litter lignin content). The results are shown in Fig. 6. The absolute concentration of SOM simulated across the LIDET sites was most strongly correlated with ANPP (R^2=0.52), but also tended to increase with MAT, albeit inconsistently (Fig. 6a; R^2=0.15). The distribution of SOM across stabilized pools strongly favored chemically-protected SOM at sites with lower temperatures, while the relative proportion of physicochemically-protected SOM increased with increasing temperature (Fig. 6b). The relative proportion of SOM in the available pool remained fairly consistent across simulated sites. Physicochemically-protected SOM was tightly positively correlated with the product of ANPP and clay content (R^2=0.96, Fig. 6c), while chemically-protected and available SOM were negatively correlated with MAT (Fig. 6d, R^2=0.40 and 0.47, respectively) and positively correlated with litter lignin content (Fig. 6e; R^2=0.68 and 0.32, respectively). The C:N of individual pools was fairly consistent across sites and tended to be higher for chemically-protected SOM (~15) than available (~8) or physicochemically-protected SOM (~10). As a result, soil C:N was largely driven across sites by the distribution of SOM across pools, especially the absolute size of the SOMp pool (Fig. 6f, R^2=0.79). Given that clay content was an important driver of physicochemically-protected SOM in the model, clay content was tightly correlated with soil C:N (R^2=0.88). Other litter characteristics and environmental factors were not strong drivers of soil C:N (R^2 for MAT: 0.42; litter lignin: 0.03; litter C:N: 0.005).

4 Discussion

Terrestrial models are increasingly representing coupled C-N biogeochemistry, and MIMICS-CN is among the first attempts to do so with a microbial explicit soil biogeochemical model that can be used to project C and N dynamics across continental-scale gradients. Our formulation and parameterization of MIMICS-CN captures site level observations of litter C loss and N immobilization at the Harvard Forest LTER site (Fig. 2). Cross-site validation of the model demonstrates that it broadly captures climate and litter quality effects on rates of C and N transformations from the LIDET observations (Figs. 3-4). Notably, the results simulated by MIMICS-CN represent N dynamics during litter decomposition about as well as a first-order model that implicitly represents microbial activity (Table 2). It also generates steady state pools and fluxes of C and N that seem reasonable compared to published syntheses (Table 3; Fig. 5). Below we discuss these dynamic and equilibrium model simulations in greater detail, as well as some of the limitations of MIMICS-CN that will be addressed in future work.

4.1 Model parameterization and validation: Cross-site litter decomposition

We first parameterized and validated MIMICS-CN using the cross-site litter decomposition study, LIDET. Previous LIDET simulations using MIMICS have successfully replicated observed C loss patterns, and adding coupled N cycling to MIMICS neither improved nor degraded simulations of LIDET litter C losses relative to the C-only model (Figs. 2-3; (Wieder et al., 2015d) report global RMSE for the C-only model = 14.6 vs. 16.0 in this study). Our results show higher than observed rates of litter C mass loss in deciduous and coniferous forest (Figs 2a, 3b; Table 2). This suggests that the partitioning of plant detrital inputs into litter pools that are chemically defined works well for initial stages of litter decay, but may not consider the changes in substrate chemistry or microbial community succession that occur in later stages of decomposition that slow rates of mass loss (Berg, 2000; Melillo et al., 1989). Models that implicitly represent microbial activity capture this phenomena by using a three pool structure (Adair et al., 2008), and future studies can consider how to more mechanistically understand interactions between initial litter quality, decomposer communities, climate, nutrient availability and late-stage litter decay rates (e.g. Craine et al., 2007; Hobbie et al., 2012; Wickings et al., 2012) in models like MIMICS-CN. In MIMICS-CN, carbon and nitrogen move together through model pools, but model dynamics are primarily driven by C, with N dynamics following suit based on pool stoichiometry. The N dynamics do, however, constrain C cycling in the model if microbes are N-limited, in which case microbes lose excess C through overflow respiration. At equilibrium, microbes in our MIMICS-CN simulations primarily obtained N through recycling of SOM pools with favorably low C:N ratios, with the result that modeled microbes were almost always C-limited at equilibrium and rarely exhibited overflow respiration. Large pulses of low-quality litter can perturb this equilibrium and induce N limitation, but in the absence of losses of or plant competition for inorganic and dissolved organic N, C cycling in MIMICS proceeds in essentially the same way with or without accounting for N.

MIMICS-CN accurately captured the stoichiometric relationships between C and N during litter decomposition (Fig. 4). This stoichiometric relationship has been well-defined in the past using theoretical microbial stoichiometry and CUE (Parton et al., 2007), but comparable soil models without explicit microbial physiology have tended to over-predict N accumulation in litterbags (Bonan et al., 2013). Moreover, models without microbial explicit physiology also show N immobilization mineralization dynamics that are completely determined by initial litter quality, whereas MIMICS simulations show greater site-level variation (Figs. 4, S2). In

MIMICS-CN, stoichiometric relationships drive litterbags to accumulate soil N until they reach a threshold C:N, after which litterbags become net sources of N. This threshold, representing the balance between microbial N requirements and availability, is a function of changes in litter stoichiometry during decomposition, as well as of the stoichiometry of microbes and their nutrient use efficiencies. By explicitly considering these dynamics MIMICS-CN has a similar or lower RMSE for N remaining in litter bags than a model that implicitly represents microbes, DAYCENT (Table 2).

MIMICS-CN and DAYCENT capture N dynamics during decomposition with similar overall degrees of fit, but for different reasons. In DAYCENT, N immobilization and loss dynamics are driven by initial litter chemistry, and good model fit to data is achieved by capturing the average N immobilized for a given litter type regardless of biome and climate conditions (see Fig. S1 and S2). By contrast, litterbag N immobilization in MIMICS-CN is driven by the availability of N in the underlying modeled soil and by site-specific effects (e.g. climate, clay content) on the simulated stocks and fluxes of N. As a result, MIMICS-CN generates greater variation in the amount N immobilized for a given litter type across sites (Figs. 3 and 4). Site-specific variability in N immobilization patterns is also clearly visible in LIDET observations (colored dots, Fig. 4), but the introduction of site-specific variability in MIMICS-CN does not substantially improve model fit to data relative to DAYCENT. Spatial variability in ecosystem processes, like N mineralization rates, may be linked to factors like local-scale microbial community composition, soil moisture, or mineralogy (Doetterl et al., 2015; Graham et al., 2016; Smithwick et al., 2005; Soranno et al., 2019). While more work needs to be done to understand the factors controlling within and among site variation in soil C-N dynamics (Bradford et al., 2017), these results highlight that the explicit representation of microbial activity in MIMICS-CN may present opportunities to explore factors responsible for biogeochemical heterogeneity across scales.

Although MIMCS-CN broadly captures appropriate climate and litter quality effects on leaf litter decomposition patterns, the model underestimates N accumulation in the highest C:N ratio litter (*Triticum aestivum*; Fig. 4f). Microbes in MIMICS-CN recycle nitrogen from necromass and necromass-derived SOM, which might allow microbes to scavenge the N required to decompose high C:N litter without having to accumulate it from the inorganic soil pool. In a real litterbag, necromass might be lost through leaching and microbial access to recycled biomass might be limited, and some microbial-derived compounds may require extensive depolymerization and proteolysis before the N is available for recycling (Schulten and Schnitzer, 1997), thus favoring N uptake from the soil pool. Alternatively, N inputs to real litterbags in the LIDET study may have come from atmospheric deposition or other unintended sources that MIMICS-CN does not address. Nonetheless, the high C:N ratio of *Triticum aestivum* is not typical of the majority of litter inputs across diverse biomes (Brovkin et al., 2012) which are well within the range that MIMICS-CN can simulate.

4.2 Model evaluation: Equilibrium C and N cycling

We conducted additional model evaluation by comparing model pools and fluxes at equilibrium to published observations. The parameter values used in the LIDET simulations produced reasonable estimates of equilibrium pools (soil organic C and N, microbial biomass C and N, and total inorganic N) and fluxes (heterotrophic respiration and N mineralization) (Table 3; Fig. 5). In combination with the LIDET results, these results indicate that MIMICS-CN can produce realistic simulations of both the short-term dynamic processes involved in litter decomposition and the soil-forming processes that produce equilibrium pools and fluxes over much longer time scales. In addition, MIMICS-CN simulates microbial stoichiometry, microbial growth and turnover, and microbially-mediated decomposition, rather than using prescribed values as in models that lack explicit representation of microbes. This increases the power of MIMICS-CN to explore the microbial and biogeochemical processes underpinning model predictions.

Continent-wide observation of soil pools and fluxes range over several orders of magnitude (Table 3), but MIMICS simulations agreed well with the median of those ranges. Observations tended to be spread over a much larger range of values than the MIMICS-CN simulations, but these simulations only included information from 13 sites while the observations included thousands of locations. The median values of observed and simulated values were within a factor of 2.5 for all pools (Fig 5). Differences in measurement depth or error in estimated bulk density values could account for some of the differences between measurements and simulations and for the spread across observed values. This is less of a concern for three of the variables used here (soil C:N, microbial biomass C as a percent of total soil C and microbial biomass N as a percent of total soil N), which are ratios that are comparable across sites. Microbial biomass C as a percent of total soil C and microbial biomass N as a percent of total soil N were highly conserved across sites, relative to soil stocks or microbial C or N, and may be particularly useful metrics for evaluating microbial explicit soil biogeochemical models since the size of the microbial biomass pool directly controls rates of SOM turnover and formation in models like MIMICS-CN. For these ratios, MIMICS-CN reproduced distributions and median values that overlapped well with observations. In future work, direct comparisons of modeled and measured values for these ratios at specific sites may shed light on the limitations of the model and the origins of data-model disagreement. However, even the simple range comparisons included here provide evidence that the mechanistic representation of soil biogeochemistry in MIMICS-CN is ecologically realistic. Examinations of model realism like this are a crucial step in transitioning from theory and small-scale model tests to applications in ESMs or at larger scales where evaluation data are more sparse.

Besides representing appropriate soil biogeochemical stocks, fluxes simulated by the models also agree well with observations. Specifically, MIMICS-CN simulations of heterotrophic respiration and net N mineralization rates fell within observed bounds, although the variation in observations was much greater than the variation in simulated values. Our simulations calculated rates at equilibrium assuming constant temperature and other factors, while real rates of these processes are driven by seasonally- and diurnally-variable temperature, soil moisture, and other factors, so predictably, our simulations produced smaller-than-observed variability in rates. MIMICS-CN produced total soil C:N values that fall within observed ranges, although observations again show greater variation of soil C:N ratios and have maximum values that are much higher than the maximum C:N ratios simulated by MIMICS-CN. SOM pools in MIMICS-CN are mostly comprised of microbial necromass, in addition to a small proportion of litter that enters SOM pools directly without first passing through microbial biomass. Increasing this proportion in the model is one way to increase the C:N of SOM pools and the overall system at equilibrium. At some sites, litter may contribute more directly to SOM pools than microbial necromass (Jilling et al., 2018). For example, forests often have a higher proportion of total soil C in the light fraction, which is almost entirely made up of plant residues, compared to agroecosystems and many grasslands (Grandy and Robertson, 2007). For those sites with large, direct contributions of plant matter to SOM, increasing the fraction of litter that passes directly into SOM in MIMICS may be appropriate.

4.3 Exploring emergent SOM dynamics

The distribution of SOM across simulated pools in MIMICS-CN (Fig. 6) illustrates how model-defined assumptions about pool stabilization mechanisms drive potential responses to environmental variables. The wide variation in SOM pool distributions among contrasting environments in our simulations provides support for experimental efforts aimed at distinguishing between SOM pools to understand SOM responses to environmental changes and potential ecosystem feedbacks. For example, global change factors like warming can cause a range of different responses among SOM pools (Conant et al., 2008; Li et al., 2013; von Lützow and Kögel-Knabner, 2009; Plante et al., 2010). Experimental studies also show that increases in SOM resulting from increased inputs are not typically evenly distributed across different SOM pools (Lajtha et al., 2017; Stewart et al., 2009), which can influence feedbacks to productivity as well as the persistence of soil C gains in response to shifts in climate. Thus, while our broad-scale projections of how and why SOM differs among pools needs to be evaluated with experiments and data synthesis across environments, they can provide a starting point for understanding SOM responses to global change factors across environments.

In MIMICS, the turnover of chemically-protected and available SOM pools is based on temperature-sensitive Michaelis-Menten kinetics and litter chemistry (the latter controlling allocation of litter pools to the different microbial functional groups). This results in SOM_C pools (analogous to light fraction or POM pools) that are negatively correlated with MAT and positively correlated with litter lignin content (Fig. 6d, 6e). Turnover of the physicochemically-protected SOM pool, on the other hand, occurs via first-order kinetics with a rate constant modified by clay content, and the equilibrium values of this pool are determined by inputs that largely come from microbial biomass and biomass turnover rates (Fig. 1). Therefore, the equilibrium values of SOM_p (analogous to heavy fraction or MAOM pools) were strongly positively correlated with the product of ANPP and clay content (Fig. 6c). This relationship broadly reflects the expected importance of total soil C inputs and their potential to be preserved after microbial processing by association with clays (Kleber et al., 2015). However, these two variables are also likely to covary with others, especially MAT, highlighting the difficulty of isolating individual mechanisms that regulate SOM.

Across the sites included in these simulations, chemically-protected SOM formed a higher proportion of total SOM at lower MAT, while physicochemically-protected SOM was favored at warmer sites (Fig. 6b). In global simulations with the carbon-only version of MIMICS, these assumptions result in MIMICS projecting longer soil C turnover soil C times and larger soil C pools in the tropics than other models (Koven et al., 2017; Wieder et al., 2018) and a higher vulnerability of high latitude soil C stocks (Wieder et al., 2015d, 2019). Evaluating the accuracy of our model assumptions and the resulting patterns in soil C and N cycling requires coupling process-level studies of the fate of decomposing litter (e.g. using isotope tracers) to broad-scale evaluation of SOM pool distributions across environmental gradients.

Soil C:N ratios simulated by MIMICS-CN across sites were highly correlated with soil clay content ($R^2=0.88$), suggesting that, in the model, soil stoichiometry emerges from the relative contributions of SOM across physicochemically- and chemically-protected pools (Fig. 6). Although the spread of C:N values across the sites simulated by MIMICS-CN was small (Fig. 6f), C:N tended to decrease with increasing temperature, and simulated soil C:N was more correlated with site temperature (R^2=0.42) than any of the litter characteristics used to drive the model, such as litter lignin (R^2=0.03) or litter C:N (R^2=0.005). This result directly contradicts a recent study using a first-order linear model which presumed that litter quality and soil quality at equilibrium were directly proportional (Menichetti et al., 2019). Although many soil biogeochemical models prescribe soil C:N ratios for individual pools, the stoichiometry of SOM in MIMICS-CN is an emergent property of the model.

The lack of correlation between simulated soil C:N and litter C:N in MIMICS-CN simulations suggests an intriguing follow-up question: in the field, is SOM stoichiometry correlated with litter quality, or is it better explained by climate, edaphic, and mineralogical gradients that impact soil microbial community composition, microbial activity, and mineral-mediated mechanisms of SOM persistence? Various regional studies provide limited support for the relationships generated by MIMICS-CN between soil C:N and MAT (Miller et al., 2004) or clay content (Hassink et al., 1993; Homann et al., 2007; Jenny, 1941), though a large-scale synthesis of measurements across all of these variables is still needed. Presently, MIMICS-CN assumes that microbial biomass stoichiometry largely controls the C:N ratios of stable SOM, with relatively minor contributions from litter quality. However, a small proportion of litter inputs become stabilized in MIMICS-CN without first passing through the stoichiometric filter of microbial biomass, and increasing this fraction in the model is a means to increase the C:N of simulated stable SOM. The strength of the mineral sink for microbial necromass in the model also impacts the relative balance of microbe- or plant-derived stable SOM, which in turn impacts modeled soil C:N. This result implies that in the field, C:N stoichiometry might be used as a means to differentiate the degree to which a given soil fraction is derived from direct plant inputs or microbial biomass, and mineralogical variables might be useful for explaining differences in fraction distributions across soils that impact C:N. Studies like (Mikutta et al., 2019) illustrate the way that C:N can be used to assess the relative contributions of plant matter or microbial residues to stable SOM. Future work will use measured C:N of soils and soil fractions and isotopic insights into the plant or microbial origins of stable SOM to improve the parameterization of this aspect of the model and better understand the relationship between mechanisms of SOM stabilization and soil stoichiometry.

4.4 Limitations and future work

MIMICS-CN combines reasonable biogeochemical simulations with the option to explore underlying microbial processes, but limitations remain. For example, MIMICS only represents two microbial groups with different stoichiometric and physiological parameters, but real soils contain a much more diverse array of microbial functional groups with different responses to environmental conditions and different couplings between C and N cycles. CUE and NUE are critical microbial parameters in MIMICS-CN, but the relationships between CUE and microbial community composition (Maynard et al., 2017), microbial growth rate (Molenaar et al., 2009; Pfeiffer et al., 2001), temperature (Allison, 2014; Dijkstra et al., 2011; Frey et al., 2013; Steinweg et al., 2008), substrate quality (Blagodatskaya et al., 2014; Frey et al., 2013; Sinsabaugh et al., 2013), or any number of other aspects of microbial metabolism are complex, difficult to quantify, and challenging to represent at the scale of a whole soil community (Geyer et al., 2016). In its current configuration, MIMICS-CN also simplifies a number of ecosystem biogeochemical processes, and there are several important pathways of N cycling currently absent from the model. For example, MIMICS-CN does not currently represent free living biological N fixation, direct mycorrhizal exchanges for plant C for microbial N, dissolved organic C or N losses, denitrification/nitrification/other inorganic N transformation and loss pathways, plant uptake of N, or inorganic N leaching beyond a simple linear decay rate. Some of these shortcomings may be remedied by integrating MIMICS with a full ecosystem biogeochemical model that represents the greater complexity of the plant-soil continuum.

MIMICS-CN provides a pathway to reconcile mechanistic explanations for phenomena like priming and plant-soil feedbacks with emergent patterns in terrestrial biogeochemistry across landscapes. MIMICS-CN and microbial models like it are a good first step towards representing the complex ecological factors that drive the coupling of soil C and N biogeochemistry, including the distribution of SOM among functionally relevant pools and SOM C:N ratios. Future work could compare model formulations that take different approaches to microbial community and stoichiometric parameters (e.g. flexible microbial parameters like C:N or CUE, additional microbial groups, partitioning microbial metabolism into a greater number of pathways) and refinement of mechanisms that confer SOM persistence. These efforts should also assess the ramifications of different choices for simulating existing data and predicting the long-term response of soil C and N cycles to global change. Our work demonstrates that MIMICS-CN can reproduce site and litter quality effects on litter decomposition C and N dynamics at a landscape scale, while also pointing to the importance of underlying, interacting microbial and biogeochemical factors in regulating SOM dynamics. Future work coupling MIMICS-CN to experiments and syntheses relating the distribution of SOM across pools to their underlying controls across gradients will improve our confidence in our ability to understand and project SOM dynamics.

Code and data availability

MIMICS-CN (v1.0) is written in R using packages rootSolve (Soetaert and Herman, 2009) and hydroGOF (Zambrano-Bigiarini, 2017). Figures were generated using packages ggplot2 (Wickham, 2016), reshape2 (Wickham, 2007), scales (Wickham, 2018), gridextra (Auguie, 2017), and cowplot (Wilke, 2016). The R scripts and datasets used to generate model results are available at https://zenodo.org/record/3534562. See Appendix A for equations.

Appendix A: Model equations

The structure and assumptions in the C-only version of MIMICS have been described previously (Wieder et al., 2014, 2015d), and the structure and assumptions in MIMIC-CN are described in section 2.1 ("Model formulation") of the methods section of this paper. The C fluxes (mg C cm^{-3} h^{-1}) from donor to receiver pools in MIMICS-CN, numbered on Fig. 1, are defined by the following:

$$\mathrm{LIT}_{m,C}\_\mathrm{MIC}_{r,C} = \mathrm{MIC}_{r,C} \times V_{\max[r1]} \times \mathrm{LIT}_{m,C} / (K_{m[r1]} + \mathrm{LIT}_{m,C}), \tag{A1}$$

$$\mathrm{LIT}_{s,C}\_\mathrm{MIC}_{r,C} = \mathrm{MIC}_{r,C} \times V_{\max[r2]} \times \mathrm{LIT}_{s,C} / (K_{m[r2]} + \mathrm{LIT}_{s,C}), \tag{A2}$$

$$\mathrm{SOM}_{a,C}\_\mathrm{MIC}_{r,C} = \mathrm{MIC}_{r,C} \times V_{\max[r3]} \times \mathrm{SOM}_{a,C} / (K_{m[r3]} + \mathrm{SOM}_{a,C}), \tag{A3}$$

$$\mathrm{MIC}_{r,C}\_\mathrm{SOM}_{C} = \mathrm{MIC}_{r,C}{}^{\beta} \times \tau_{[r],} \tag{A4}$$

$$\mathrm{LIT}_{m,C}\_\mathrm{MIC}_{K,C} = \mathrm{MIC}_{K,C} \times V_{\max[K1]} \times \mathrm{LIT}_{m,C} / (K_{m[K1]} + \mathrm{LIT}_{m,C}), \tag{A5}$$

$$\mathrm{LIT}_{s,C}\_\mathrm{MIC}_{K,C} = \mathrm{MIC}_{K,C} \times V_{\max[K2]} \times \mathrm{LIT}_{s,C} / (K_{m[K2]} + \mathrm{LIT}_{s,C}), \tag{A6}$$

$$\mathrm{SOM}_{a,C}\_\mathrm{MIC}_{K,C} = \mathrm{MIC}_{K,C} \times V_{\max[K3]} \times \mathrm{SOM}_{a,C} / (K_{m[K3]} + \mathrm{SOM}_{a,C}), \tag{A7}$$

$$\mathrm{MIC}_{K,C}\_\mathrm{SOM}_{C} = \mathrm{MIC}_{K,C}{}^{\beta} \times \tau_{[K]}, \tag{A8}$$

$$\mathrm{SOM}_{p,C}\_\mathrm{SOM}_{a,C} = \mathrm{SOM}_{p,C} \times D, \tag{A9}$$

$$\begin{aligned}\mathrm{SOM}_{c,C}\_\mathrm{SOM}_{a,C} = {} & (\mathrm{MIC}_{r,C} \times V_{\max[r2]} \times \mathrm{SOM}_{c,C} / (KO_{[r]} \times K_{m[r2]} + \mathrm{SOM}_{c,C})) \; + \\ & (\mathrm{MIC}_{K,C} \times V_{\max[K2]} \times \mathrm{SOM}_{c,C} / (KO_{[K]} \times K_{m[K2]} + \mathrm{SOM}_{c,C})).\end{aligned} \tag{A10}$$

where pools and parameters are described in section 2.1 and Table 1, respectively. The N fluxes (mg N cm^{-3} h^{-1}) from donor to receiver pools in MIMICS-CN are calculated based on the C fluxes between pools and the C:N ratio of donor pools. These fluxes are numbered on Fig. 1 and defined by the following:

$LIT_{m,N}\_MIC_{r,N} = A1 \times LIT_{m,N} / LIT_{m,C}$, (A11)

$LIT_{s,N}\_MIC_{r,N} = A2 \times LIT_{s,N} / LIT_{s,C}$, (A12)

$SOM_{a,N}\_MIC_{r,N} = A3 \times SOM_{a,N} / SOM_{a,C}$, (A13)

$MIC_{r,N}\_SOM_{,N} = A4 \times MIC_{r,N} / MIC_{r,C}$, (A14)

$LIT_{m,N}\_MIC_{K,N} = A5 \times LIT_{m,N} / LIT_{m,C}$, (A15)

$LIT_{s,N}\_MIC_{K,N} = A6 \times LIT_{s,N} / LIT_{s,C}$, (A16)

$SOM_{a,N}\_MIC_{K,N} = A7 \times SOM_{a,N} / SOM_{a,C}$, (A17)

$MIC_{K,N}\_SOM_{,N} = A8 \times MIC_{K,N} / MIC_{K,C}$, (A18)

$SOM_{p,N}\_SOM_{a,N} = A9 \times SOM_{p,N} / SOM_{p,C}$, (A19)

$SOM_{c,N}\_SOM_{a,N} = A10 \times SOM_{c,N} / SOM_{c,C}$. (A20)

Each time step, the microbial pools in MIMICS-CN take up inorganic N from the DIN pool proportional to the biomass in each pool. Subsequently, the C:N ratio of all the inputs to each microbial pool is calculated, and the microbial pools spill either excess C or excess N to maintain a model-defined C:N ratio of microbial biomass. The algorithm that determines the release of excess C or N is determined using the following equations:

$DINup_r = (1 - N_{leak}) \times DIN \times MIC_{r,C} / (MIC_{r,C} + MIC_{K,C})$, (A21)

$DINup_K = (1 - N_{leak}) \times DIN \times MIC_{K,C} / (MIC_{r,C} + MIC_{K,C})$, (A22)

$upMIC_{r,C} = CUE_{[1]} \times (A1 + A3) + CUE_{[2]} \times (A2)$, (A23)

$upMIC_{r,N} = NUE \times (A11 + A13 + A12) + A21$, (A24)

$CNup_r = A23 / A24$, (A25)

$$\mathrm{Overflow_r = A23 - (A24 \times min(CN_r, A25)),} \tag{A26}$$

$$\mathrm{Nspill_r = A24 - (A23 / max(CN_r, A25)),} \tag{A27}$$

$$\mathrm{upMIC_{K,C} = CUE_{[3]} \times (A5 + A7) + CUE_{[4]} \times (A6),} \tag{A28}$$

$$\mathrm{upMIC_{K,N} = NUE \times (A15 + A17 + A16) + A22,} \tag{A29}$$

$$\mathrm{CNup_K = A28 / A29,} \tag{A30}$$

$$\mathrm{Overflow_K = A28 - (A29 \times min(CN_K, A30)),} \tag{A31}$$

$$\mathrm{Nspill_K = A29 - (A28 / max(CN_K, A30)).} \tag{A32}$$

Inorganic N leaches slowly from the model according to a model-defined rate:

$$\mathrm{LeachingLoss} = N_{leak} \times \mathrm{DIN.} \tag{A33}$$

Given the fluxes defined above, the changes in C and N pools in each hourly timestep (mg C or N cm^{-3}) are described by the following:

$$\frac{\mathrm{dLIT_{m,C}}}{\mathrm{dt}} = \mathrm{I_{LIT_{m,C}}} \times \left(1\text{-}\mathrm{f_{i,met}}\right) - \mathrm{A1 - A5,} \tag{A34}$$

$$\frac{\mathrm{dLIT_{s,C}}}{\mathrm{dt}} = \mathrm{I_{LIT_{s,C}}} \times \left(1\text{-}\mathrm{f_{i,struc}}\right) - \mathrm{A2 - A6,} \tag{A35}$$

$$\frac{\mathrm{dMIC_{r,C}}}{\mathrm{dt}} = \mathrm{CUE_{[1]} \times (A1 + A3) + CUE_{[2]} \times (A2) - A4 - Overflow_r,} \tag{A36}$$

$$\frac{\mathrm{dMIC_{K,C}}}{\mathrm{dt}} = \mathrm{CUE_{[3]} \times (A5 + A7) + CUE_{[4]} \times (A6) - A8 - Overflow_K,} \tag{A37}$$

$$\frac{\mathrm{dSOM_{p,C}}}{\mathrm{dt}} = \mathrm{I_{LIT_{m,C}} \times f_{i,met} + (f_{p,r} \times A4) + (f_{p,K} \times A8) - A9,} \tag{A38}$$

$$\frac{\mathrm{dSOM_{c,C}}}{\mathrm{dt}} = \mathrm{I_{LIT_{s,C}} \times f_{i,struc} + (f_{c,r} \times A4) + (f_{c,K} \times A8) - A10,} \tag{A39}$$

$$\frac{\mathrm{dSOM_{a,C}}}{\mathrm{dt}} = \mathrm{(f_{a,r} \times A4) + (f_{a,K} \times A8) + A9 + A10 - A3 - A7,} \tag{A40}$$

$$\frac{dLIT_{m,N}}{dt} = \frac{I_{LIT_{m,C}} \times \left(1\text{-}f_{i,met}\right)}{CN_m} - A11 - A15, \tag{A41}$$

$$\frac{dLIT_{s,N}}{dt} = \frac{I_{LIT_{s,C}} \times \left(1\text{-}f_{i,struc}\right)}{CN_s} - A12 - A16, \tag{A42}$$

$$\frac{dMIC_{r,N}}{dt} = NUE \times (A11 + A13 + A12) - A14 + DINup_r - Nspill_r, \tag{A43}$$

$$\frac{dMIC_{K,N}}{dt} = NUE \times (A15 + A17 + A16) - A18 + DINup_K - Nspill_K, \tag{A44}$$

$$\frac{dSOM_{p,N}}{dt} = \frac{I_{LIT_{m,C}} \times \left(f_{i,met}\right)}{CN_m} + (f_{p,r} \times A14) + (f_{p,K} \times A18) - A19, \tag{A45}$$

$$\frac{dSOM_{c,N}}{dt} = \frac{I_{LIT_{s,C}} \times \left(f_{i,struc}\right)}{CN_s} + (f_{c,r} \times A14) + (f_{c,K} \times A18) - A20, \tag{A46}$$

$$\frac{dSOM_{a,N}}{dt} = (f_{a,r} \times A14) + (f_{a,K} \times A18) + A19 + A20 - A13 - A17, \tag{A47}$$

$$\frac{dDIN}{dt} = (1 - NUE) \times (A11 + A12 + A13 + A15 + A16 + A17) + $$

$$Nspill_r + Nspill_K - DINup_r - DINup_K - LeachingLoss. \tag{A48}$$

Table 1. Parameters used in MIMICS-CN for both LIDET and equilibrium simulations.

Parameter	Description	Value	Units
f_{met}	Partitioning of inputs to metabolic litter pool	0.85 - 0.013 (lignin/N)	-
f_{i}	Fraction of litter inputs transferred to SOM	0.05, 0.3	-
V_{slope} (Met-r, Met-K, Struc-r)	Regression coefficient	0.063	$\ln(\mathrm{mg\ C\ (mg\ MIC)^{-1}\ h^{-1}})°\mathrm{C}^{-1}$
V_{slope} (Struc-K, Avail-r, Avail-K)	Regression coefficient	0.043	$\ln(\mathrm{mg\ C\ (mg\ MIC)^{-1}\ h^{-1}})°\mathrm{C}^{-1}$
V_{int}	Regression intercept	5.47	$\ln(\mathrm{mg\ C\ (mg\ MIC)^{-1}\ h^{-1}})$
a_{V}	Tuning coefficient	4.8×10^{-7}	-
V_{mod}	Modifies V_{max}	10, 1.5, 10, 3, 2.25, 2	-
V_{max}	Temperature-sensitive maximum reaction velocity (T is mean annual soil temperature)	$e^{(V\mathrm{slope} \times T + V\mathrm{int})} \times a_{\mathrm{v}} \times V_{\mathrm{mod}}$	$\mathrm{mg\ C\ (mg\ MIC)^{-1}}$
K_{slope} (Met-r, Met-K, Avail-r, Avail-K)	Regression coefficient	0.017	$\ln(\mathrm{mg\ C\ cm^{-3}})°\mathrm{C}^{-1}$
K_{slope} (Struc-r, Struc-K)	Regression coefficient	0.027	$\ln(\mathrm{mg\ C\ cm^{-3}})°\mathrm{C}^{-1}$
K_{int}	Regression intercept	3.19	$\ln(\mathrm{mg\ C\ cm^{-3}})$
a_{K}	Tuning coefficient	0.5	-
P_{scalar}	Physical protection scalar used in K_{mod}	$(2 \times e^{-2 \times \sqrt{(f\mathrm{clay})}})^{-1}$	-
K_{mod}	Modifies K_{m}	$0.125, 0.5, 0.25 \times P_{\mathrm{scalar}}$, $0.5, 0.25, 0.167 \times P_{\mathrm{scalar}}$	-
KO	Further modifies K_{m} for oxidation of $\mathrm{SOM_c}$	6, 6	-
K_{m}	Half saturation constant (T is mean annual soil temperature)	$e^{(K\mathrm{slope} \times T + V\mathrm{int})} \times a_{\mathrm{k}} \times K_{\mathrm{mod}}$	$\mathrm{mg\ C\ cm^{-3}}$
τ	Microbial biomass turnover rate	$2.4 \times 10^{-4} \times e^{0.3\,(f\mathrm{met})} \times \tau_{\mathrm{mod1}} \times \tau_{\mathrm{mod2}}$, $1.1 \times 10^{-4} \times e^{0.1\,(f\mathrm{met})} \times \tau_{\mathrm{mod1}} \times \tau_{\mathrm{mod2}}$	$\mathrm{h^{-1}}$
τ_{mod1}	Modifies microbial turnover rate	$0.6 < \sqrt{(\mathrm{NPP}/100)} < 1.3$	-
τ_{mod2}	Modifies microbial turnover rate	$\tau \times 0.55 / (.45 \times \mathrm{Inputs})$	-
β	Exponent that modifies turnover rate	2	-
CUE	Microbial carbon use efficiency	0.55, 0.25, 0.75, 0.35	$\mathrm{mg\ mg^{-1}}$
NUE	Proportion of mineralized N captured by microbes	0.85	$\mathrm{mg\ mg^{-1}}$
$\mathrm{CN_s}$	C:N of structural litter	$(\mathrm{Measured\ CN} - \mathrm{CN_m} \times f_{\mathrm{met}}) / (1-f_{\mathrm{met}})$	$\mathrm{mg\ mg^{-1}}$
$\mathrm{CN_m}$	C:N of metabolic litter	15	$\mathrm{mg\ mg^{-1}}$
$\mathrm{CN_r}$	C:N of copiotrophic microbial pool	6	$\mathrm{mg\ mg^{-1}}$
$\mathrm{CN_k}$	C:N of oligotrophic microbial pool	10	$\mathrm{mg\ mg^{-1}}$
f_{p}	Fraction of τ partitioned to $\mathrm{SOM_p}$	$0.015 \times e^{1.3\,(f\mathrm{clay})}$, $0.01 \times e^{0.8\,(f\mathrm{clay})}$	-
f_{c}	Fraction of τ partitioned to $\mathrm{SOM_c}$	$0.3 \times e^{-3\,(f\mathrm{met})}$, $0.9 \times e^{-3\,(f\mathrm{met})}$	-
f_{a}	Fraction of τ partitioned to $\mathrm{SOM_a}$	$1 - (f_{\mathrm{p}} + f_{\mathrm{c}})$	-
D	Desorption rate from $\mathrm{SOM_p}$ to $\mathrm{SOM_a}$	$10^{-6} \times e^{-4.5\,(f\mathrm{clay})}$	$\mathrm{h^{-1}}$
N_{leak}	Rate of loss of inorganic N pool	0.2	$\mathrm{h^{-1}}$

Table 2. Goodness-of-fit statistics comparing MIMICS-CN and DAYCENT simulations to observations of C and N in decomposing litterbags in the LIDET study, aggregated by biome. DAYCENT results are subset from simulations in Bonan et al. (2013) to match the sites included in MIMICS-CN simulations. The values shown are the number of observations (n), Pearson's correlation coefficient squared (R^2), root mean square error (RMSE), and bias calculated between observed and simulated percent C and N remaining. For more details on the sites grouped into each biome, see Wieder et al. (2015).

		MIMICS-CN Carbon			DAYCENT Carbon			MIMICS-CN Nitrogen			DAYCENT Nitrogen		
Biome	n	R^2	RMSE	bias	R^2	RMSE	bias	R^2	RMSE	bias	R^2	RMSE	bias
Tundra	114	0.74	12.56	9.49	0.78	8.32	3.21	0.33	0.32	0.09	0.41	0.31	0.00
Boreal	60	0.61	14.30	9.32	0.73	9.06	-0.55	0.64	0.28	0.07	0.72	0.27	-0.14
Conifer	60	0.79	18.61	-16.42	0.89	9.09	5.93	0.73	0.20	0.05	0.79	0.26	0.13
Deciduous	94	0.59	16.40	-8.92	0.80	12.36	9.20	0.51	0.31	-0.13	0.63	0.33	0.18
Humid	151	0.50	17.24	-3.23	0.61	15.18	-4.22	0.14	0.44	-0.13	0.24	0.45	-0.04
Arid	113	0.61	16.67	2.09	0.68	19.90	11.63	0.32	0.29	0.16	0.01	0.49	0.20
Tropical	46	0.57	15.29	7.75	0.64	20.81	17.04	0.46	0.45	0.36	0.20	0.55	0.35
All	638	0.63	16.00	-0.12	0.67	14.36	4.73	0.29	0.34	0.03	0.30	0.40	0.08

Table 3. Ranges of MIMICS-CN estimates of steady-state values for a variety of soil pools and fluxes, compared against observed ranges from several continent-wide data synthesis studies. The ranges of values included for MIMICS-CN are derived from simulations of sites included in the LIDET study.

	MIMICS-CN range	Published range	Reference
Total C (mg cm^{-3})*	7.0-50	3.9-89	Zak et al. 1994
		2.7-360	Xu, Thornton and Post 2013
		5.2-610	Cleveland and Liptzin 2007
Total N (mg cm^{-3})*	0.60-5.1	0.38-5.1	Zak et al. 1994
		0.66-22	Xu, Thornton and Post 2013
		0.39-24	Cleveland and Liptzin 2007
Soil C:N	9.6-12	4.0-40	Colman and Schimel 2013
		10-28	Zak et al. 1994
		11-31	Xu, Thornton and Post 2013
		2.0-82	Cleveland and Liptzin 2007
Inorganic nitrogen (μg cm^{-3})	0.01-0.06	0.12-8.1	Zak et al. 1994
Respiration (μg C cm^{-3} hr^{-1})	0.02-0.28	0.01-0.70	Colman and Schimel 2013
		0.21-0.91	Zak et al. 1994
Net N mineralization (μg N cm^{-3} hr^{-1})	0-0.01	0-0.10	Colman and Schimel 2013
		0.004-0.058	Zak et al. 1994
Microbial biomass C (mg cm^{-3})	0.15-1.3	0.03-1.3	Zak et al. 1994
		0.01-5.3	Xu, Thornton and Post 2013
		0.08-39	Cleveland and Liptzin 2007
Microbial biomass N (mg cm^{-3})	0.02-0.16	0.006-0.33	Zak et al. 1994
		0.042-0.64	Xu, Thornton and Post 2013
		0.018-4.9	Cleveland and Liptzin 2007
Microbial biomass C as % of soil C	0.95-4.8	0.18-3.3	Zak et al. 1994
		0.99-5.0	Xu, Thornton and Post 2013
		0.27-93	Cleveland and Liptzin 2007
Microbial biomass N as % of soil N	1.2-5.9	1.1-15	Zak et al. 1994
		2.3-5.7	Xu, Thornton and Post 2013
		0.48-64	Cleveland and Liptzin 2007

*Depths simulated by MIMICS-CN are for the top 30 cm of soil, whereas published ranges represent measurements ranging from the top 5 to top 30 cm.

Figure 1. Overview of the pools and fluxes of C and N in MIMICS-CN. Black outlines indicate pools that contain C; green outlines indicate pools that contain N. Litter inputs (I) are determined based on site-specific net primary productivity and partitioned between metabolic and structural litter pools (LITm and LITs) using a site-specific litter quality metric (fmet) calculated using litter lignin and N content. Temperature-sensitive forward Michaelis-Menten kinetics (Vmax and Km, red lines) determine the flux of litter pool C and N and available SOM C and N (SOMa) into microbial biomass (MICr and MICK). Fluxes of C into microbial pools result in respiration losses according to a defined carbon use efficiency (CUE). Microbes maintain biomass stoichiometry by spilling excess C as overflow respiration or excess N into the dissolved inorganic nitrogen pool (DIN) based on a prescribed biomass C:N. Microbial biomass turnover (τ, blue) varies by functional type (MICr and MICK) and is proportional to the square of microbial biomass. Microbial biomass turns over into available (SOMa), physicochemically-stabilized (SOMp) and chemically-stabilized (SOMc) soil organic matter pools. Inorganic N (DIN) leaks from the model at a first-order rate. Numbers in parentheses indicate the equations in Appendix A that correspond to each depicted flux. Parameter values, units and descriptions are given in Table 1.

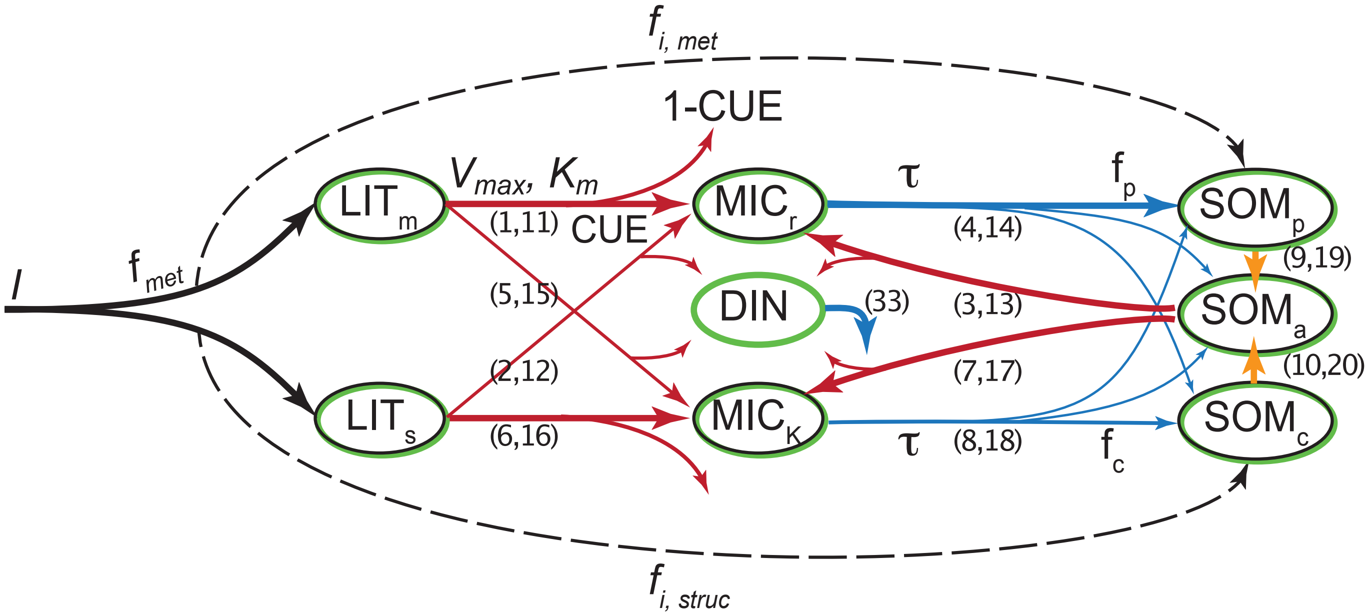


Figure 2. Litter decomposition timeseries simulated by MIMICS-CN (lines with shaded area) compared to observations (points and error bars) of (a) percent mass remaining and (b) percent of initial N remaining over ten years for six different litter types at the Harvard Forest LTER. Litter decomposition data came from the LIDET study (Parton et al., 2007; Bonan et al., 2013; mean ±1 SD). Spread in the observations and model are largely generated by the effects of initial litter quality on decomposition rates and N dynamics. Model parameters were calibrated to fit MIMICS-CN to observations from Harvard Forest (Table 1).

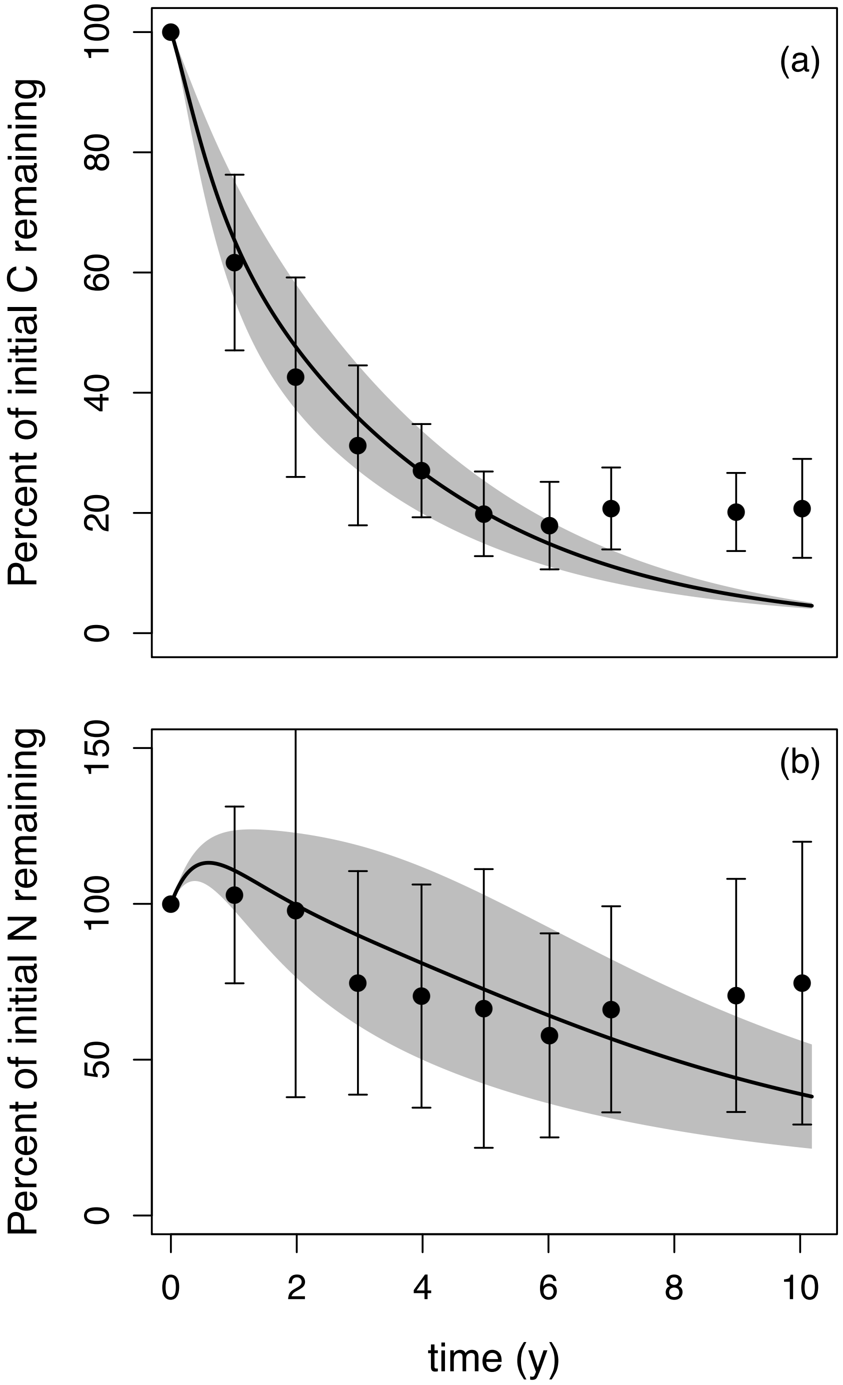


Figure 3. MIMICS-CN simulations of percent C remaining (top) and N remaining (bottom) in litterbags in the LIDET study versus observed values, colored by litter type (left) or biome (right). Dashed line shows the 1:1 line.

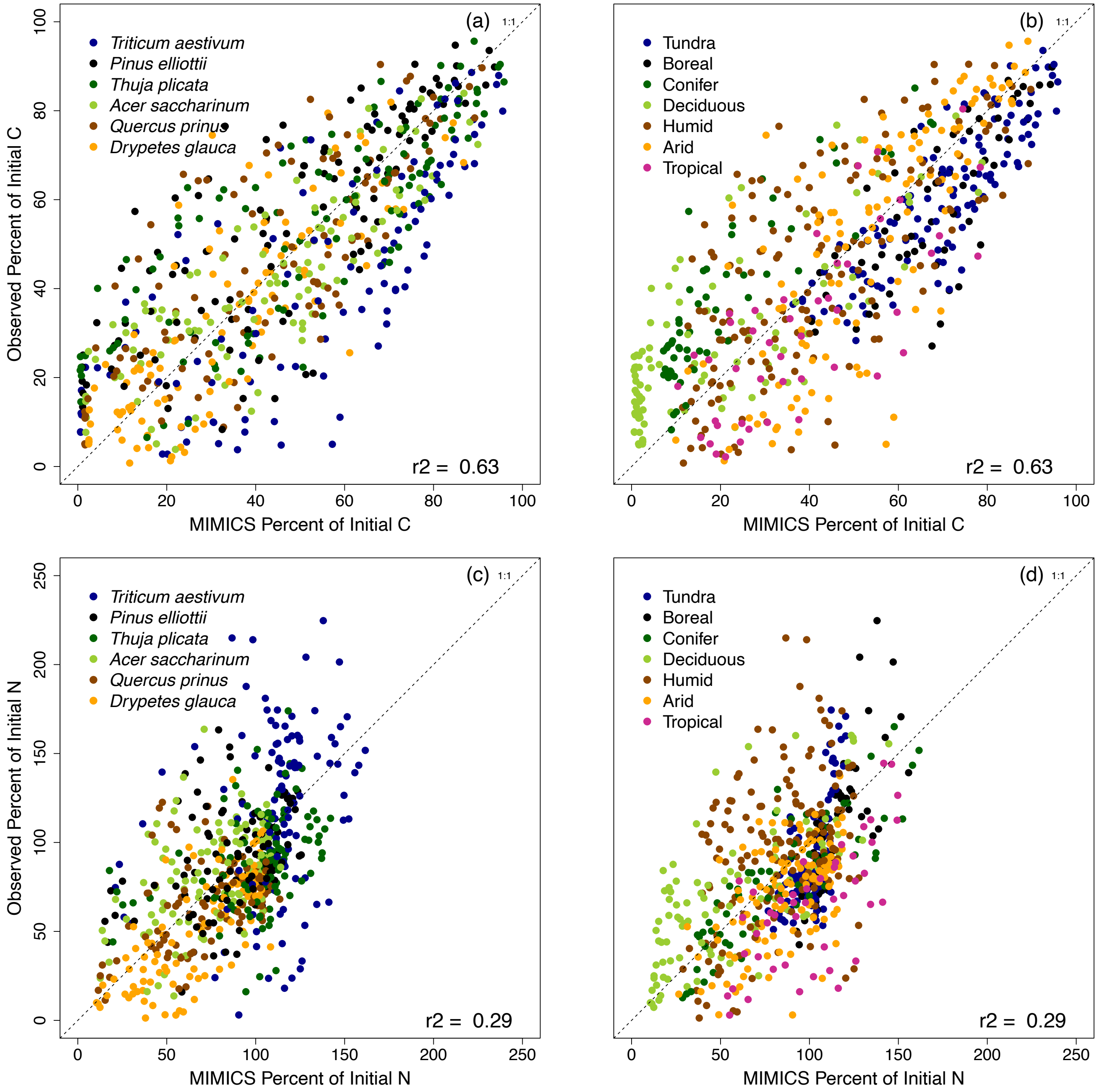


Figure 4. MIMICS-CN simulations of immobilization-mineralization thresholds across litters of different quality. Litter quality (in terms of C:N and lignin content) decreases from upper left panel to lower right panel. Red triangles show model simulations of C losses vs N losses from litterbags in the LIDET study. Colored dots show observed C vs N losses across biomes (Parton et al. 2007).

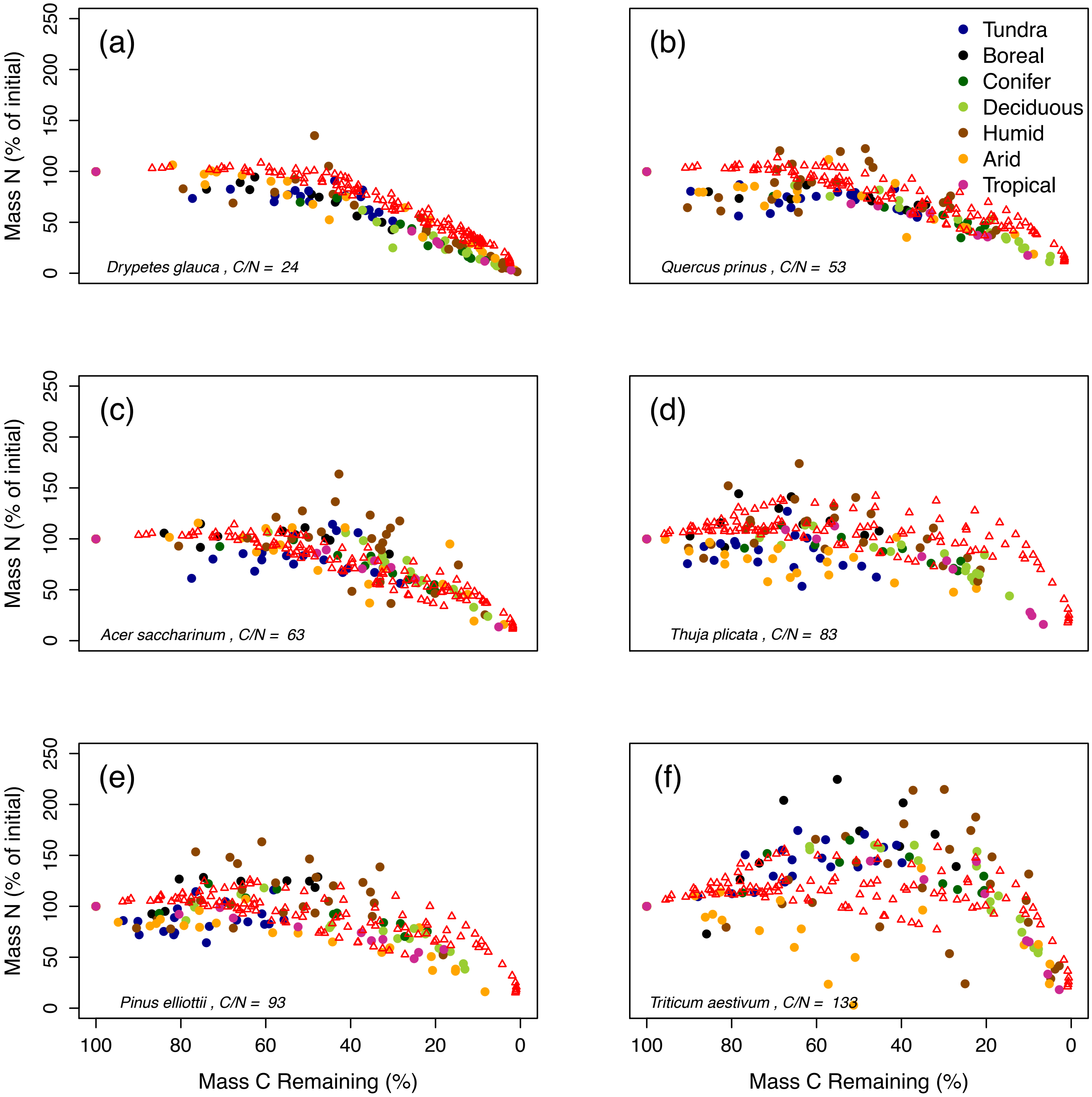


Figure 5. Distributions of MIMICS-CN estimates of steady-state values for a variety of soil pools and fluxes, compared against observed ranges from several continent-wide data synthesis studies. Black lines show the median value across all observations; red lines show median value of MIMICS-CN simulations.

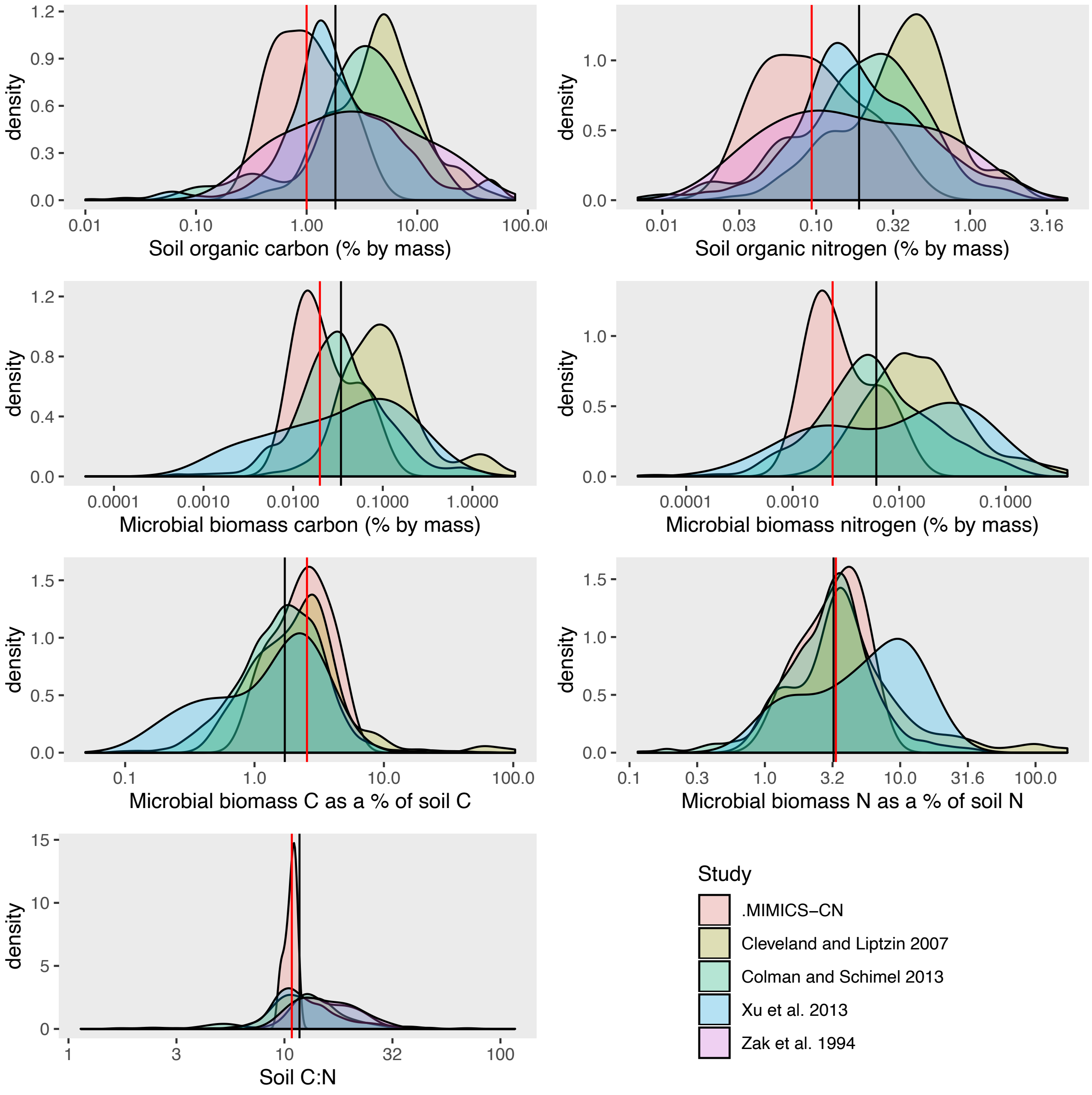
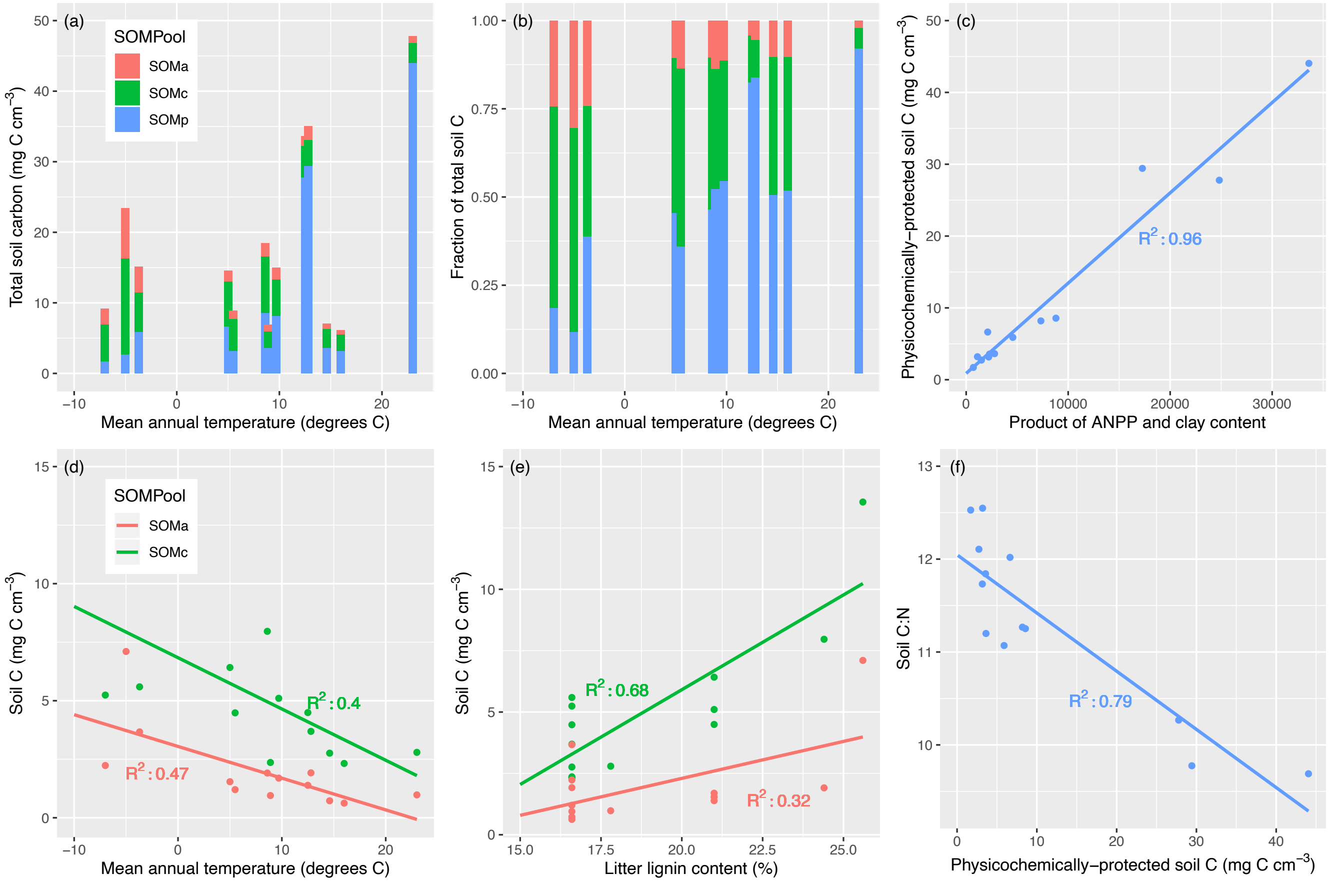


Figure 6. Variation in steady state SOM pools and environmental factors controlling their distribution in MIMICS-CN simulations across LIDET sites. Top panels show the (a) total C stocks in physicochemically-protected, chemically-protected, and available SOM pools (SOMp, SOMc, SOMa pools, respectively) arranged by the site mean annual temperature (MAT), or the (b) relative fraction of each SOM pool arranged in the same way. Upper right and bottom panels show the correlations between C in each SOM pool and environmental drivers including: (c) SOMp vs. the product of annual net primary productivity (ANPP) and clay content, (d) SOMc and SOMa vs. MAT, and (e) SOMc and SOMa vs. lignin content of litter inputs at each site. Finally, (f) soil stoichiometry is largely determined by the fraction of total SOM pools that are considered physicochemically protected.



CHAPTER III: SOIL CARBON AND NITROGEN COUPLINGS IN OBSERVATIONS AND MODELS

In preparation for review at Biogeochemistry

Abstract

Soils couple carbon (C) and nitrogen (N) cycles through ecological stoichiometry, and the ratio of soil C and N (i.e. soil C:N ratio) serves as an easily-measured, widely available proxy for more advanced measures of soil chemistry that can be used to evaluate soil models. Previous syntheses have explored the aboveground controls on soil stoichiometry (e.g. plant cover, climate), but relatively less focus has been placed on the belowground controls on soil stoichiometry (e.g., edaphic factors, mineralogy) and the relative importance of these different drivers. Here, we strive to understand the relative importance of drivers of soil stoichiometry across a global-scale synthesis of measurements and soil models that couple C and N cycling. We evaluated patterns in the Soil Data Harmonization (SoDaH) dataset against patterns generated by two soil models using multiple linear regressions. We used one model that explicitly represents microbial biomass, coupled C and N cycles, and multiple pools of stabilized organic matter (MIMICS-CN) and another model that uses a cascade of linearly decomposing pools (CASA-CNP) to simulate a global range of sites and assess relationships between soil C:N and edaphic and climate variables. Both the SoDaH database and our model simulations indicated strong negative relationships between soil clay content and bulk soil C:N. However, mean annual temperature and plant litter C:N had stronger impacts on soil C:N in one of our model simulations (MIMICS-CN) than in the database. In addition, pH was an important driver of soil C:N relationships in the database that was not represented in the models. Our results highlight the importance of more widespread measurements of plant litter chemistry in conjunction with soil variables. In addition, our results point to the importance of bottom-up controls (i.e. mineralogy) in determining soil C:N in observational data, and we discuss future

work for specifying and quantifying these bottom-up controls and altering models to match the observed importance of such controls.

1. Introduction

Soils are a massive global store of carbon (C) and nitrogen (N) (Batjes, 1996; Jackson et al., 2017) with an unclear but potentially high vulnerability to global change that could exacerbate climate change through increasing greenhouse gas fluxes (Bowles et al., 2018; Conant et al., 2008; Davidson and Janssens, 2006; Dungait et al., 2012; Wagai et al., 2013). Projecting future changes in C and N cycling is key to forecasting global change. However, C and N cycles cannot be fully understood or accurately modeled in isolation because of their many interactions and feedbacks (Bonan and Levis, 2010; Sokolov et al., 2008; Zaehle and Dalmonech, 2011). For example, N availability in soil can limit the ability of plants to respond to elevated CO_2 with increased productivity (Norby et al., 2010). Concurrently, C availability can limit the ability of decomposer microbes to generate enzymes to release N from soil organic matter (SOM) (Soong et al., 2020). Representing these dynamics in numerical models is challenging, but there is a strong need for models that can accurately simulate C and N fluxes and states in diverse soils across biomes as they respond to global change. Models, however, can only advance as quickly as scientific understanding does, and fundamental theoretical questions still exist about the nature, origin, and vulnerability of soil organic matter.

For a long time, soil scientists assumed the longevity of SOM was due to its chemical recalcitrance and that the oldest SOM was made up of the recalcitrant leftovers from plant residue decomposition (Schmidt et al., 2011; Weil and Brady, 2016). Recent developments in soil science highlight the role of microbial necromass in forming SOM (Bradford et al., 2013; Cotrufo et al., 2013; Kallenbach et al., 2015; Liang et al., 2019; Ludwig et al., 2015), and the most modern theories posit that SOM is made up of a combination of substances derived from both plant and microbial origins (Lavallee et al., 2020; Lehmann and Kleber, 2015). New research is also beginning to unpack the role of edaphic factors in driving the distribution of SOM across primarily plant-derived or microbial-derived forms (Haddix et al., 2020) and the susceptibility of different forms to global change (Rocci et al., 2021). For example, soil minerals differ in their capacity to stabilize SOM and preferentially bond to different organic moieties (Jilling et al., 2018). Moreover, environmental factors like pH, N availability and precipitation can drive the composition and activity of microbial communities that break down plant matter and SOM (Evans and Wallenstein, 2012; Fierer and Jackson, 2006; Moore et al., 2021), ultimately driving rates of decomposition. Unpacking this web of factors remains a sticky problem in part because soil measurements struggle to distinguish between SOM stabilized by chemical recalcitrance, mineral association, or other means (Lavallee et al., 2020; Poeplau et al., 2018).

Nutrient stoichiometry provides a window into the chemical makeup of SOM that can help resolve the challenge of unravelling the controls on plant- or microbially-derived SOM. The ratio of C and N in particular serves as a useful, easily-measured proxy for more complex analyses of SOM chemistry (Bailey et al., 2018). The C:N ratio of plant inputs to soils is generally much higher (i.e. less enriched in N) than decomposer organisms (Sterner and Elser 2000), and the C:N ratio of SOM emerges in part from the biochemistry of soil organisms as they process plant matter. Both microbial and soil C:N are tightly conserved across sites compared to other indicators of SOM chemistry. Previous large-scale syntheses have found well-constrained stoichiometric ratios for soil and soil microbes (Xu et al. (2013): 16.9 and 7 respectively; Cleveland and Liptzin (2007): 14.3 and 8.6 respectively) with variation across climate and vegetation gradients. Given the relative ease with which soil C and N can be measured, soil C:N values are available across a large number of sites and ecosystems, making C:N a particularly valuable quantity for testing model assumptions about the balance of factors in driving soil processes. Proxy variables like this are an important and necessary tool for assessing broad-scale patterns in soil chemistry (Bailey et al., 2018) when more specific measurements of soil chemical classes are not available at the same scale.

Previous syntheses have gathered stoichiometric data from across a global range of sites and explored the effects of climate and vegetation on observed soil C:N distributions (Cleveland and Liptzin, 2007; Homann et al., 2007; Tian et al., 2010; Tipping et al., 2016; Xu et al., 2013), but less emphasis has been placed on exploring the belowground, bottom-up controls on soil stoichiometry. While aboveground drivers matter, soil C:N can also be influenced by a variety of other factors (Fig. 1). In general, microbial residues are enriched in N relative to plant residues (Melillo et al., 1989). Once embedded in the soil matrix, many factors can influence the stability and relative proportion of plant or microbial residues across soil profiles and landscapes. For example, surface charges on clays form stable associations with charged moieties like amino groups (Jilling et al., 2018), and clay fractions in soils tend to be enriched in N relative to sand fractions (Haddix et al., 2016). Minerals with variable surface properties can modify the sorptive fractionation of N-rich molecules in clay fractions (Kramer et al., 2017; Mikutta et al., 2019). The inherent C:N stoichiometry of plants and microbes themselves also influences soil C:N, since soil C:N emerges in part from microbial decomposition and transformation of plant litter inputs. Additionally, climate (temperature, precipitation) and edaphic (pH, mineralogy) factors can further influence the composition and activity of soil microbes (Delgado-Baquerizo et al., 2018; Fierer and Jackson, 2006), with downstream impacts on soil C:N.

A large number of factors contributes to the emergence of soil C:N values, but for the sake of understanding broad trends and evaluating models, it is useful to simplify this diversity of drivers down to broad categories that correspond to widely-available measured variables. We base our theoretical framework of soil C:N ratio on understanding that soil contains a mix of organic matter along the continuum from unprocessed plant matter to stabilized microbial residues (Lehmann and Kleber, 2015), whereby three categories of factors might influence the C:N ratio of SOM (Fig. 1). First, the stoichiometry of plant inputs to soil should clearly play a role as the source of the vast majority of C and N in soil. Second, climate factors and microbial traits (e.g. growth rate, growth efficiency, nutrient efficiency, biomass C:N) should drive the rate of transformation of plant litter into nitrogen-enriched microbial biomass and therefore influence soil C:N. Finally, edaphic factors like clay content should drive the long-term stabilization of microbial residues and influence the proportion of organic matter made up of those N-enriched residues. Within these three broad categories that we might expect to drive soil stoichiometry, we can select specific variables that soil models take as inputs, namely plant litter C:N, mean annual temperature, and clay content. Although not collected in previous syntheses (Cleveland and Liptzin, 2007; Xu et al., 2013), these variables are a good starting place to compare patterns in models and data because models take these variables as inputs and a large number of measured records are available for these three variables from other sources.

Soil models that include both C and N cycling vary in their assumption about the controls on various forms of stabilized SOM. These assumptions can be mapped onto the framework presented above. For this work, we chose to compare two global-scale models of soil that couple C and N cycles and vary in their simulated controls on SOM cycling: the Carnegie-Ames-Stanford Approach model of Carbon, Nitrogen and Phosphorus (CASA-CNP) model and the MIcrobial-MIneral Carbon Stabilization model with coupled nitrogen cycling (MIMICS-CN). CASA-CNP represents soil as a cascade of linearly decomposing pools in which plant litter inputs are transferred sequentially into SOM pools of increasing stability and N enrichment (Randerson et al., 1996; Wang et al., 2010). On the other hand, MIMICS-CN focuses on microbial controls over the transformation and stabilization of SOM (Kyker-Snowman et al., 2020; Wieder et al., 2014, 2015d). Thus, CASA-CNP emphasizes climate and the stoichiometry of plant litter inputs in determining the C:N ratio of SOM, whereas MIMICS-CN potentially puts greater emphasis on physicochemical stabilization of microbial necromass. The differences in the theoretical assumptions and structure of these two models should generate soil C:N patterns that emphasize different drivers of SOM stoichiometry. For example, the focus on microbial necromass in forming SOM in MIMICS-CN should result in an emphasis on bottom-up controls on soil C:N: (e.g. clay) relative to CASA-CNP, which should emphasize top-down controls (e.g. plant litter chemistry). These model-generated hypotheses are testable against data but doing so requires careful alignment of measured and modeled variables.

Novel data products provide opportunities to test the processes that drive emergent patterns in soil stoichiometry as well as their application in models. In this paper, we conduct a series of data-model comparisons to explore real-world drivers and test model assumptions of soil stoichiometric patterns. Specifically, our questions are:

1) How important are plant litter chemistry, climate, and soil abiotic controls over soil C:N stoichiometry across real soils and model simulations?

2) Are there important observed drivers of soil C:N that models miss and that modelers should work to include in the near term?

We investigated these questions with a combination of model simulations and large-scale data exploration. First, we used a recent synthesis of soil data from the Soils Data Harmonization (SoDaH) database (Wieder et al., 2021) to quantify the influence of environmental drivers of soil C:N. Second, we compared these findings with results simulated by two global scale models that represent coupled C:N stoichiometry, CASA-CNP (Wang et al., 2009) and MIMICS-CN (Wieder et al., 2018). Finally, we used simple regressions to explore variables from the SoDaH database that emerged as important drivers of soil C:N but were absent from models.

2. Methods

We addressed our questions first by exploring patterns in a large-scale synthesis dataset, then by simulating large-scale patterns in models of soil C and N cycling, and finally by comparing patterns across models and data. Although direct comparisons of models and data at the same sites would have been ideal, the number of sites in the measured database with the necessary measured data to run the models was extremely limited, and direct site comparisons would have produced a very limited sample set. Instead, we compared broad patterns in relationships between variables across models and data. In doing so, we took advantage of the wide breadth of measurements in the SoDaH database while deriving unique insights into the controls on soil C:N stoichiometry across these observations and models. Our main interest here was to quantify the broader patterns present in the SoDaH database and evaluate the extent to which the selected soil models show qualitatively similar results.

With this approach, selecting the right variables to explore was an important consideration. Given our expectation that litter input stoichiometry, climate factors, and edaphic factors should all play a role in controlling soil C:N (Fig. 1), we selected litter C:N, mean annual temperature (MAT), and clay content as representative variables that could be compared between models and data. In addition, to explore drivers of soil C:N not represented in models but available at a substantial number of sites in the SoDaH database, we explored measured patterns in soil C:N versus mean annual precipitation (MAP), soil pH, depth, and extractable soil metals indicative of soil mineralogy (i.e. dithionite- or oxalate-extractable iron, aluminum and silica).

2.1 Models

For our model simulations, we chose two global-scale models that represented C and N cycling through soil pools. For comparison's sake, we chose one model that employs a cascade of linearly decomposing soil pools (CASA-CNP), similar to what most Earth system models use, as well as a newer model that controls decomposition and soil formation using simulated microbial growth and biomass (MIMICS-CN). With both models, we simulated steady-state pool values at a global range of sites using a biogeochemical model testbed (Wieder et al., 2018).

2.1.1 Model descriptions

The Carnegie Ames Stanford Approach - Carbon Nitrogen Phosphorus model (CASA-CNP)

The CASA-CNP model follows the basic structure that most Earth system models use to represent surface mineral soils currently, i.e. a cascade of linearly decomposing pools (Wang et al., 2010). The CASA-CNP parameterization determines turnover times and a range for litter and soil C:N ratios for each vegetation type simulated by the model. The soil C:N ratio emerges in CASA-CNP from the distribution of SOM across these pools, which vary in their prescribed stoichiometry.

The MIcrobial-MIneral Carbon Stabilization model with coupled nitrogen cycling (MIMICS-CN)

The MIMICS-CN model simulates both C and N cycles in surface mineral soils using the biomass of two microbial groups with distinct traits (one fast-growing, low-C:N and one slow-growing, higher C:N) to drive rates of decomposition and the transformation of plant litter into stabilized organic matter. MIMICS-CN explicitly represents microbial biomass and metabolism and represents 2 pools of plant litter (considered metabolic and structural) and 3 pools of stabilized organic matter that differ in their persistence: the physically-protected pool is not susceptible to microbial enzymatic degradation and decomposes according to first-order kinetics modified by clay content, the chemically-protected pool is susceptible to microbial degradation and decomposes according to biomass-driven reverse Michaelis-Menten kinetics, and both of these pools feed into an available pool that microbes access and decompose according to biomass-driven reverse Michaelis-Menten kinetics (Kyker-Snowman et al., 2020; Wieder et al., 2014, 2015d). The C:N stoichiometry of the metabolic litter pool is fixed, while the C:N of the structural pool is allowed to vary so that the C and N inputs to these pools are equal to the C and N in unpartitioned total litter inputs. In MIMICS-CN, overall soil C:N is driven by the stoichiometry and activity of the microbes that process plant litter inputs into stable organic matter, the stoichiometry of plant litter inputs to soil, and the distribution of SOM across stabilized pools with different emergent C:N values.

2.1.2 Model simulations

We conducted model simulations to explore the relationship between stoichiometry and other soil variables using the soil biogeochemical model testbed developed in Wieder et al. (2018, 2019). The biogeochemical testbed was developed to allow streamlined testing of multiple models of soil using the same set of globally-gridded daily forcing data, here derived from the Community Land Model version 5 (CLM5; Lawrence et al., 2019) that uses historical climate data from the Global Soil Wetness Project reanalysis (GSWP3; Dirmeyer et al., 2006). Forcing data required to run the testbed included daily estimates of gross primary productivity, air and soil temperature, and soil moisture. Litter input quantity, timing, quality, and C:N stoichiometry was calculated using the vegetation component of CASA-CNP (Wang et al., 2010). Other data inputs required by the soil models included soil clay content, derived from the CLM surface dataset (Lawrence et al., 2019), and plant functional type, derived from the 1-km International Geosphere-Biosphere Program Data and Information System (IGBP DISCover) dataset (Loveland et al., 2000). For more details on the functioning of the testbed, see Wieder et al. (2018, 2019). The testbed and inputs were used to spin up each soil model to steady state for soil organic matter pools in the top meter of soil. The steady-state values of C and N across soil pools were output from each model, summed across pools, and divided to determine whole-soil C:N values for each model and each globally-gridded point. We plotted simple regressions across all locations to illustrate the emergent relationships between the three variables and soil C:N (Figure 2).

2.2 The Soil Data Harmonization database (SoDaH)

The SoDaH database was synthesized from observations taken across a variety of research networks that included the Long-Term Ecological Research (LTER) network, the Nutrient Network (NutNet), the National Ecological Observation Network (NEON), the Critical Zone Observatory (CZO) network, and the Detrital Input and Removal Treatment (DIRT) experiments, with the aim to understand the controls on the distribution, stability, and potential vulnerability of SOM pools (Wieder et al., 2021). The SoDaH database was born from the desire to integrate the deep knowledge about soil properties across multiple research networks into a meaningful data product that could be used to understand soil functioning across broad swaths of time and space. The database includes measurements of over 150 soil variables from 215 sites across the world, and the database is unique in its inclusion of time series data and environmental manipulation experiments across many of those sites. The specific measurements available for a given location vary widely, but the harmonization process used to generate the database makes it possible to parse the data for the variables of interest and look at broad patterns.

2.2.1 Statistical analyses of C:N and edaphic variables in SoDaH

To assess our model-generated relationships against measurements in the SoDaH database, we filtered the database to exclude values that fell outside the assumptions of our model simulations. The version of MIMICS-CN we used was designed to simulate surface mineral soils, so we removed organic horizons from the database and filtered the database to exclude measurements from experimentally treated samples, leaving only control samples. We also restricted our analyses to the top 20 cm of soil to limit the impact of depth on our results, leaving 1204 observations. Our models simulated soil down to 1 m but only as a single layer averaging across the entire depth profile; therefore, we expected to see some differences between model results and observations resulting from the difference in represented depths (i.e. bias in model results toward lower C:N values), but we still felt this was a fair comparison. Plant litter C:N values were somewhat sparse in the database, but we generated values for many of the NEON sites by averaging the C:N of aboveground leaf and needle inputs. We generated the same simple regressions as for our model results of soil C:N against plant litter C:N, mean annual temperature, and percent clay (Figure 2). We then explored additional correlations within the database and identified several variables other than the three already selected that were worth including in further analyses with the database, namely depth, mean annual precipitation (MAP), pH, and extractable metals.

2.3 Data-model comparison with linear mixed models

To obtain a deeper quantitative understanding of how different variables ranked in their importance in driving soil C:N, we developed a suite of multiple linear regressions (MLR) that we used with different subsets of the data and different sets of our variables. Importantly, the results of these models are only comparable within a given data subset because MLR are only comparable for the same set of response variables, but our goal was not to compare values directly but rather to compare the relative importance of variables in driving soil C:N across these different statistical models. Additionally, measurements of litter C:N were only available for a smaller subset of our data (primarily sites from the National Ecological Observation Network or NEON), so we chose to evaluate the importance of two other variables (clay, MAT) across the larger dataset while exploring litter C:N patterns within the smaller subset of data from NEON sites. For each MLR, we scaled variables to have a mean of 0 and a standard deviation of 1 so the regression coefficients would be comparable, and we checked that the residuals followed a normal distribution.

On the entire SoDaH dataset, we ran two MLRs: one which contained our two variables of interest related to the testbed models (clay, MAT), and one which additionally included three other variables that showed up as potentially important in our initial explorations of the data (pH, MAP, depth down to 20 cm). For the smaller subset of NEON data (for which we had fewer measurements of a larger number of variables), we ran four MLRs: two with the same variables as the models used with the entire dataset, one with the variables available within our models (clay, MAT, litter C:N), and one with a wider array of variables available only within this smaller subset of data (clay, MAT, litter C:N, pH, MAP, depth, and also extractable metals).

For our testbed model results, we generated two MLRs each: one with the two variables shared with the larger dataset (clay, MAT) and one with the three variables shared with the smaller dataset (clay, MAT, litter C:N).

For all of our statistical models, we generated relative importance values using the relaimpo package in R (Grömping, 2006) to understand how different variables controlled soil C:N values for a given subset of the data and variables. We reported the relative importance for each independent variable, which is the percentage of the total explained variance that a given individual variable explains. We compared model and data results as appropriate.

Finally, based on the statistical model results, we generated additional simple regressions of variables that emerged as important drivers of soil C:N to understand patterns in the SoDaH data that models miss.

3. Results

We explored important drivers of soil C:N across models and data. Our results emphasize the role of clay content in driving soil C:N in both data and the MIMICS model, while litter C:N and MAT were less important in data but more important in model results. In our in-depth analysis of the SoDaH database, the explained variance and the relative importance of each variable differed across subsets of the data and variables, but other factors that emerged as important included depth, pH, and extractable soil metals.

3.1 Broad patterns in soil stoichiometry across data and models

3.1.1. Simple regressions

We used simple regressions to give a sense of how models and measured data compared (Figure 3). Results diverged among measured values and the two soil models we included. Across the simple regressions in Figure 3, the three representative variables we chose (litter C:N, MAT, clay) explained considerably more variation in the MIMICS-CN than the data or the CASA-CNP model, and the measured correlations were not strong overall across such a large number of observations. Measured soil C:N was weakly positively correlated with plant litter C:N ($R^2 = 0.19$) and weakly negatively correlated with percent clay ($R^2 = 0.15$), but mean annual temperature did not seem to show any specific relationship with soil C:N ($R^2 = 0.005$). Relative to data, values of soil C:N simulated by MIMICS-CN were more strongly correlated with all three variables, and correlation coefficients were comparable across variables. By comparison, values of soil C:N simulated by CASA-CNP were weakly correlated with all three variables.

In general, observations captured greater natural variability in our simple regressions than the MIMICS-CN model, which is unsurprising given that models are necessarily simplified representations of the natural world. The range of values for C:N generated by MIMICS-CN was very small and biased towards lower, more nitrogen-enriched values, generally falling between 9 and 13. While many real soils do fall in this range, the natural measured range in soil C:N and the range captured by CASA-CNP was much greater. In addition, our observations were limited to the top 20 cm of soil while our models both simulated soils down to 1 m, so it would be reasonable to models to be biased towards lower soil C:N values and a different range than observations.

3.1.2 MLR and relative importance results

Our MLR and relative importance analyses allowed us to clarify relationships between variables in our model outputs and data results. Similar to our simple regressions, results diverged across models and data. Within measured results, no MLR generated an R^2 greater than 0.52, indicating that other unmeasured variables likely contribute to the patterns explored here (Table 1). The lowest R^2 was 0.15 for the MLR model that used only clay and MAT to explain variance across the entire SoDaH database. MLR models that included more explanatory variables or that were applied only to the NEON subset of the SoDaH database generally explained more of the overall variance in the data (Table 1).

Despite generally low R^2 values, across every MLR model using observations, clay emerged as a variable with high relative importance in driving soil C:N (Table 1; 33.4% of the total variance explained in the most parsimonious model using the larger dataset). As a reminder, we focused on relative importance values as indications of the contributions of individual variables in driving soil C:N, where "relative importance" is the percentage of the total explained variance that a given individual variable explains. In the entire SoDaH dataset (SoDaH-5), MAP and pH also emerged as important drivers (33.6% and 31.4% respectively), while depth and MAT were of lesser importance (0.3% and 1.3% respectively). Within the NEON subset of the SoDaH data, the same variables still emerged as important drivers of soil C:N, but MAT also showed moderate importance. Dithionite-extractable iron and aluminum also emerged as moderately important drivers in the statistical model that included them, while other variables did not have much relative importance.

The two soil models we evaluated showed MLR and relative importance results that differed both from measured results and from each other. In MIMICS-CN, MAT, clay and litter C:N held roughly equal relative importance in driving soil C:N (31.1%, 31.4%, and 37.5%, respectively), and these three variables explained a large overall proportion of the variance in model results (R^2=0.80). In CASA-CNP, on the other hand, these three variables explained a very small proportion of the overall variance in model results (R^2=0.06), and the results indicated that litter C:N was the most important driver of soil C:N (64.6%), while MAT was roughly half as important (32.7%) and clay was hardly important at all (2.8%), though caution should be taken in interpreting the relative importance of variables that explained so little variance.

3.2 Key drivers of C:N absent from models

Based on the MLR results, we delved deeper into the relationships between soil C:N, MAP, pH, and extractable metals to assess whether there are important drivers of soil C:N that models miss. We plotted the relationship between soil C:N and both MAP and MAT to discern any obvious climate-driven patterns; however, no clear pattern emerged (Fig. 4). On the other hand, we plotted the relationship between soil C:N and both pH and clay (Fig. 5) and observed a response surface with high soil C:N values for acidic, low-clay soils and declining soil C:N with both increasing clay content and increasing pH. pH and clay content impacted soil C:N independently, such that there were an abundance of both high-clay, low-pH and low-clay, high-pH sites with low soil C:N (i.e. clay and pH did non covary).

Finally, we plotted simple regressions of soil C:N against extractable metals (Fig. 6) to better understand the relationships between soil C:N and both dithionite-extractable iron and aluminum, which both emerged as important variables in our MLR analysis. Soil C:N showed a weakly positive correlation with dithionite-extractable aluminum and a weakly negative relationship with dithionite-extractable iron, though both correlations were too weak to draw strong conclusions.

4. Discussion

Soil C:N stoichiometry is an easily-measured quantity that provides a window into the chemical nature of SOM and allows us to evaluate the performance of soil models that couple C and N cycles in a unique way relative to soil C or N stocks alone. Our analyses allowed us to quantify the relative importance of top-down and bottom-up controls on soil C:N (Fig. 1) in modern models of soil C and N and in data, and to explore important variables missing from models. We evaluated modeled and measured drivers of soil C:N using global simulations of two distinct models and a new global database of soil variables. We discerned that clay content was an important control on soil C:N in both data and one of our models, and that litter and climate variables were differently important in data compared to our models. We also discovered a number of other drivers of soil C:N in the database, notably pH and several extractable metals.

4.1 Broad patterns in soil stoichiometry across data and models

The relative strength of different variables driving soil C:N across model results and data points out both strengths and flaws in the design of models. In SoDaH-CN, all of the statistical models we evaluated explained only some of the variance in the data, with R^2 values that consistently fell higher than CASA-CNP but lower than MIMICS-CN. Therefore, neither CASA-CNP nor MIMICS-CN simulations fell precisely in line with measured data: CASA-CNP underemphasized the importance of the included drivers of soil C:N, while MIMICS-CN overemphasized them. Within the limited variance explained in the measured data by our included variables, clay content consistently generated the highest relative importance, with climate and litter C:N variables appearing with relative importance values that were roughly half of clay content. pH and some extractable metals also showed relative importance values that fell between clay and climate variables, further emphasizing the role of edaphic variables in controlling soil stoichiometry. Although unmeasured variables explain a considerable portion of the variance in soil C:N in SoDaH, our results overall highlight the primary importance of bottom-up controls on soil C:N (Fig. 1), with top-down litter and climate controls appearing secondarily important.

For MIMICS-CN, clay content, MAT and litter C:N were roughly equal in their relative importance, meaning that MIMICS-CN overemphasized the role of both MAT and plant litter chemistry in driving soil C:N relative to data. This makes sense given that MIMICS-CN strongly emphasizes the role of microbial processing in transforming high-CN plant litter into low-CN stable organic matter. MAT controls the rate of microbial processing of litter in MIMICS-CN, which in turn pushes the distribution of C and N at high-temperature sites away from high-CN plant litter pools and towards low-CN stable organic matter pools. Coupled with the fact that MIMICS-CN generated soil C:N values in a range that was considerably lower than the measured range, MIMICS-CN may be too heavily weighted towards the microbial end of the soil spectrum and may need to be adjusted to stabilize more unprocessed plant residues, perhaps through some kind of soil aggregation mechanism. Real-world variability in soil C:N was much greater than MIMICS-CN would suggest, though this is often a hazard of models.

Results simulated with CASA-CNP diverged from both MIMICS-CN and measured data, indicating fundamental differences in the theory underlying the design of CASA-CNP. In CASA-CNP, the C:N stoichiometry of various soil pools is prescribed as internal parameters to the model, and total soil stoichiometry for a given site is dictated by the balance of SOM across these pools. As a result, the three variables included here as hypothesized drivers of soil stoichiometry (MAT, litter C:N, clay) explained hardly any of the variance in soil C:N in CASA-

CNP simulations ($R^2 = 0.06$). This points to a flaw in the cascading pool design of CASA-CNP and models like it: soil stoichiometry is disconnected from many environmental variables and is not allowed to vary with variables that are known to be important drivers of soil C:N. This may limit the ability of models with this design to accurately capture the feedbacks between soil C and N cycles and may cause compensating errors in the simulation of each elemental cycle individually (for example, errors in the estimation of plant-available N released from decaying SOM).

Although our initial data-model comparisons point to some potential strengths and areas for improvement in existing models, confirmation of the widespread controls on soil C:N is limited by the availability of soil measurements aligned at the same sites. In particular, despite the breadth of sites, measurements, and research networks included in SoDaH, very limited measurements were available for plant litter chemistry at sites where soil C:N was also measured. This is an inexpensive and straightforward problem to correct if more empirical studies would measure both soil C:N and plant litter C:N (finding the true C:N of total plant litter inputs is complicated by roots, root exudates, and various other pathways for plant inputs to soil, but even a rough estimate of aboveground plant input C:N would be a useful proxy for evaluating models). In addition, our study pointed to the importance of edaphic variables in driving soil C:N. More widespread measurements of specific soil mineralogy coupled to detailed mechanistic studies exploring the affinities of different minerals for N-enriched organic moieties (e.g. amino acids) may provide clarity about the role of edaphic factors in filtering SOM and enriching soil C:N.

4.2 Key drivers of C:N absent from models

Our results also pointed to pH and extractable metals as important measured variables controlling soil C:N that are absent from the models we used. pH is a master variable in soil that controls many processes, from mineral and dissolved elemental reactivity (Neina, 2019; Rasmussen et al., 2018; Weil and Brady, 2016) to the makeup of microbial communities in soil (Fierer and Jackson, 2006). One interesting interpretation of our results relates to modern research highlighting the important role of soil mineralogy in determining how much and which SOM compounds become stabilized (Mikutta et al., 2019). Coupled with the result that clay content was an important driver of soil C:N across both MIMICS-CN and data, the correlation of soil C:N and pH might suggest that specific mineralogy (and not simply clay-sized minerals) may be important in stabilizing high-N residues from dissolved organic matter and microbial necromass. Mineral reactivity and importance in stabilizing SOM has been shown to depend on pH (Rasmussen et al., 2018), which may help to explain our results. Previous research has also shown that soil minerals show preferential affinity and bonding strengths for specific organic moieties, especially N-rich moieties like amino groups, and act as a bottom-up filter on the chemical makeup of stabilized SOM (Mikutta et al., 2019). Changes in organo-mineral associations with depth have been used to explain observed depth patterns in soil C:N and ^{15}N isotopic enrichment, in contrast to the historical view that soil C:N decreases with depth are due to increasing degree of residue processing (Kramer et al., 2017). This concept of a "mineral filter" (Mikutta et al., 2019) acting as a bottom-up control on SOM composition is supported overall by our analyses (i.e. the high relative importance of clay, pH and specific extractable metals) and could be explored further in models with minor modifications.

The hypothesis that specific mineral interactions control soil C:N through sorptive fractionation could be explored in models like MIMICS-CN with some additional data collection. Further field experiments could be used to develop a mechanistic relationship between pH and the minerals that are important in stabilizing certain forms of organic matter. Using such a relationship, pH is an easily measured quantity that could be used to improve models, for example by making the model coefficient of clay stabilization dependent on pH. The importance of dithionite-extractable iron in driving soil C:N in our results also supports the idea that specific minerals may play a role in filtering the chemistry of SOM during stabilization. Chemical extractions are more expensive and somewhat rarer in measured data, but such relationships may be useful in identifying the specific minerals (e.g. iron oxides) that play a key role in stabilizing low-C:N microbial residues. Specific mineral representation is currently beyond the scope of models like MIMICS-CN and CASA-CNP, but understanding the role of different minerals in stabilizing SOM on a large scale could improve the theory of SOM stabilization in a way that could eventually be simplified for use in models.

4.3 Summary and future directions

Soil C and N cycles interact and modulate one another, and we cannot accurately understand or model either in isolation. As models increasingly connect C and N cycles, it is crucial that we evaluate the mechanistic connections in models and data and ensure that they are in accord. In our study, we evaluated one traditionally-structured model and one contemporary microbially-explicit model against a large scale data synthesis. We found that neither model perfectly captured the mechanistic controls on soil C:N in the data: the traditionally-structured model failed to capture the influence of climate, litter or edaphic control over C:N, while the contemporary model overemphasized all three. Moving forward, broader measurements of the variables that control soil C:N and deeper experiments into the mineral-mediated mechanistic controls on SOM stoichiometry will help to connect data and models and reduce uncertainties in large-scale coupled biogeochemical simulations.

Table 1. Results from multiple linear regression analyses of SoDaH data, NEON sites included within SoDaH, MIMICS-CN and CASA-CNP. For each MLR, we scaled variables to have a mean of 0 and a standard deviation of 1 so the regression coefficients would be comparable. Relative importance percentages show the percentage of the total variance explained by each statistical model that a given individual variable explains. "NA" indicates a variable that was not included in a given analysis (or statistical model). Statistical models with the same "Grouping" letter included the same subset of variables.

						Relative importance percentage											
MLR name	Type	n	R^2	AIC	Grouping	MAT	Clay	Litter C:N	MAP	Depth	pH	Fe_ox	Al_ox	Ai_ox	Fe_dith	Si_dith	Al_dith
SoDaH-2	Observation	1204	0.15	3195	a	2.0%	98.0%	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
SoDaH-5	Observation	1204	0.27	2669	c	1.3%	33.4%	NA	33.6%	0.3%	31.4%	NA	NA	NA	NA	NA	NA
NEON-2	Observation	239	0.25	616	a	31.7%	68.3%	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
NEON-3	Observation	239	0.28	607	b	20.9%	46.5%	32.6%	NA	NA	NA	NA	NA	NA	NA	NA	NA
NEON-5	Observation	239	0.32	600	c	20.0%	37.3%	NA	15.8%	2.4%	24.5%	NA	NA	NA	NA	NA	NA
NEON-7	Observation	239	0.52	386	d	9.7%	15.5%	9.4%	8.7%	1.9%	10.5%	4.5%	9.3%	4.9%	11.1%	1.8%	12.5%
MIMICS-CN-2	Model	2697	0.54	5554	a	51.6%	48.4%	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
MIMICS-CN-3	Model	2697	0.80	3318	b	31.1%	31.4%	37.5%	NA	NA	NA	NA	NA	NA	NA	NA	NA
CASA-CNP-2	Model	2697	0.03	7588	a	92.8%	7.2%	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
CASA-CNP-3	Model	2697	0.06	7500	b	32.7%	2.8%	64.6%	NA	NA	NA	NA	NA	NA	NA	NA	NA

Figure 1. Current state of knowledge about how plant, microbial, climate, and mineralogical factors combine to influence the coupling of soil C and N biogeochemical cycles.

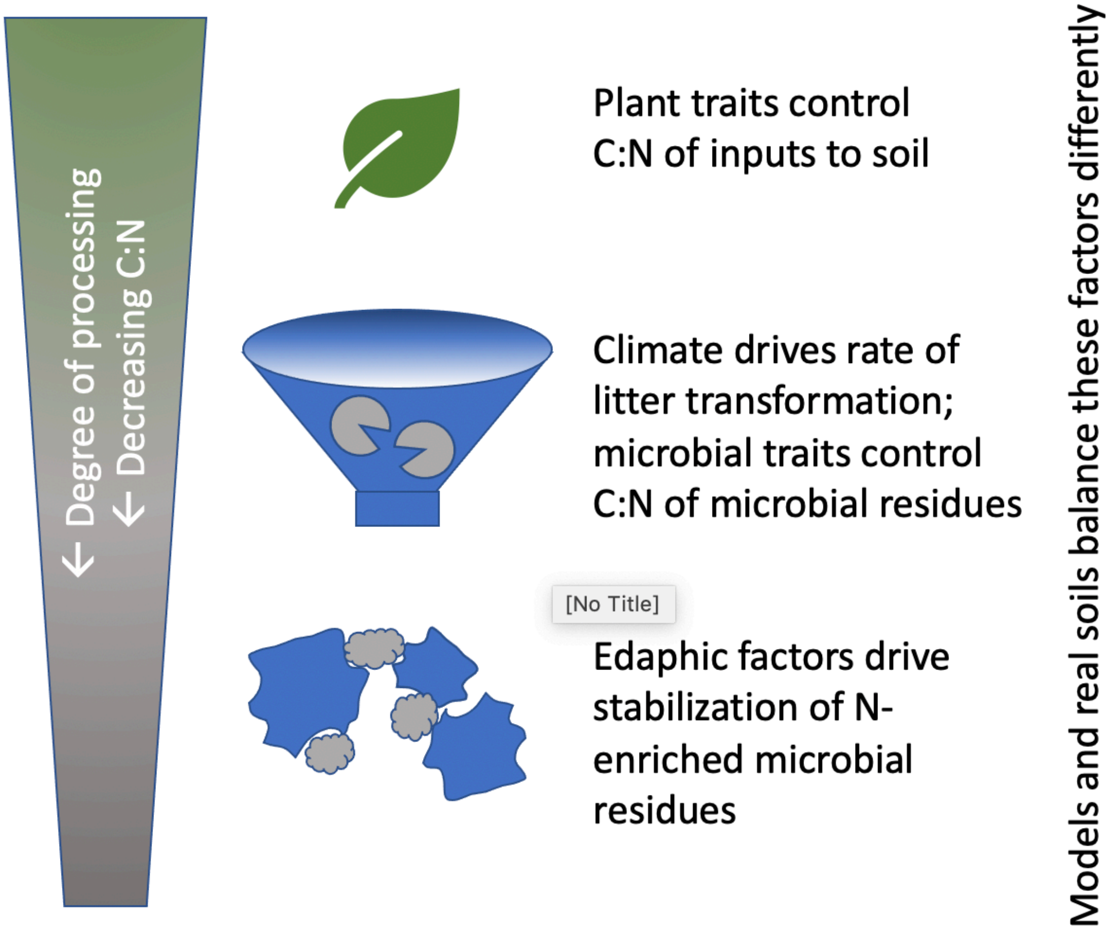


Figure 2. Overlapping measured variables between data subsets and model inputs.

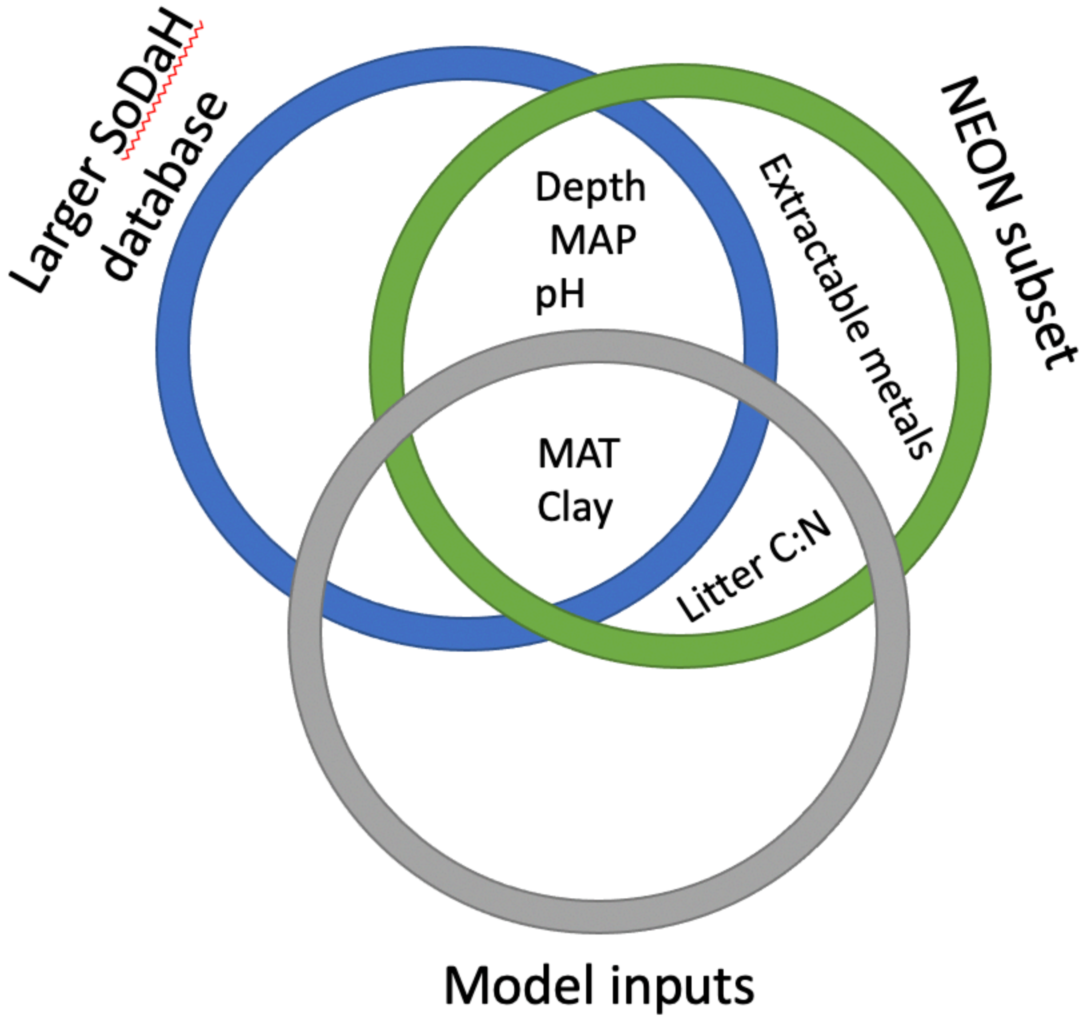


Figure 3. Drivers of soil C:N across the SoDaH database (top row) and the MIMICS-CN and CASA-CNP models (bottom row). Lines and correlation coefficients show linear regressions of the data. Mean annual temperature (MAT) is given in units of degrees Celsius.

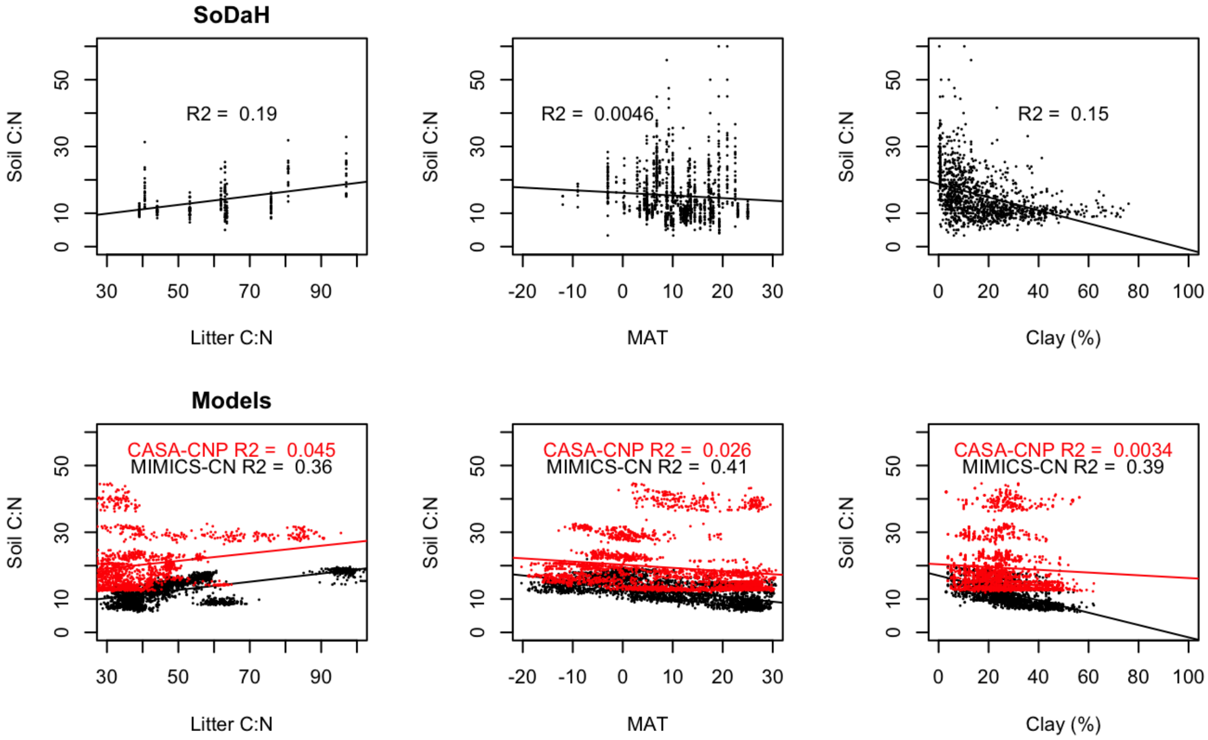


Figure 4. Soil C vs. N as a function of mean annual precipitation (MAP, mm) and mean annual temperature (MAT, degrees Celsius) in the SoDaH database.

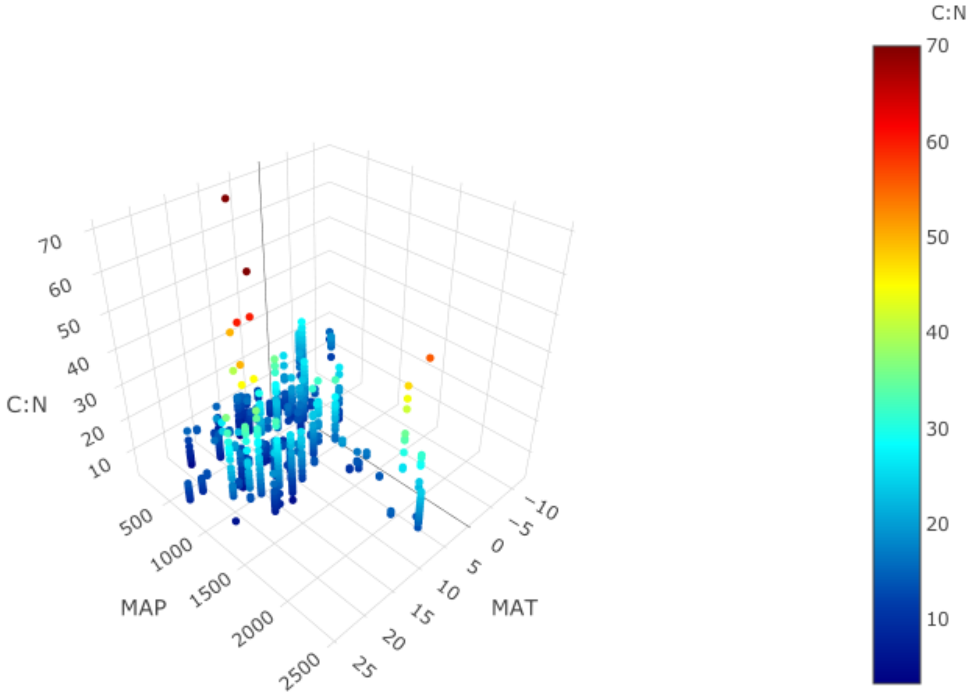


Figure 5. Soil C vs. N as a function of pH and clay in the SoDaH database. Lines show linear regressions of the data.

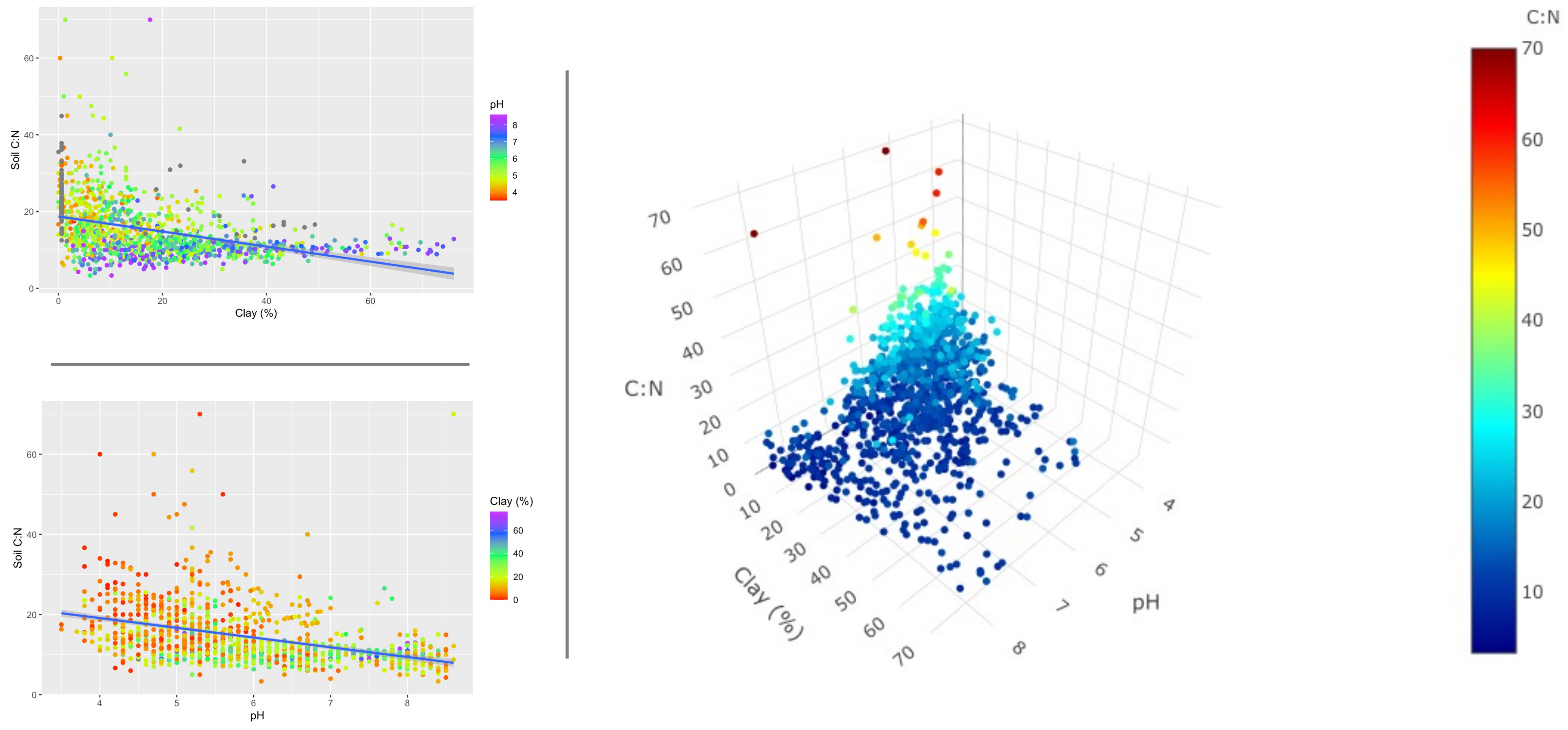
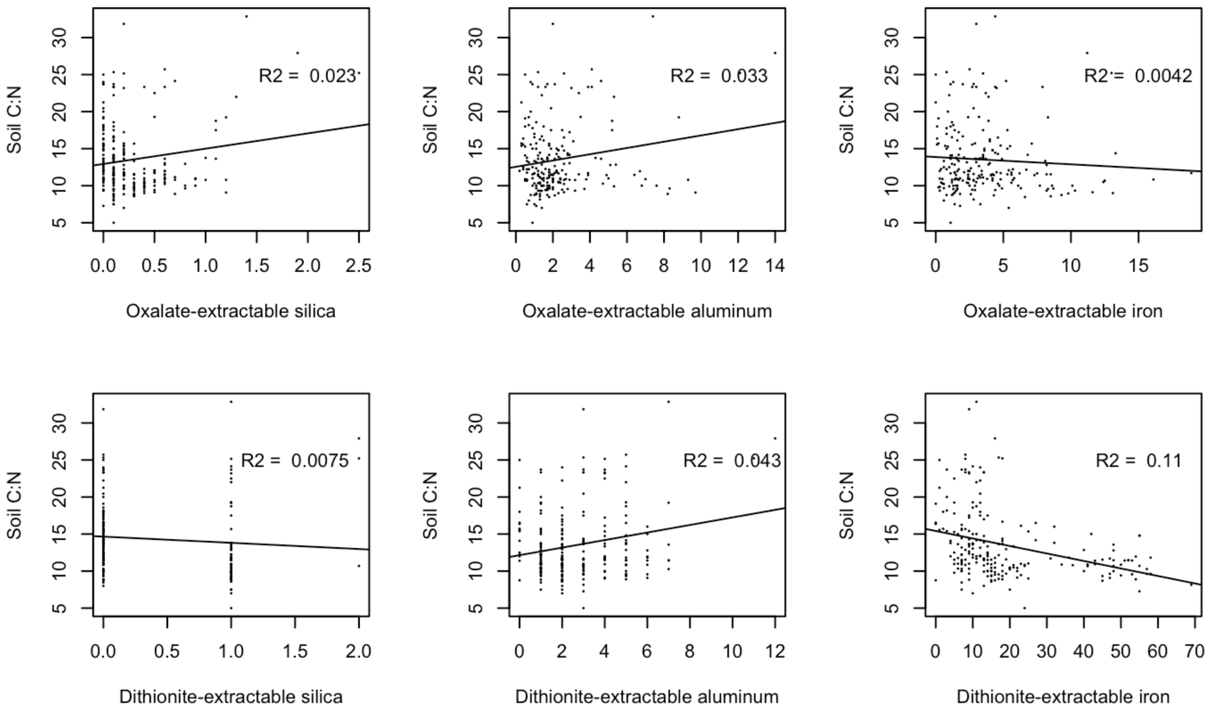


Figure 6. Soil C vs. N as a function of extractable metals in the SoDaH database. Lines show linear regressions of the data. All extractable metals shown are in units of mg/g soil.



SYNTHESIS

Ecology matters in the Earth system. Plants, fauna, and microorganisms on the land surface control terrestrial flows of water, elements like C and N, and energy, with extensive feedbacks to climate processes. For example, plant communities determine the rates of terrestrial exchanges of CO_2 and water through photosynthesis (Bonan, 2016), and plant-soil feedbacks determine the rates that plants can take up soil N and facilitate photosynthesis (Terrer et al., 2018; Zaehle et al., 2014). Soil microbial communities determine rates of decomposition and release of C and nutrients from plant litter (Gan et al., 2013; Kaiser et al., 2014; Nottingham et al., 2013; Rinkes et al., 2013), and microbial community composition dictates the impact of climate factors on rates of decomposition (Averill et al., 2015; Blagodatskaya et al., 2010; Edwards and Zak, 2011; Kaiser et al., 2015). Ecological communities even control rates of energy transfer with the land surface – for example, plant community composition and phenology determine absorption of radiative energy through the albedo effect and land surface modification of atmospheric currents through surface roughness (Bonan, 2008, 2016). We can therefore improve the realism of ESMs by adding more ecological process representations, but we must be careful not to add them blindly.

Adding new ecological processes to ESMs without careful consideration can damage the performance of models in a variety of ways. Overly complex calculations can increase the computational costs of ESMs to an unreasonable degree. New processes that have not been adequately validated can introduce new errors to models that propagate throughout other connected processes, leading ultimately to decreasing accuracy in the projection of future behavior in terrestrial ecosystems. Even if a process has performed well against data, models sometimes suffer from the principle of "equifinality" whereby many model formulations can reproduce the same data, and models sometimes reproduce data well using an inaccurate underlying process representation. This principle highlights the need to constantly evaluate and improve the underlying process representations in existing ESMs to reflect emerging empirical and theoretical understanding of terrestrial ecosystems.

Given the inherent hazards of adding new process representations to ESMs, it is critical to mitigate these risks via thoughtful procedures for evaluating new process models along the way to ESM incorporation. Chapter 1 of this dissertation describes such a pathway made up of three stages: initial process assessment, simple model development, and model testing within an ESM. As described in Chapter 1, the path to ESM incorporation is nonlinear, often requiring revisiting intermediate steps before proceeding to the next stage. Following a procedure like that described in Chapter 1 allows large-scale models to grow and develop in a responsible way.

Specific examples help to illustrate model development pathways like that described on Chapter 1. Models of soil biogeochemistry provide a particularly good example of the iterative nature of model development, both within and outside ESMs. Historical process models of soil (and most of the soil models used within ESMs today) follow a decomposition scheme whereby plant matter decays linearly along a cascade of increasingly stable pools. This scheme is based on the outmoded idea that chemical recalcitrance is primarily responsible for the long-term stability of SOM, and the models that are built off this method are slowly falling out of favor to make way for models that more faithfully represent microbial controls over the transformation and stabilization of SOM. As soil models evolve, it’s critical to interrogate new process representations with a variety of data and from a range of perspectives to ensure that one flawed process representation is not being replaced with something equally flawed.

Chapters 2 and 3 of this dissertation illustrate responsible evaluation of a novel soil process representation. In Chapter 2, I described the initial development of a coupled C and N version of the MIMICS model, followed by evaluation of the model's performance against a continent-wide database of litter decomposition experiments. This evaluation provided a baseline for understanding the new model's ability to dynamically replicate patterns in both C and N decomposition across a range of climate and soil conditions. Although this was only a first step in examining of the performance of the new model, it served as compelling evidence that the model fundamentally functions and reasonably captures real decomposition dynamics, and furthermore suggests that the model is worthy of additional development and examination using diverse, large-scale datasets.

In Chapter 3 of this dissertation, I continued the evaluation of the MIMICS-CN model by comparing its performance against a more traditional model structure (CASA-CNP) and a large-scale database (SoDaH). In this chapter, I focused on the ratio of soil C and N in models and data and used statistical analysis of model results and data to discern the most important drivers of soil C:N. I learned that MIMICS-CN balances temperature, plant litter chemistry, and soil clay content roughly equally as controls over soil C:N, while CASA-CNP showed little control of any of these three variables over soil C:N, and measured values showed an intermediate level of control of these three values over soil C:N with an emphasis on clay content over the other two. On the whole, this suggests that MIMICS-CN captures controls over soil C:N that a more traditional model (CASA-CNP) fundamentally misses, but that MIMICS-CN fails to capture all of the variability and drivers of soil C:N present in measured data.

Development and testing with the MIMICS model extends considerably beyond the work in this dissertation. Recent work has added soil moisture controls on microbial growth (Wieder et al., 2018) and depth resolution (Wang et al., 2021) to MIMICS. Furthermore, the parameter estimates used in MIMICS have been refined (Zhang et al., 2020) and the model's overall performance has been extensively tested from several angles against global soils data and other soil models (Basile et al., 2020; Koven et al., 2017; Shi et al., 2018; Sulman et al., 2018). As the complexity of MIMICS and the variables it includes expands, it is crucial that the model is both developed and then tested using data that represent our most modern understanding of the functioning of soil systems.

In addition to the included variables discussed above, MIMICS is uncommon amongst soil models in its capacity to represent aspects of microbial community and physiology. Given the rapid expansion of the field of soil microbiology (Fierer et al., 2021), future work could use MIMICS to explore the large-scale ramifications of global-change-driven shifts in microbial communities. For example, the r and K strategists represented in MIMICS may respond differently to shifting patterns in precipitation and soil moisture, and MIMICS enables us to connect such responses to overall responses of soil C and N cycles. Beyond the simplified representation of r and K strategists in MIMICS, other logical divisions of microbes into functional groups may emerge from soil microbiological research that could be incorporated into MIMICS. In addition, various aspects of microbial physiology and metabolism (such as CUE or growth rate) may have different sensitivities to global change factors like rising temperatures. MIMICS provides a unique testbed for exploring the large-scale ramifications of microbial responses to global change for soil C and N cycling around the world.

Many forms of ecology, from soil metagenomics to plant-soil nutrient interactions and beyond, have the potential to improve the realism of ESMs and ultimately increase our ability to project and understand the future of our changing planet. As we do so, it is critical that all new model developments undergo rigorous testing at every stage to ensure that process representations ranging in scale from micro- to macroscopic align with our empirical understanding of Earth system mechanics. This dissertation has outlined a roadmap for the careful progress of models along the path to ESMs and demonstrated this roadmap with the development and multi-faceted evaluation of a microbial-explicit model of soil C and N cycling. With the paired forward progress of empirical understanding and modeled process representation in the future, ESMs and ecological research in soils and beyond will strengthen one another and jointly improve our understanding of the functioning of natural systems.

LIST OF REFERENCES

Abramoff, R. Z., Davidson, E. A. and Finzi, A. C.: A parsimonious modular approach to building a mechanistic belowground carbon and nitrogen model, J. Geophys. Res. Biogeosciences, 122(9), 2418–2434, doi:10.1002/2017JG003796, 2017.

Adair, E. C., Parton, W. J., Del Grosso, S. J., Silver, W. L., Harmon, M. E., Hall, S. A., Burke, I. C. and Hart, S. C.: Simple three-pool model accurately describes patterns of long-term litter decomposition in diverse climates, Glob. Chang. Biol., 14(11), 2636–2660, doi:10.1111/j.1365-2486.2008.01674.x, 2008.

Adams, G. S., Converse, B. A., Hales, A. H. and Klotz, L. E.: People systematically overlook subtractive changes, Nature, 592(7853), 258–261, doi:10.1038/s41586-021-03380-y, 2021.

Ainsworth, E. A. and Long, S. P.: What have we learned from 15 years of free-air CO2enrichment (FACE)? A meta-analytic review of the responses of photosynthesis, canopy properties and plant production to rising CO2, New Phytol., 165(2), 351–372, doi:10.1111/j.1469-8137.2004.01224.x, 2005.

Allison, S. D.: Modeling adaptation of carbon use efficiency in microbial communities, Front. Microbiol., 5, 1–9, doi:10.3389/fmicb.2014.00571, 2014.

Anderson-Cook, C. M., Lu, L. and Parker, P. A.: Effective interdisciplinary collaboration between statisticians and other subject matter experts, Qual. Eng., 31(1), 164–176, 2019.

Auguie, B.: gridExtra: functions in Grid graphics. R Package Version 2.3, CRAN Proj., 2017.

Averill, C. and Waring, B.: Nitrogen limitation of decomposition and decay: How can it occur?, Glob. Chang. Biol., (June), 1–11, doi:10.1111/gcb.13980, 2017.

Averill, C., Rousk, J. and Hawkes, C.: Microbial-mediated redistribution of ecosystem nitrogen cycling can delay progressive nitrogen limitation, Biogeochemistry, 126(1–2), 11–23, doi:10.1007/s10533-015-0160-x, 2015.

Bailey, V. L., Bond-Lamberty, B., DeAngelis, K., Grandy, A. S., Hawkes, C. V, Heckman, K., Lajtha, K., Phillips, R. P., Sulman, B. N., Todd-Brown, K. E. O. and Wallenstein, M. D.: Soil carbon cycling proxies: Understanding their critical role in predicting climate change feedbacks, Glob. Chang. Biol., 24(3), 895–905, doi:10.1111/gcb.13926, 2018.

Bala Chaudhary, V. and Berhe, A. A.: Ten simple rules for building an antiracist lab, PLoS Comput. Biol., 16(10), 1–9, doi:10.1371/journal.pcbi.1008210, 2020.

Basile, S. J., Lin, X., Wieder, W. R., Hartman, M. D. and Keppel-aleks, G.: Leveraging the signature of heterotrophic respiration on atmospheric CO 2 for model benchmarking, , 1293–1308, 2020.

Baskaran, P., Hyvönen, R., Berglund, S. L., Clemmensen, K. E., Ågren, G. I., Lindahl, B. D. and Manzoni, S.: Modelling the influence of ectomycorrhizal decomposition on plant nutrition and soil carbon sequestration in boreal forest ecosystems, New Phytol., 213(3), 1452–1465, doi:10.1111/nph.14213, 2017.

Batjes, N. H.: Total carbon and nitrogen in the soils of the world, Eur. J. Soil Sci., (47), 151–163, 1996.

Berg, B.: Litter decomposition and organic matter turnover in northern forest soils, For. Ecol. Manage., 133, 13–22, 2000.

Bernard, R. E. and Cooperdock, E. H. G.: No progress on diversity in 40 years, Nat. Geosci., 11(5), 292–295, doi:10.1038/s41561-018-0116-6, 2018.

Blagodatskaya, E., Blagodatsky, S., Dorodnikov, M. and Kuzyakov, Y.: Elevated atmospheric CO2 increases microbial growth rates in soil: results of three CO2 enrichment experiments, Glob. Chang. Biol., 16(2), 836–848, doi:10.1111/j.1365-2486.2009.02006.x, 2010.

Blagodatskaya, E., Blagodatsky, S., Anderson, T. H. and Kuzyakov, Y.: Microbial growth and carbon use efficiency in the rhizosphere and root-free soil, PLoS One, 9(4), doi:10.1371/journal.pone.0093282, 2014.

Blankinship, J. C., Berhe, A. A., Crow, S. E., Druhan, J. L., Heckman, K. A., Keiluweit, M., Lawrence, C. R., Marín-Spiotta, E., Plante, A. F., Rasmussen, C., Schädel, C., Schimel, J. P., Sierra, C. A., Thompson, A., Wagai, R. and Wieder, W. R.: Improving understanding of soil organic matter dynamics by triangulating theories, measurements, and models, Biogeochemistry, 2, doi:10.1007/s10533-018-0478-2, 2018.

Bonan, G. B.: Land-atmosphere CO2 exchange simulated by a land surface process model coupled to an atmospheric general circulation model, J. Geophys. Res., 100(D2), 2817–2831, doi:10.1029/94JD02961, 1995.

Bonan, G. B.: Ecological Climatology: Concepts and Applications, 2nd Edition, Geogr. Res., 48(2), 221–222, doi:10.1111/j.1745-5871.2009.00640.x, 2008.

Bonan, G. B.: Forests, Climate, and Public Policy: A 500-Year Interdisciplinary Odyssey, Annu. Rev. Ecol. Evol. Syst., 47(1), 97–121, doi:10.1146/annurev-ecolsys-121415-032359, 2016.

Bonan, G. B.: Climate Change and Terrestrial Ecosystem Modeling, Cambridge University Press., 2019.

Bonan, G. B. and Doney, S. C.: Climate, ecosystems, and planetary futures: The challenge to predict life in Earth system models, Science (80-.)., 359(6375), doi:10.1126/science.aam8328, 2018.

Bonan, G. B. and Levis, S.: Quantifying carbon-nitrogen feedbacks in the Community Land Model (CLM4), Geophys. Res. Lett., 37(7), 1–6, doi:10.1029/2010GL042430, 2010.

Bonan, G. B., Levis, S., Sitch, S., Vertenstein, M. and Oleson, K. W.: A dynamic global vegetation model for use with climate models: Concepts and description of simulated vegetation dynamics, Glob. Chang. Biol., 9(11), 1543–1566, doi:10.1046/j.1365-2486.2003.00681.x, 2003.

Bonan, G. B., Lawrence, P. J., Oleson, K. W., Levis, S., Jung, M., Reichstein, M., Lawrence, D. M. and Swenson, S. C.: Improving canopy processes in the Community Land Model version 4 (CLM4) using global flux fields empirically inferred from FLUXNET data, J. Geophys. Res., 116(G2), 1–22, doi:10.1029/2010jg001593, 2011.

Bonan, G. B., Hartman, M. D., Parton, W. J. and Wieder, W. R.: Evaluating litter decomposition in earth system models with long-term litterbag experiments: an example using the Community Land Model version 4 (CLM4)., Glob. Chang. Biol., 19(3), 957–74, doi:10.1111/gcb.12031, 2013.

Bowles, T. M., Jackson, L. E. and Cavagnaro, T. R.: Mycorrhizal fungi enhance plant nutrient acquisition and modulate nitrogen loss with variable water regimes, Glob. Chang. Biol., 24(1), e171–e182, doi:10.1111/gcb.13884, 2018.

Bracco, A., Long, M. C., Levine, N. M., Thomas, R. Q., Deutsch, C. and McKinley, G. A.: NCAR'S summer colloquium: Capacity building in cross-disciplinary research of earth system carbon-climate connections, Bull. Am. Meteorol. Soc., 96(8), 1381–1384, doi:10.1175/BAMS-D-13-00246.1, 2015.

Bradford, M. A., Keiser, A. D., Davies, C. A., Mersmann, C. A. and Strickland, M. S.: Empirical evidence that soil carbon formation from plant inputs is positively related to microbial growth, Biogeochemistry, 113(1–3), 271–281, doi:10.1007/s10533-012-9822-0, 2013.

Bradford, M. A., Wieder, W. R., Bonan, G. B., Fierer, N., Raymond, P. A. and Crowther, T. W.: Managing uncertainty in soil carbon feedbacks to climate change, Nat. Clim. Chang., 6(8), 751–758, doi:10.1038/nclimate3071, 2016.

Bradford, M. A., Ciska, G. F., Bonis, A., Bradford, E. M., Classen, A. T., Cornelissen, J. H. C., Crowther, T. W., De Long, J. R., Freschet, G. T., Kardol, P., Manrubia-Freixa, M., Maynard, D. S., Newman, G. S., Logtestijn, R. S. P., Viketoft, M., Wardle, D. A., Wieder, W. R., Wood, S. A. and Van Der Putten, W. H.: A test of the hierarchical model of litter decomposition, Nat. Ecol. Evol., 1(12), 1836–1845, doi:10.1038/s41559-017-0367-4, 2017.

Brovkin, V., Van Bodegom, P. M., Kleinen, T., Wirth, C., Cornwell, W. K., Cornelissen, J. H. C. and Kattge, J.: Plant-driven variation in decomposition rates improves projections of global litter stock distribution, Biogeosciences, 9(1), 565–576, doi:10.5194/bg-9-565-2012, 2012.

Carey, C. C., Farrell, K. J., Hounshell, A. G. and O'Connell, K.: Macrosystems EDDIE teaching modules significantly increase ecology students' proficiency and confidence working with ecosystem models and use of systems thinking, Ecol. Evol., 10(22), 12515–12527, doi:10.1002/ece3.6757, 2020.

Charney, J., Stone, P. H. and Quirk, W. J.: Drought in the Sahara: A biogeophysical feedback mechanism, Science (80-.)., 187(4175), 434–435, doi:10.1126/science.187.4175.434, 1975.

Cheng, S. J.: Modeling global change ecology in a high–carbon dioxide world, [online] Available from: https://eos.org/meeting-reports/modeling-global-change-ecology-in-a-high-carbon-dioxide-world, 2018.

Cheng, S. J., Hess, P. G., Wieder, W. R., Thomas, R. Q., Nadelhoffer, K. J., Vira, J., Lombardozzi, D. L., Gundersen, P., Fernandez, I. J., Schleppi, P., Gruselle, M.-C., Moldan, F. and Goodale, C. L.: Decadal impacts of nitrogen additions on temperate forest carbon sinks: A data-model comparison, Biogeosciences Discuss., doi:10.5194/bg-2018-505, 2019.

Cheng, W. and Kuzyakov, Y.: Root effects on soil organic matter decomposition, in Roots and soil management: interactions between roots and the soil. Agronomy monograph No. 48., 2005.

Clark, M. P., Slater, A. G., Rupp, D. E., Woods, R. A., Vrugt, J. A., Gupta, H. V., Wagener, T. and Hay, L. E.: Framework for Understanding Structural Errors (FUSE): A modular framework to diagnose differences between hydrological models, Water Resour. Res., 44(12), 1–14, doi:10.1029/2007wr006735, 2008.

Clark, M. P., Kavetski, D. and Fenicia, F.: Pursuing the method of multiple working hypotheses for hydrological modeling, Water Resour. Res., 47(9), 1–16, doi:10.1029/2010WR009827, 2011.

Clark, M. P., Nijssen, B., Lundquist, J. D., Kavetski, D., Rupp, D. E., Woods, R. A., Freer, J. E., Gutmann, E. D., Wood, A. W., Brekke, L. D., Arnold, J. R., Gochis, D. J. and Rasmussen, R. M.: A unified approach for process-based hydrologic modeling: 1. Modeling concept, Water Resour. Res., 51(4), 2498–2514, doi:10.1002/2015WR017198, 2015.

Cleveland, C. C. and Liptzin, D.: C:N:P stoichiometry in soil: is there a “Redfield ratio” for the microbial biomass?, Biogeochemistry, 85(3), 235–252, doi:10.1007/s10533-007-9132-0, 2007.

Collier, N., Hoffman, F. M., Lawrence, D. M., Keppel-Aleks, G., Koven, C. D., Riley, W. J., Mu, M. and Randerson, J. T.: The International Land Model Benchmarking (ILAMB) System: Design, Theory, and Implementation, J. Adv. Model. Earth Syst., 10(11), 2731–2754, doi:10.1029/2018MS001354, 2018.

Colman, B. P. and Schimel, J. P.: Drivers of microbial respiration and net N mineralization at the continental scale, Soil Biol. Biochem., 60, 65–76, doi:10.1016/j.soilbio.2013.01.003, 2013.

Conant, R. T., Steinweg, J. M., Haddix, M. L., Paul, E. A., Plante, A. F. and six: Experimental warming shows that decomposition temperature sensitivity increase with soil organic matter recalcitrance, Ecology, 89(9), 2384–2391 [online] Available from: http://www.esajournals.org/doi/pdf/10.1890/08-0137.1, 2008.

Cooley, E.: Training an Interdisciplinary Team in Communication and Decision-Making Skills, Small Gr. Res., 25(1), 5–25, doi:10.1177/1046496494251002, 1994.

Cotrufo, M. F., Wallenstein, M. D., Boot, C. M., Denef, K. and Paul, E.: The Microbial Efficiency-Matrix Stabilization (MEMS) framework integrates plant litter decomposition with soil organic matter stabilization: Do labile plant inputs form stable soil organic matter?, Glob. Chang. Biol., 19(4), 988–995, doi:10.1111/gcb.12113, 2013.

Cox, P. M., Betts, R. A., Jones, C. D., Spall, S. A. and Tollerdell, I. J.: Accelaration of global warming due to carbon-cycle feedback in a coupled climate model, Nature, 408, doi:10.1038/35047138, 2000.

Craine, J., Morrow, C. and Fierer, N.: Microbial nitrogen limitation increases decomposition, Ecology, 88(8), 2105–2113, 2007.

Dagon, K., Sanderson, B. M., Fisher, R. A. and Lawrence, D. M.: A machine learning approach to emulation and biophysical parameter estimation with the Community Land Model, version 5, Adv. Stat. Climatol. Meteorol. Oceanogr., 6(2), 223–244, doi:10.5194/ascmo-6-223-2020, 2020.

Danabasoglu, G., Lamarque, J. F., Bacmeister, J., Bailey, D. A., DuVivier, A. K., Edwards, J., Emmons, L. K., Fasullo, J., Garcia, R., Gettelman, A., Hannay, C., Holland, M. M., Large, W. G., Lauritzen, P. H., Lawrence, D. M., Lenaerts, J. T. M., Lindsay, K., Lipscomb, W. H., Mills, M. J., Neale, R., Oleson, K. W., Otto-Bliesner, B., Phillips, A. S., Sacks, W., Tilmes, S., van Kampenhout, L., Vertenstein, M., Bertini, A., Dennis, J., Deser, C., Fischer, C., Fox-Kemper, B., Kay, J. E., Kinnison, D., Kushner, P. J., Larson, V. E., Long, M. C., Mickelson, S., Moore, J. K., Nienhouse, E., Polvani, L., Rasch, P. J. and Strand, W. G.: The Community Earth System Model Version 2 (CESM2), J. Adv. Model. Earth Syst., 12(2), 1–35, doi:10.1029/2019MS001916, 2020.

Davidson, E. A. and Janssens, I. A.: Temperature sensitivity of soil carbon decomposition and feedbacks to climate change., Nature, 440(March), 165–173, doi:10.1038/nature04514, 2006.

Delgado-Baquerizo, M., Oliverio, A. M., Brewer, T. E., Benavent-gonzález, A., Eldridge, D. J., Bardgett, R. D., Maestre, F. T., Singh, B. K. and Fierer, N.: Bacteria Found in Soil, Science (80-.)., 325(February), 320–325, 2018.

Denning, A. S., Randall, D. A., Collatz, G. J. and Sellers, P. J.: Simulations of terrestrial carbon metabolism and atmospheric CO2 in a general circulation model. Part 2: Simulated CO2 concentrations, Tellus, Ser. B Chem. Phys. Meteorol., 48(4), 543–567, doi:10.3402/tellusb.v48i4.15931, 1996.

Deutsch, C. A., Tewksbury, J. J., Tigchelaar, M., Battisti, D. S., Merrill, S. C., Huey, R. B. and Naylor, R. L.: Increase in crop losses to insect pests in a warming climate, Science (80-.)., 361(6405), 916–919, doi:10.1126/science.aat3466, 2018.

Dickinson, R. E.: Modeling evapotranspiration for three-dimensional global climate models, Clim. Process. Clim. Sensit., 29, 58–72, 1984.

Dickinson, R. E. and Henderson-Sellers, A.: Modelling tropical deforestation: A study of GCM land-surface parametrizations, Q. J. R. Meteorol. Soc., 114(480), 439–462, doi:10.1002/qj.49711448009, 1988.

Dickinson, R. E., Jager, J., Washington, W. M. and Wolski, R.: Boundary Subroutine for the NCAR Global Climate Model, NCAR Tech. Note NCAR/TN-173+IA, 1981.

Dickinson, R. E., Henderson-Sellers, A., Kennedy, P. J. and Wilson, M. F.: Biosphere-

Atmosphere Transfer Scheme (BATS) for the NCAR Community Climate Model, NCAR Tech. Note NCAR/TN-275+STR, 1986.

Dietze, M. C.: Prediction in ecology: A first-principles framework, Ecol. Appl., 27(7), 2048–2060, doi:10.1002/eap.1589, 2017.

Dignac, M.-F., Derrien, D., Barré, P., Barot, S., Cécillon, L., Chenu, C., Chevallier, T., Freschet, G. T., Garnier, P., Guenet, B., Hedde, M., Klumpp, K., Lashermes, G., Maron, P.-A., Nunan, N., Roumet, C. and Basile-Doelsch, I.: Increasing soil carbon storage: mechanisms, effects of agricultural practices and proxies. A review, Agron. Sustain. Dev., 37(2), 14, doi:10.1007/s13593-017-0421-2, 2017.

Dijkstra, F. A., Carrillo, Y., Pendall, E. and Morgan, J. A.: Rhizosphere priming: A nutrient perspective, Front. Microbiol., 4(JUL), doi:10.3389/fmicb.2013.00216, 2013.

Dijkstra, P., Thomas, S. C., Heinrich, P. L., Koch, G. W., Schwartz, E. and Hungate, B. A.: Effect of temperature on metabolic activity of intact microbial communities: Evidence for altered metabolic pathway activity but not for increased maintenance respiration and reduced carbon use efficiency, Soil Biol. Biochem., 43(10), 2023–2031, doi:10.1016/j.soilbio.2011.05.018, 2011.

Dirmeyer, P. A., Gao, X., Zhao, M., Guo, Z., Oki, T. and Hanasaki, N.: GSWP-2: Multimodel analysis and implications for our perception of the land surface, Bull. Am. Meteorol. Soc., 87(10), 1381–1397, doi:10.1175/BAMS-87-10-1381, 2006.

Doetterl, S., Stevens, A., Six, J., Merckx, R., Van Oost, K., Casanova Pinto, M., Casanova-Katny, A., Muñoz, C., Boudin, M., Zagal Venegas, E. and Boeckx, P.: Soil carbon storage controlled by interactions between geochemistry and climate, Nat. Geosci., 8(10), 780–783, doi:10.1038/ngeo2516, 2015.

Duffy, M. A., García-Robledo, C., Gordon, S. P., Grant, N. A., Green, D. A., Kamath, A., Penczykowski, R. M., Rebolleda-Gómez, M., Wale, N. and Zaman, L.: Model systems in ecology, evolution, and behavior: A call for diversity in our model systems and discipline, Am. Nat., 198(1), 53–68, doi:10.1086/714574, 2021.

Dufresne, J. L., Foujols, M. A., Denvil, S., Caubel, A., Marti, O., Aumont, O., Balkanski, Y., Bekki, S., Bellenger, H., Benshila, R., Bony, S., Bopp, L., Braconnot, P., Brockmann, P., Cadule, P., Cheruy, F., Codron, F., Cozic, A., Cugnet, D., de Noblet, N., Duvel, J. P., Ethé, C., Fairhead, L., Fichefet, T., Flavoni, S., Friedlingstein, P., Grandpeix, J. Y., Guez, L., Guilyardi, E., Hauglustaine, D., Hourdin, F., Idelkadi, A., Ghattas, J., Joussaume, S., Kageyama, M., Krinner, G., Labetoulle, S., Lahellec, A., Lefebvre, M. P., Lefevre, F., Levy, C., Li, Z. X., Lloyd, J., Lott, F., Madec, G., Mancip, M., Marchand, M., Masson, S., Meurdesoif, Y., Mignot, J., Musat, I., Parouty, S., Polcher, J., Rio, C., Schulz, M., Swingedouw, D., Szopa, S., Talandier, C., Terray, P., Viovy, N. and Vuichard, N.: Climate change projections using the IPSL-CM5 Earth System Model: From CMIP3 to CMIP5., 2013.

Dungait, J. A. J., Hopkins, D. W., Gregory, A. S. and Whitmore, A. P.: Soil organic matter turnover is governed by accessibility not recalcitrance, Glob. Chang. Biol., 18(6), 1781–1796, doi:10.1111/j.1365-2486.2012.02665.x, 2012.

Dunne, J. P., Horowitz, L. W., Adcroft, A. J., Ginoux, P., Held, I. M., John, J. G., Krasting, J. P., Malyshev, S., Naik, V., Paulot, F., Shevliakova, E., Stock, C. A., Zadeh, N., Balaji, V., Blanton, C., Dunne, K. A., Dupuis, C., Durachta, J., Dussin, R., Gauthier, P. P. G., Griffies, S. M., Guo, H., Hallberg, R. W., Harrison, M., He, J., Hurlin, W., McHugh, C., Menzel, R., Milly, P. C. D., Nikonov, S., Paynter, D. J., Ploshay, J., Radhakrishnan, A., Rand, K., Reichl, B. G., Robinson, T., Schwarzkopf, D. M., Sentman, L. T., Underwood, S., Vahlenkamp, H., Winton, M., Wittenberg, A. T., Wyman, B., Zeng, Y. and Zhao, M.: The GFDL Earth System Model Version 4.1 (GFDL-ESM 4.1): Overall Coupled Model Description and Simulation Characteristics, J. Adv. Model. Earth Syst., 12(11), 1–56, doi:10.1029/2019MS002015, 2020.

Duursma, R. A.: Plantecophys - An R package for analysing and modelling leaf gas exchange data, PLoS One, 10(11), 1–13, doi:10.1371/journal.pone.0143346, 2015.

Dybzinski, R., Farrior, C. E. and Pacala, S. W.: Increased forest carbon storage with increased atmospheric CO2 despite nitrogen limitation: A game-theoretic allocation model for trees in competition for nitrogen and light, Glob. Chang. Biol., 21(3), 1182–1196, doi:10.1111/gcb.12783, 2015.

Edburg, S. L., Hicke, J. A., Lawrence, D. M. and Thornton, P. E.: Simulating coupled carbon and nitrogen dynamics following mountain pine beetle outbreaks in the western United States, J. Geophys. Res. Biogeosciences, 116(4), 1–15, doi:10.1029/2011JG001786, 2011.

Edwards, I. P. and Zak, D. R.: Fungal community composition and function after long-term exposure of northern forests to elevated atmospheric CO2 and tropospheric O3, Glob. Chang. Biol., 17(6), 2184–2195, doi:10.1111/j.1365-2486.2010.02376.x, 2011.

Edwards, P. N.: History of climate modeling, Wiley Interdiscip. Rev. Clim. Chang., 2(1), 128–139, doi:10.1002/wcc.95, 2011.

Emery, N. C., Bledsoe, E. K., Hasley, A. O. and Eaton, C. D.: Cultivating inclusive instructional and research environments in ecology and evolutionary science, Ecol. Evol., 11(4), 1480–1491, doi:10.1002/ece3.7062, 2021.

Evans, S. E. and Wallenstein, M. D.: Soil microbial community response to drying and rewetting stress: Does historical precipitation regime matter?, Biogeochemistry, 109(1–3), 101–116, doi:10.1007/s10533-011-9638-3, 2012.

Farrior, C. E., Dybzinski, R., Levin, S. A. and Pacala, S. W.: Competition for water and light in closed-canopy forests: A tractable model of carbon allocation with implications for carbon sinks, Am. Nat., 181(3), 314–330, doi:10.1086/669153, 2013.

Fatichi, S., Manzoni, S., Or, D. and Paschalis, A.: A Mechanistic Model of Microbially Mediated Soil Biogeochemical Processes: A Reality Check, Global Biogeochem. Cycles, 33(6), 620–648, doi:10.1029/2018GB006077, 2019.

Fer, I., Gardella, A. K., Shiklomanov, A. N., Campbell, E. E., Cowdery, E. M., De Kauwe, M. G., Desai, A., Duveneck, M. J., Fisher, J. B., Haynes, K. D., Hoffman, F. M., Johnston, M. R., Kooper, R., LeBauer, D. S., Mantooth, J., Parton, W. J., Poulter, B., Quaife, T., Raiho, A., Schaefer, K., Serbin, S. P., Simkins, J., Wilcox, K. R., Viskari, T. and Dietze, M. C.: Beyond ecosystem modeling: A roadmap to community cyberinfrastructure for ecological data-model integration, Glob. Chang. Biol., 27(1), 13–26, doi:10.1111/gcb.15409, 2021.

Field, A. P. and Gillett, R.: How to do a meta-analysis, Br. J. Math. Stat. Psychol., 63(3), 665–694, doi:10.1348/000711010X502733, 2010.

Fierer, N. and Jackson, R. B.: The diversity and biogeography of soil bacterial communities., Proc. Natl. Acad. Sci., 103(3), 626–631, doi:10.1073/pnas.0507535103, 2006.

Fierer, N., Bradford, M. a and Jackson, R. B.: Toward an ecological classification of soil bacteria., Ecology, 88(6), 1354–64 [online] Available from: http://www.ncbi.nlm.nih.gov/pubmed/17601128, 2007.

Fierer, N., Wood, S. A. and Bueno de Mesquita, C. P.: How microbes can, and cannot, be used to assess soil health, Soil Biol. Biochem., 153(December 2020), 108111, doi:10.1016/j.soilbio.2020.108111, 2021.

Fisher, R. A. and Koven, C. D.: Perspectives on the Future of Land Surface Models and the Challenges of Representing Complex Terrestrial Systems, J. Adv. Model. Earth Syst., 12(4), doi:10.1029/2018MS001453, 2020.

Foley, J. A., Prentice, I. C., Ramankutty, N., Levis, S., Pollard, D., Sitch, S. and Haxeltine, A.: An integrated biosphere model of land surface processes, terrestrial carbon balance, and vegetation dynamics, Global Biogeochem. Cycles, 10(4), 603–628, doi:10.1029/96GB02692, 1996.

Ford, H., Brick, C., Azmitia, M., Blaufuss, K. and Dekens, P.: Women from some minorities get too few talks, Nature, 576(December), 1–4, 2019.

Franklin, O., Harrison, S. P., Dewar, R., Farrior, C. E., Brännström, Å., Dieckmann, U., Pietsch, S., Falster, D., Cramer, W., Loreau, M., Wang, H., Mäkelä, A., Rebel, K. T., Meron, E., Schymanski, S. J., Rovenskaya, E., Stocker, B. D., Zaehle, S., Manzoni, S., van Oijen, M., Wright, I. J., Ciais, P., van Bodegom, P. M., Peñuelas, J., Hofhansl, F., Terrer, C., Soudzilovskaia, N. A., Midgley, G. and Prentice, I. C.: Organizing principles for vegetation dynamics, Nat. Plants, 6(5), 444–453, doi:10.1038/s41477-020-0655-x, 2020.

Franks, P. J., Bonan, G. B., Berry, J. A., Lombardozzi, D. L., Holbrook, N. M., Herold, N. and Oleson, K. W.: Comparing optimal and empirical stomatal conductance models for application in Earth system models, Glob. Chang. Biol., 24(12), 5708–5723, doi:10.1111/gcb.14445, 2018.

Frey, S. D., Lee, J., Melillo, J. M. and Six, J.: The temperature response of soil microbial efficiency and its feedback to climate, Nat. Clim. Chang., 3(4), 395–398, doi:10.1038/nclimate1796, 2013.

Friedlingstein, P., Jones, M. W., O'Sullivan, M., Andrew, R. M., Hauck, J., Peters, G. P., Peters, W., Pongratz, J., Sitch, S., Le Quéré, C., Bakker, D. C. E., Canadell, J. G., Ciais, P., Jackson, R. B., Anthoni, P., Barbero, L., Bastos, A., Bastrikov, V., Becker, M., Bopp, L., Buitenhuis, E., Chandra, N., Chevallier, F., Chini, L. P., Currie, K. I., Feely, R. A., Gehlen, M., Gilfillan, D., Gkritzalis, T., Goll, D. S., Gruber, N., Gutekunst, S., Harris, I., Haverd, V., Houghton, R. A., Hurtt, G., Ilyina, T., Jain, A. K., Joetzjer, E., Kaplan, J. O., Kato, E., Klein Goldewijk, K., Korsbakken, J. I., Landschützer, P., Lauvset, S. K., Lefèvre, N., Lenton, A., Lienert, S., Lombardozzi, D., Marland, G., McGuire, P. C., Melton, J. R., Metzl, N., Munro, D. R., Nabel, J. E. M. S., Nakaoka, S.-I., Neill, C., Omar, A. M., Ono, T., Peregon, A., Pierrot, D., Poulter, B., Rehder, G., Resplandy, L., Robertson, E., Rödenbeck, C., Séférian, R., Schwinger, J., Smith, N., Tans, P. P., Tian, H., Tilbrook, B., Tubiello, F. N., van der Werf, G. R., Wiltshire, A. J. and Zaehle, S.: Global Carbon Budget 2019, Earth Syst. Sci. Data, 1, 1783–1838, 2019.

Friend, A. D.: Terrestrial plant production and climate change, J. Exp. Bot., 61(5), 1293–1309, doi:10.1093/jxb/erq019, 2010.

Fung, I. Y., Doney, S. C., Lindsay, K. and John, J.: Evolution of carbon sinks in a changing climate, Proc. Natl. Acad. Sci. U. S. A., 102(32), 11201–11206, doi:10.1073/pnas.0504949102, 2005.

Gan, H., Zak, D. R. and Hunter, M. D.: Chronic nitrogen deposition alters the structure and function of detrital food webs in a northern hardwood ecosystem, Ecol. Appl., 23(6), 1311–1321, doi:10.1890/12-1895.1, 2013.

Georgiou, K., Abramoff, R. Z., Harte, J., Riley, W. J. and Torn, M. S.: Microbial community-level regulation explains soil carbon responses to long-term litter manipulations, Nat. Commun., 8(1), 1223, doi:10.1038/s41467-017-01116-z, 2017.

German, D. P., Marcelo, K. R. B., Stone, M. M. and Allison, S. D.: The Michaelis-Menten kinetics of soil extracellular enzymes in response to temperature: a cross-latitudinal study, Glob. Chang. Biol., 18(4), 1468–1479, doi:10.1111/j.1365-2486.2011.02615.x, 2012.

Geyer, K. M., Kyker-Snowman, E., Grandy, A. S. and Frey, S. D.: Microbial carbon use efficiency: accounting for population, community, and ecosystem-scale controls over the fate of metabolized organic matter, Biogeochemistry, 127(2–3), 173–188, doi:10.1007/s10533-016-0191-y, 2016.

Giles, S., Jackson, C. and Stephen, N.: Barriers to fieldwork in undergraduate geoscience degrees, Nat. Rev. Earth Environ., 1(2), 77–78, doi:10.1038/s43017-020-0022-5, 2020.

Gleixner, G.: Soil organic matter dynamics: A biological perspective derived from the use of compound-specific isotopes studies, Ecol. Res., 28(5), 683–695, doi:10.1007/s11284-012-1022-9, 2013.

Golaz, J. C., Caldwell, P. M., Van Roekel, L. P., Petersen, M. R., Tang, Q., Wolfe, J. D., Abeshu, G., Anantharaj, V., Asay-Davis, X. S., Bader, D. C., Baldwin, S. A., Bisht, G., Bogenschutz, P. A., Branstetter, M., Brunke, M. A., Brus, S. R., Burrows, S. M., Cameron-Smith, P. J., Donahue, A. S., Deakin, M., Easter, R. C., Evans, K. J., Feng, Y., Flanner, M., Foucar, J. G., Fyke, J. G., Griffin, B. M., Hannay, C., Harrop, B. E., Hoffman, M. J., Hunke, E. C., Jacob, R. L., Jacobsen, D. W., Jeffery, N., Jones, P. W., Keen, N. D., Klein, S. A., Larson, V. E., Leung, L. R., Li, H. Y., Lin, W., Lipscomb, W. H., Ma, P. L., Mahajan, S., Maltrud, M. E., Mametjanov, A., McClean, J. L., McCoy, R. B., Neale, R. B., Price, S. F., Qian, Y., Rasch, P. J., Reeves Eyre, J. E. J., Riley, W. J., Ringler, T. D., Roberts, A. F., Roesler, E. L., Salinger, A. G., Shaheen, Z., Shi, X., Singh, B., Tang, J., Taylor, M. A., Thornton, P. E., Turner, A. K., Veneziani, M., Wan, H., Wang, H., Wang, S., Williams, D. N., Wolfram, P. J., Worley, P. H., Xie, S., Yang, Y., Yoon, J. H., Zelinka, M. D., Zender, C. S., Zeng, X., Zhang, C., Zhang, K., Zhang, Y., Zheng, X., Zhou, T. and Zhu, Q.: The DOE E3SM Coupled Model Version 1: Overview and Evaluation at Standard Resolution, J. Adv. Model. Earth Syst., 11(7), 2089–2129, doi:10.1029/2018MS001603, 2019.

Graham, E. B., Knelman, J. E., Schindlbacher, A., Siciliano, S., Breulmann, M., Yannarell, A., Beman, J. M., Abell, G., Philippot, L., Prosser, J., Foulquier, A., Yuste, J. C., Glanville, H. C., Jones, D. L., Angel, R., Salminen, J., Newton, R. J., Bürgmann, H., Ingram, L. J., Hamer, U., Siljanen, H. M. P., Peltoniemi, K., Potthast, K., Bañeras, L., Hartmann, M., Banerjee, S., Yu, R. Q., Nogaro, G., Richter, A., Koranda, M., Castle, S. C., Goberna, M., Song, B., Chatterjee, A., Nunes, O. C., Lopes, A. R., Cao, Y., Kaisermann, A., Hallin, S., Strickland, M. S., Garcia-Pausas, J., Barba, J., Kang, H., Isobe, K., Papaspyrou, S., Pastorelli, R., Lagomarsino, A., Lindström, E. S., Basiliko, N. and Nemergut, D. R.: Microbes as Engines of Ecosystem Function: When Does Community Structure Enhance Predictions of Ecosystem Processes?, Front. Microbiol., 7(February), 1–10, doi:10.3389/fmicb.2016.00214, 2016.

Grandy, A. S. and Neff, J. C.: Molecular C dynamics downstream: The biochemical decomposition sequence and its impact on soil organic matter structure and function, Sci. Total Environ., 404(2–3), 297–307, doi:10.1016/j.scitotenv.2007.11.013, 2008.

Grandy, A. S. and Robertson, G. P.: Land-use intensity effects on soil organic carbon accumulation rates and mechanisms, Ecosystems, 10(1), 58–73, doi:10.1007/s10021-006-9010-y, 2007.

Grömping, U.: Relative importance for linear regression in R: The package relaimpo, J. Stat. Softw., 17(1), 1–27, doi:10.18637/jss.v017.i01, 2006.

Gruber, N. and Galloway, J. N.: An Earth-system perspective of the global nitrogen cycle, Nature, 451(7176), 293–296, doi:10.1038/nature06592, 2008.

Guenet, B., Moyano, F. E., Peylin, P., Ciais, P. and Janssens, I. A.: Towards a representation of priming on soil carbon decomposition in the global land biosphere model ORCHIDEE (version 1.9.5.2), Geosci. Model Dev., 9(2), 841–855, doi:10.5194/gmd-9-841-2016, 2016.

Haddix, M. L., Paul, E. A. and Cotrufo, M. F.: Dual, differential isotope labeling shows the preferential movement of labile plant constituents into mineral-bonded soil organic matter, Glob. Chang. Biol., 22(6), 2301–2312, doi:10.1111/gcb.13237, 2016.

Haddix, M. L., Gregorich, E. G., Helgason, B. L., Janzen, H., Ellert, B. H. and Francesca Cotrufo, M.: Climate, carbon content, and soil texture control the independent formation and persistence of particulate and mineral-associated organic matter in soil, Geoderma, 363(December 2019), 114160, doi:10.1016/j.geoderma.2019.114160, 2020.

Hampton, S. E., Jones, M. B., Wasser, L. A., Schildhauer, M. P., Supp, S. R., Brun, J., Hernandez, R. R., Boettiger, C., Collins, S. L., Gross, L. J., Fernández, D. S., Budden, A., White, E. P., Teal, T. K., Labou, S. G. and Aukema, J. E.: Skills and Knowledge for Data-Intensive Environmental Research, Bioscience, 67(6), 546–557, doi:10.1093/biosci/bix025, 2017.

Hararuk, O., Smith, M. J. and Luo, Y.: Microbial models with data-driven parameters predict stronger soil carbon responses to climate change, Glob. Chang. Biol., 2439–2453, doi:10.1111/gcb.12827, 2015.

Harmon, M. E., Silver, W. L., Fasth, B., Chen, H., Burke, I. C., Parton, W. J., Hart, S. C., Currie, W. S. and LIDET: Long-term patterns of mass loss during the decomposition of leaf and fine root litter: an intersite comparison, Glob. Chang. Biol., 15, 1320–1338, doi:10.1111/j.1365-2486.2008.01837.x, 2009.

Hassink, J., Bouwman, L. A., Zwart, K. B., Bloem, J. and Brussaard, L.: Relationships between soil texture, physical protection of organic matter, soil biota, and c and n mineralization in grassland soils, Geoderma, 57(1–2), 105–128, doi:10.1016/0016-7061(93)90150-J, 1993.

Heckman, K., Grandy, A. S., Gao, X., Keiluweit, M., Wickings, K., Carpenter, K., Chorover, J. and Rasmussen, C.: Sorptive fractionation of organic matter and formation of organo-hydroxy-aluminum complexes during litter biodegradation in the presence of gibbsite, Geochim. Cosmochim. Acta, 121, 667–683, doi:10.1016/j.gca.2013.07.043, 2013.

Hobbie, S. E., Eddy, W. C., Buyarski, C. R., Carol Adair, E., Ogdahl, M. L. and Weisenhorn, P.: Response of decomposing litter and its microbial community to multiple forms of nitrogen enrichment, Ecol. Monogr., 82(3), 389–405, doi:10.1890/11-1600.1, 2012.

Homann, P. S., Kapchinske, J. S. and Boyce, A.: Relations of mineral-soil C and N to climate and texture: Regional differences within the conterminous USA, Biogeochemistry, 85(3), 303–316, doi:10.1007/s10533-007-9139-6, 2007.

Huang, Y., Guenet, B., Wang, Y. L. and Ciais, P.: Global simulation and evaluation of soil organic matter and microbial carbon and nitrogen stocks using the microbial decomposition model ORCHIMIC v2.0, Global Biogeochem. Cycles [online] Available from: https://doi.org/10.1029/2020gb006836, 2021.

Huang, Y. Y., Guenet, B., Ciais, P., Janssens, I. A., Soong, J. L., Wang, Y., Goll, D., Blagodatskaya, E., Huang, Y. Y., Dev, M., Huang, Y. Y., Guenet, B., Ciais, P., Janssens, I. A., Soong, J. L., Wang, Y., Goll, D., Blagodatskaya, E. and Huang, Y. Y.: ORCHIMIC (v1.0), a microbe-driven model for soil organic matter decomposition designed for large-scale applications, Geosci. Model Dev., 11, 2111–2138, doi:10.5194/gmd-11-2111-2018, 2018a.

Huang, Y. Y. Y., Guenet, B., Ciais, P., Janssens, I. A., Soong, J. L., Wang, Y., Goll, D. and Blagodatskaya, E.: ORCHIMIC (v1.0), a microbe-driven model for soil organic matter decomposition designed for large-scale applications, Geosci. Model Dev., 11(February), doi:10.5194/gmd-11-2111-2018, 2018b.

Hunt, S. K., Galatowitsch, M. L. and McIntosh, A. R.: Interactive effects of land use, temperature, and predators determine native and invasive mosquito distributions, Freshw. Biol., 62(9), 1564–1577, doi:10.1111/fwb.12967, 2017.

Ito, A., Hajima, T., Lawrence, D. M., Brovkin, V., Delire, C., Guenet, B., Jones, C. D., Malyshev, S., Materia, S., McDermid, S. P., Peano, D., Pongratz, J., Robertson, E., Shevliakova, E., Vuichard, N., Wårlind, D., Wiltshire, A. and Ziehn, T.: Soil carbon sequestration simulated in CMIP6-LUMIP models: Implications for climatic mitigation, Environ. Res. Lett., 15(12), doi:10.1088/1748-9326/abc912, 2019.

Jackson, R. B., Lajtha, K., Crow, S. E., Hugelius, G., Kramer, M. G. and Piñeiro, G.: The Ecology of Soil Carbon: Pools, Vulnerabilities, and Biotic and Abiotic Controls, Annu. Rev. Ecol. Evol. Syst., doi:10.1146/annurev-ecolsys-112414-054234, 2017.

Jenkinson, D. S. and Rayner, J. H.: The turnover of soil organic matter in some of the rothamsted classical experiments, Soil Sci., doi:10.1097/00010694-197705000-00005, 1977.

Jenny, H.: Factors of Soil Formation, Soil Sci., doi:10.1097/00010694-194111000-00009, 1941.

Jiang, C., Ryu, Y., Wang, H. and Keenan, T. F.: An optimality-based model explains seasonal variation in C3 plant photosynthetic capacity, Glob. Chang. Biol., 26(11), 6493–6510, doi:10.1111/gcb.15276, 2020.

Jilling, A., Keiluweit, M., Contosta, A. R., Frey, S., Schimel, J., Schnecker, J., Smith, R. G., Tiemann, L. and Grandy, A. S.: Minerals in the rhizosphere: overlooked mediators of soil nitrogen availability to plants and microbes, Biogeochemistry, doi:10.1007/s10533-018-0459-5, 2018.

Jobbágy, E. G. and Jackson, R. B.: The vertical distribution of soil organic carbon and its relation to climate and vegetation, Ecol. Appl., 10(2), 423–436, doi:10.1890/1051-0761(2000)010[0423:TVDOSO]2.0.CO;2, 2000.

Jung, M., Schwalm, C., Migliavacca, M., Walther, S., Camps-Valls, G., Koirala, S., Anthoni, P., Besnard, S., Bodesheim, P., Carvalhais, N., Chevallier, F., Gans, F., S Goll, D., Haverd, V., Köhler, P., Ichii, K., K Jain, A., Liu, J., Lombardozzi, D., E M S Nabel, J., A Nelson, J., O'Sullivan, M., Pallandt, M., Papale, D., Peters, W., Pongratz, J., Rödenbeck, C., Sitch, S., Tramontana, G., Walker, A., Weber, U. and Reichstein, M.: Scaling carbon fluxes from eddy covariance sites to globe: Synthesis and evaluation of the FLUXCOM approach, Biogeosciences, 17(5), 1343–1365, doi:10.5194/bg-17-1343-2020, 2020.

Kaiser, C., Franklin, O., Dieckmann, U. and Richter, A.: Microbial community dynamics alleviate stoichiometric constraints during litter decay., Ecol. Lett., 17(6), 680–90, doi:10.1111/ele.12269, 2014.

Kaiser, C., Franklin, O., Richter, A. and Dieckmann, U.: Social dynamics within decomposer communities lead to nitrogen retention and organic matter build-up in soils, Nat. Commun., 6, 1–10, doi:10.1038/ncomms9960, 2015.

Kallenbach, C. M., Grandy, A. and Frey, S. D.: Direct evidence for microbial-derived soil organic matter formation and its ecophysiological controls, Nat. Commun., in revisio, 1–10, doi:10.1038/ncomms13630, 2016.

Kallenbach, C. M. M., Grandy, A. S. S., Frey, S. D. D. and Diefendorf, A. F. F.: Microbial physiology and necromass regulate agricultural soil carbon accumulation, Soil Biol. Biochem., 91, 279–290, doi:10.1016/j.soilbio.2015.09.005, 2015.

Kattge, J. and Knorr, W.: Temperature acclimation in a biochemical model of photosynthesis: A reanalysis of data from 36 species, Plant, Cell Environ., 30(9), 1176–1190, doi:10.1111/j.1365-3040.2007.01690.x, 2007.

Kattge, J., Knorr, W., Raddatz, T. and Wirth, C.: Quantifying photosynthetic capacity and its relationship to leaf nitrogen content for global-scale terrestrial biosphere models, Glob. Chang. Biol., 15(4), 976–991, doi:10.1111/j.1365-2486.2008.01744.x, 2009.

Kattge, J., Díaz, S., Lavorel, S., Prentice, I. C., Leadley, P., Bönisch, G., Garnier, E., Westoby, M., Reich, P. B., Wright, I. J., Cornelissen, J. H. C., Violle, C., Harrison, S. P., Van Bodegom, P. M., Reichstein, M., Enquist, B. J., Soudzilovskaia, N. A., Ackerly, D. D., Anand, M., Atkin, O., Bahn, M., Baker, T. R., Baldocchi, D., Bekker, R., Blanco, C. C., Blonder, B., Bond, W. J., Bradstock, R., Bunker, D. E., Casanoves, F., Cavender-Bares, J., Chambers, J. Q., Chapin, F. S., Chave, J., Coomes, D., Cornwell, W. K., Craine, J. M., Dobrin, B. H., Duarte, L., Durka, W., Elser, J., Esser, G., Estiarte, M., Fagan, W. F., Fang, J., Fernández-Méndez, F., Fidelis, A., Finegan, B., Flores, O., Ford, H., Frank, D., Freschet, G. T., Fyllas, N. M., Gallagher, R. V., Green, W. A., Gutierrez, A. G., Hickler, T., Higgins, S. I., Hodgson, J. G., Jalili, A., Jansen, S., Joly, C. A., Kerkhoff, A. J., Kirkup, D., Kitajima, K., Kleyer, M., Klotz, S., Knops, J. M. H., Kramer, K., Kühn, I., Kurokawa, H., Laughlin, D., Lee, T. D., Leishman, M., Lens, F., Lenz, T., Lewis, S. L., Lloyd, J., Llusià, J., Louault, F., Ma, S., Mahecha, M. D., Manning, P., Massad, T., Medlyn, B. E., Messier, J., Moles, A. T., Müller, S. C., Nadrowski, K., Naeem, S., Niinemets, Ü., Nöllert, S., Nüske, A., Ogaya, R., Oleksyn, J., Onipchenko, V. G., Onoda, Y., Ordoñez, J., Overbeck, G., et al.: TRY - a global database of plant traits, Glob. Chang. Biol., 17(9), 2905–2935, doi:10.1111/j.1365-2486.2011.02451.x, 2011.

De Kauwe, M. G., Medlyn, B. E., Zaehle, S., Walker, A. P., Dietze, M. C., Hickler, T., Jain, A. K., Luo, Y., Parton, W. J., Prentice, I. C., Smith, B., Thornton, P. E., Wang, S., Wang, Y. P., Wårlind, D., Weng, E., Crous, K. Y., Ellsworth, D. S., Hanson, P. J., Seok Kim, H., Warren, J. M., Oren, R. and Norby, R. J.: Forest water use and water use efficiency at elevated CO2: A model-data intercomparison at two contrasting temperate forest FACE sites, Glob. Chang. Biol., 19(6), 1759–1779, doi:10.1111/gcb.12164, 2013.

De Kauwe, M. G., Medlyn, B. E., Zaehle, S., Walker, A. P., Dietze, M. C., Wang, Y. P., Luo, Y., Jain, A. K., El-Masri, B., Hickler, T., Wårlind, D., Weng, E., Parton, W. J., Thornton, P. E., Wang, S., Prentice, I. C., Asao, S., Smith, B., Mccarthy, H. R., Iversen, C. M., Hanson, P. J., Warren, J. M., Oren, R. and Norby, R. J.: 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 Phytol., 203(3), 883–899, doi:10.1111/nph.12847, 2014.

Kennedy, D., Swenson, S., Oleson, K. W., Lawrence, D. M., Fisher, R., Lola da Costa, A. C. and Gentine, P.: Implementing Plant Hydraulics in the Community Land Model, Version 5, J. Adv.

Model. Earth Syst., 11(2), 485–513, doi:10.1029/2018MS001500, 2019.

Kleber, M., Eusterhues, K., Keiluweit, M., Mikutta, C., Mikutta, R. and Nico, P. S.: Mineral-Organic Associations: Formation, Properties, and Relevance in Soil Environments, Adv. Agron., 130, 1–140, doi:10.1016/bs.agron.2014.10.005, 2015.

Koven, C. D., Hugelius, G., Lawrence, D. M. and Wieder, W. R.: Higher climatological temperature sensitivity of soil carbon in cold than warm climates, Nat. Clim. Chang., 7(11), 817–822, doi:10.1038/nclimate3421, 2017.

Kramer, M. G., Lajtha, K. and Aufdenkampe, A. K.: Depth trends of soil organic matter C:N and 15N natural abundance controlled by association with minerals, Biogeochemistry, 136(3), 237–248, doi:10.1007/s10533-017-0378-x, 2017.

Kuziemsky, C. E., Borycki, E. M., Purkis, M. E., Black, F., Boyle, M., Cloutier-Fisher, D., Fox, L. A., MacKenzie, P., Syme, A., Tschanz, C., Wainwright, W. and Wong, H.: An interdisciplinary team communication framework and its application to healthcare “e-teams” systems design, BMC Med. Inform. Decis. Mak., 9(1), 1–15, doi:10.1186/1472-6947-9-43, 2009.

Kyker-Snowman, E., Wieder, W. R., Frey, S. D. and Grandy, A. S.: Stoichiometrically coupled carbon and nitrogen cycling in the MIcrobial-MIneral Carbon Stabilization model version 1.0 (MIMICS-CN v1.0), Geosci. Model Dev., 13(9), doi:10.5194/gmd-13-4413-2020, 2020.

Lade, S. J., Steffen, W., de Vries, W., Carpenter, S. R., Donges, J. F., Gerten, D., Hoff, H., Newbold, T., Richardson, K. and Rockström, J.: Human impacts on planetary boundaries amplified by Earth system interactions, Nat. Sustain., 3(2), 119–128., 2019.

Lajtha, K., Bowden, R. D., Crow, S., Fekete, I., Kotroczo, Z., Plante, A., Simpson, M. and Nadelhoffer, K.: The Detrital Input and Removal Treatment (DIRT) Network, Elsevier Inc., 2017.

Lal, R.: Soil carbon sequestration impacts on global climate change and food security, Science (80-.)., 304(5677), 1623–1627 [online] Available from: //000221934300036, 2004.

Lavallee, J. M., Soong, J. L. and Cotrufo, M. F.: Conceptualizing soil organic matter into particulate and mineral-associated forms to address global change in the 21st century, Glob. Chang. Biol., 26(1), 261–273, doi:10.1111/gcb.14859, 2020.

Law, B. E., Hudiburg, T. W., Berner, L. T., Kent, J. J., Buotte, P. C. and Harmon, M. E.: Land use strategies to mitigate climate change in carbon dense temperate forests, Proc. Natl. Acad. Sci. U. S. A., 115(14), 3663–3668, doi:10.1073/pnas.1720064115, 2018.

Lawrence, D. M., Fisher, R. A., Koven, C. D., Oleson, K. W., Swenson, S. C., Bonan, G., Collier, N., Ghimire, B., van Kampenhout, L., Kennedy, D., Kluzek, E., Lawrence, P. J., Li, F., Li, H., Lombardozzi, D., Riley, W. J., Sacks, W. J., Shi, M., Vertenstein, M., Wieder, W. R., Xu, C., Ali, A. A., Badger, A. M., Bisht, G., van den Broeke, M., Brunke, M. A., Burns, S. P., Buzan, J., Clark, M., Craig, A., Dahlin, K., Drewniak, B., Fisher, J. B., Flanner, M., Fox, A. M., Gentine, P., Hoffman, F., Keppel-Aleks, G., Knox, R., Kumar, S., Lenaerts, J., Leung, L. R., Lipscomb, W. H., Lu, Y., Pandey, A., Pelletier, J. D., Perket, J., Randerson, J. T., Ricciuto, D. M., Sanderson, B. M., Slater, A., Subin, Z. M., Tang, J., Thomas, R. Q., Val Martin, M. and Zeng, X.: The Community Land Model Version 5: Description of New Features, Benchmarking, and Impact of Forcing Uncertainty, J. Adv. Model. Earth Syst., 11(12), 4245–4287, doi:10.1029/2018MS001583, 2019.

LeBauer, D. and Treseder, K.: Nitrogen Limitation of Net Primary Productivity, Ecology, 89(2), 371–379, doi:10.1016/j.agee.2013.04.020, 2008.

Lehmann, J. and Kleber, M.: The contentious nature of soil organic matter., Nature, 528(7580), 60–8, doi:10.1038/nature16069, 2015.

Leuzinger, S. and Thomas, R. Q.: Meetings How do we improve Earth system models ? Integrating Earth system models , ecosystem models , New Phytol., 191, 15–18, 2011.

Li, D., Schädel, C., Haddix, M. L., Paul, E. A., Conant, R., Li, J., Zhou, J. and Luo, Y.: Differential responses of soil organic carbon fractions to warming: Results from an analysis with data assimilation, Soil Biol. Biochem., 67, 24–30, doi:10.1016/j.soilbio.2013.07.008, 2013.

Li, J., Wang, G., Allison, S. D., Mayes, M. a. and Luo, Y.: Soil carbon sensitivity to temperature and carbon use efficiency compared across microbial-ecosystem models of varying complexity, Biogeochemistry, 119(1–3), 67–84, doi:10.1007/s10533-013-9948-8, 2014.

Li, X. and Xiao, J.: A global, 0.05-degree product of solar-induced chlorophyll fluorescence derived from OCO-2, MODIS, and reanalysis data, Remote Sens., 11(5), doi:10.3390/rs11050517, 2019.

Liang, Z., Olesen, J. E., Jensen, J. L. and Elsgaard, L.: Nutrient availability affects carbon turnover and microbial physiology differently in topsoil and subsoil under a temperate grassland, Geoderma, 336(August 2018), 22–30, doi:10.1016/j.geoderma.2018.08.021, 2019.

Lombardozzi, D., Levis, S., Bonan, G. and Sparks, J. P.: Predicting photosynthesis and transpiration responses to ozone: Decoupling modeled photosynthesis and stomatal conductance, Biogeosciences, 9(8), 3113–3130, doi:10.5194/bg-9-3113-2012, 2012.

Lombardozzi, D., Sparks, J. P. and Bonan, G.: Integrating O3 influences on terrestrial processes: photosynthetic and stomatal response data available for regional and global modeling, Biogeosciences, 10, 6815–6831, doi:10.5194/bg-10-6815-2013, 2013.

Lombardozzi, D., Levis, S., Bonan, G., Hess, P. G. and Sparks, J. P.: The influence of chronic ozone exposure on global carbon and water cycles, J. Clim., 28(1), 292–305, doi:10.1175/JCLI-D-14-00223.1, 2015a.

Lombardozzi, D. L., Bonan, G. B., Smith, N. G., Dukes, J. S. and Fisher, R. A.: Temperature acclimation of photosynthesis and respiration: A key uncertainty in the carbon cycle-climate feedback, Geophys. Res. Lett., 42(20), 8624–8631, doi:10.1002/2015GL065934, 2015b.

Louis, B. P., Maron, P. A., Viaud, V., Leterme, P. and Menasseri-Aubry, S.: Soil C and N models that integrate microbial diversity, Environ. Chem. Lett., 1–14, doi:10.1007/s10311-016-0571-5, 2016.

Loveland, T. R., Reed, B. C., Ohlen, D. O., Brown, J. F., Zhu, Z., Yang, L. and Merchant, J. W.: Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data, Int. J. Remote Sens., 21(6–7), 1303–1330, doi:10.1080/014311600210191, 2000.

Ludwig, M., Achtenhagen, J., Miltner, A., Eckhardt, K. U., Leinweber, P., Emmerling, C. and Thiele-Bruhn, S.: Microbial contribution to SOM quantity and quality in density fractions of temperate arable soils, Soil Biol. Biochem., doi:10.1016/j.soilbio.2014.12.002, 2015.

Luo, Y., Ahlström, A., Allison, S. D., Batjes, N. H., Brovkin, V., Carvalhais, N., Chappell, A., Ciais, P., Davidson, E. A., Finzi, A., Georgiou, K., Guenet, B., Hararuk, O., Harden, J. W., He, Y., Hopkins, F., Jiang, L., Koven, C., Jackson, R. B., Jones, C. D., Lara, M. J., Liang, J., Mcguire, A. D., Parton, W., Peng, C., Randerson, J. T., Salazar, A., Sierra, C. A., Smith, M. J., Tian, H., Todd-Brown, K. E. O., Torn, M., Groenigen, K. J., Wang, Y. P., West, T. O., Wei, Y., Wieder, W. R., Xia, J., Xu, X., Xu, X. and Zhou, T.: Toward more realistic projections of soil carbon dynamics by Earth system models, Global Biogeochem. Cycles, 1–17, doi:10.1002/2015GB005239.Received, 2016.

von Lützow, M. and Kögel-Knabner, I.: Temperature sensitivity of soil organic matter decomposition-what do we know?, Biol. Fertil. Soils, 46(1), 1–15, doi:10.1007/s00374-009-0413-8, 2009.

Manzoni, S. and Porporato, A.: Soil carbon and nitrogen mineralization: Theory and models across scales, Soil Biol. Biochem., 41(7), 1355–1379, doi:10.1016/j.soilbio.2009.02.031, 2009.

Marín-Spiotta, E., Barnes, R. T., Berhe, A. A., Hastings, M. G., Mattheis, A., Schneider, B. and Williams, B. M.: Hostile climates are barriers to diversifying the geosciences, Adv. Geosci., 53, 117–127, 2020.

Martin, L. J., Blossey, B. and Ellis, E.: Mapping where ecologists work: Biases in the global distribution of terrestrial ecological observations, Front. Ecol. Environ., 10(4), 195–201, doi:10.1890/110154, 2012.

Mattheis, A., Murphy, M. and Marin-Spiotta, E.: Examining intersectionality and inclusivity in geosciences education research: A synthesis of the literature 2008–2018, J. Geosci. Educ., 67(4), 505–517, doi:10.1080/10899995.2019.1656522, 2019.

Maynard, D. S., Crowther, T. W. and Bradford, M. A.: Fungal interactions reduce carbon use efficiency, Ecol. Lett., In Press, doi:10.1111/ele.12801, 2017.

Medlyn, B. E., Zaehle, S., De Kauwe, M. G., Walker, A. P., Dietze, M. C., Hanson, P. J., Hickler, T., Jain, A. K., Luo, Y., Parton, W., Prentice, I. C., Thornton, P. E., Wang, S., Wang, Y.-P., Weng, E., Iversen, C. M., McCarthy, H. R., Warren, J. M., Oren, R. and Norby, R. J.: Using ecosystem experiments to improve vegetation models, Nat. Clim. Chang., 5(6), 528–534, doi:10.1038/nclimate2621, 2015.

Melillo, J. M., Aber, J. D., Linkins, A. E., Ricca, A., Fry, B. and Nadelhoffer, K. J.: Carbon and nitrogen dynamics along the decay continuum: Plant litter to soil organic matter, Plant Soil, 115(2), 189–198, doi:10.1007/BF02202587, 1989.

Menichetti, L., Ågren, G. I., Barré, P., Moyano, F. and Kätterer, T.: Generic parameters of first-order kinetics accurately describe soil organic matter decay in bare fallow soils over a wide edaphic and climatic range, Sci. Rep., 9(1), 1–12, doi:10.1038/s41598-019-55058-1, 2019.

Mercado, L. M., Medlyn, B. E., Huntingford, C., Oliver, R. J., Clark, D. B., Sitch, S., Zelazowski, P., Kattge, J., Harper, A. B. and Cox, P. M.: Large sensitivity in land carbon storage due to geographical and temporal variation in the thermal response of photosynthetic capacity, New Phytol., 218(4), 1462–1477, doi:10.1111/nph.15100, 2018.

Metcalfe, D. B., Hermans, T. D. G., Ahlstrand, J., Becker, M., Berggren, M., Björk, R. G., Björkman, M. P., Blok, D., Chaudhary, N., Chisholm, C., Classen, A. T., Hasselquist, N. J., Jonsson, M., Kristensen, J. A., Kumordzi, B. B., Lee, H., Mayor, J. R., Prevéy, J., Pantazatou, K., Rousk, J., Sponseller, R. A., Sundqvist, M. K., Tang, J., Uddling, J., Wallin, G., Zhang, W., Ahlström, A., Tenenbaum, D. E. and Abdi, A. M.: Patchy field sampling biases understanding of climate change impacts across the Arctic, Nat. Ecol. Evol., 2(9), 1443–1448, doi:10.1038/s41559-018-0612-5, 2018.

Mikutta, R., Turner, S., Schippers, A., Gentsch, N., Meyer-Stüve, S., Condron, L. M., Peltzer, D. A., Richardson, S. J., Eger, A., Hempel, G., Kaiser, K., Klotzbücher, T. and Guggenberger, G.: Microbial and abiotic controls on mineral-associated organic matter in soil profiles along an ecosystem gradient, Sci. Rep., 9(1), 1–9, doi:10.1038/s41598-019-46501-4, 2019.

Miller, A. J., Amundson, R., Burke, I. C. and Yonker, C.: The effect of climate and cultivation on soil organic C and N, Biogeochemistry, 67(1), 57–72, doi:10.1023/B:BIOG.0000015302.16640.a5, 2004.

Minasny, B., Malone, B. P., McBratney, A. B., Angers, D. A., Arrouays, D., Chambers, A., Chaplot, V., Chen, Z. S., Cheng, K., Das, B. S., Field, D. J., Gimona, A., Hedley, C. B., Hong, S. Y., Mandal, B., Marchant, B. P., Martin, M., McConkey, B. G., Mulder, V. L., O'Rourke, S., Richer-de-Forges, A. C., Odeh, I., Padarian, J., Paustian, K., Pan, G., Poggio, L., Savin, I., Stolbovoy, V., Stockmann, U., Sulaeman, Y., Tsui, C. C., Vï¿½gen, T. G., van Wesemael, B. and Winowiecki, L.: Soil carbon 4 per mille, Geoderma, 292, 59–86, doi:10.1016/j.geoderma.2017.01.002, 2017.

Molenaar, D., van Berlo, R., de Ridder, D. and Teusink, B.: Shifts in growth strategies reflect tradeoffs in cellular economics., Mol. Syst. Biol., 5(323), 323, doi:10.1038/msb.2009.82, 2009.

Moore, J. A. M., Anthony, M. A., Pec, G. J., Trocha, L. K., Trzebny, A., Geyer, K. M., van Diepen, L. T. A. and Frey, S. D.: Fungal community structure and function shifts with atmospheric nitrogen deposition, Glob. Chang. Biol., 27(7), 1349–1364, doi:10.1111/gcb.15444, 2021.

Moorhead, D. L. and Sinsabaugh, R. L.: A theoretical model of litter decay and microbial interaction, Ecol. Monogr., 76(2), 151–174, 2006.

Mooshammer, M., Wanek, W., Zechmeister-Boltenstern, S. and Richter, A.: Stoichiometric imbalances between terrestrial decomposer communities and their resources: mechanisms and implications of microbial adaptations to their resources, Front. Microbiol., 5(February), 1–10, doi:10.3389/fmicb.2014.00022, 2014.

Morales, N., Bisbee O'Connell, K., McNulty, S., Berkowitz, A., Bowser, G., Giamellaro, M. and Miriti, M. N.: Promoting inclusion in ecological field experiences: Examining and overcoming barriers to a professional rite of passage, Bull. Ecol. Soc. Am., 101(4), 1–10, doi:10.1002/bes2.1742, 2020.

Morford, S. L., Houlton, B. Z. and Dahlgren, R. A.: Increased forest ecosystem carbon and nitrogen storage from nitrogen rich bedrock, Nature, 477(7362), 78–84, doi:10.1038/nature10415, 2011.

Nancarrow, S., Booth, A., Ariss, S., Smith, T., Enderby, P. and Roots, A.: Ten principles of good interdisciplinary team work, Hum. Resour. Health, 11(19), 1–11, 2013.

Neina, D.: The Role of Soil pH in Plant Nutrition and Soil Remediation, Appl. Environ. Soil Sci., 2019(3), doi:10.1155/2019/5794869, 2019.

Norby, R. J., Warren, J. M., Iversen, C. M., Medlyn, B. E. and McMurtrie, R. E.: CO2 enhancement of forest productivity constrained by limited nitrogen availability, Proc. Natl. Acad. Sci., 107(45), 19368–19373, doi:10.1073/pnas.1006463107, 2010.

Nottingham, A. T., Turner, B. L., Winter, K., Chamberlain, P. M., Stott, A. and Tanner, E. V. J.: Root and arbuscular mycorrhizal mycelial interactions with soil microorganisms in lowland tropical forest, FEMS Microbiol. Ecol., 85(1), 37–50, doi:10.1111/1574-6941.12096, 2013.

O'Rourke, M., Crowley, S., Eigenbrode, S. and Wulfhorst, J.: Enhancing Communication & Collaboration in Interdisciplinary Research, SAGE Publications, Inc., 2014.

Parton, W., Silver, W. L., Burke, I. C., Grassens, L., Harmon, M. E., Currie, W. S., King, J. Y., Adair, E. C., Brandt, L. A., Hart, S. C. and Fasth, B.: Global-Scale Similarities in Nitrogen Release Patterns During Long-Term Decomposition, Science (80-.)., 315(January), 361–364, 2007.

Parton, W. J. and Rasmussen, P. E.: Long-Term Effects of Crop Management in Wheat-Fallow: II. CENTURY Model Simulations, Soil Sci. Soc. Am. J., 58(2), 530, doi:10.2136/sssaj1994.03615995005800020040x, 1994.

Pfeiffer, T., Schuster, S. and Bonhoeffer, S.: Cooperation and competition in the evolution of ATP-producing pathways., Science (80-.)., 292(5516), 504–7, doi:10.1126/science.1058079, 2001.

Phillips, R. P., Meier, I. C., Bernhardt, E. S., Grandy, S., Wickings, K., Finzi, A. C., Knops, J., Grandy, a S., Wickings, K. and Finzi, A. C.: Roots and fungi accelerate carbon and nitrogen cycling in forests exposed to elevated CO2., Ecol. Lett., 15(9), 1042–9, doi:10.1111/j.1461-0248.2012.01827.x, 2012.

Plante, A. F., Conant, R. T., Carlson, J., Greenwood, R., Shulman, J. M., Haddix, M. L. and Paul, E. A.: Decomposition temperature sensitivity of isolated soil organic matter fractions, Soil Biol. Biochem., 42(11), 1991–1996, doi:10.1016/j.soilbio.2010.07.022, 2010.

Poeplau, C., Don, A., Six, J., Kaiser, M., Benbi, D., Chenu, C., Cotrufo, M. F., Derrien, D., Gioacchini, P., Grand, S., Gregorich, E., Griepentrog, M., Gunina, A., Haddix, M., Kuzyakov, Y., Kühnel, A., Macdonald, L. M., Soong, J., Trigalet, S., Vermeire, M.-L., Rovira, P., van Wesemael, B., Wiesmeier, M., Yeasmin, S., Yevdokimov, I. and Nieder, R.: Isolating organic carbon fractions with varying turnover rates in temperate agricultural soils – A comprehensive method comparison, Soil Biol. Biochem., 125(June 2017), 10–26, doi:10.1016/j.soilbio.2018.06.025, 2018.

Powlson, D. S., Smith, P. and Smith, J. U.: Evaluation of soil organic matter models : using existing long-term datasets. [online] Available from: https://books.google.com/books?hl=en&lr=&id=8FHmCAAAQBAJ&oi=fnd&pg=PA4&dq=powlson+model&ots=AewsEbBF_S&sig=rt3mNe5oZS3J7YHzvcf_NAnTpeA (Accessed 9 August 2018), 1996.

Prentice, I. C., Dong, N., Gleason, S. M., Maire, V. and Wright, I. J.: Balancing the costs of carbon gain and water transport: Testing a new theoretical framework for plant functional ecology, Ecol. Lett., 17(1), 82–91, doi:10.1111/ele.12211, 2014.

Prentice, I. C., Liang, X., Medlyn, B. E. and Wang, Y.-P.: Reliable, robust and realistic: the three R's of next-generation land-surface modelling, Atmos. Chem. Phys, 15, 5987–6005, doi:10.5194/acp-15-5987-2015, 2015.

Randerson, J. T., Thompson, M. V., Malmstrom, C. M., Field, C. B. and Fung, I. Y.: Substrate limitations for heterotrophs: Implications for models that estimate the seasonal cycle of atmospheric CO2, Global Biogeochem. Cycles, 10(4), 585–602, doi:10.1029/96GB01981, 1996.

Rasmussen, C., Heckman, K., Wieder, W. R., Keiluweit, M., Lawrence, C. R., Berhe, A. A., Blankinship, J. C., Crow, S. E., Druhan, J. L., Hicks Pries, C. E., Marin-Spiotta, E., Plante, A. F., Schädel, C., Schimel, J. P., Sierra, C. A., Thompson, A. and Wagai, R.: Beyond clay: towards an improved set of variables for predicting soil organic matter content, Biogeochemistry, 137(3), 297–306, doi:10.1007/s10533-018-0424-3, 2018.

Reed, S. C., Yang, X. and Thornton, P. E.: Incorporating phosphorus cycling into global modeling efforts: a worthwhile, tractable endeavor, New Phytol., n/a-n/a, doi:10.1111/nph.13521, 2015.

Richardson, L. F.: Weather Prediction by Numerical Process, Cambridge University Press., 1922.

Rinkes, Z. L., Sinsabaugh, R. L., Moorhead, D. L., Grandy, a S. and Weintraub, M. N.: Field and lab conditions alter microbial enzyme and biomass dynamics driving decomposition of the same leaf litter., Front. Microbiol., 4(September), 260, doi:10.3389/fmicb.2013.00260, 2013.

Robertson, A. D., Paustian, K., Ogle, S., Wallenstein, M. D., Lugato, E. and Cotrufo, M. F.:

Unifying soil organic matter formation and persistence frameworks: the MEMS model, Biogeosciences Discuss., 1–36, doi:10.5194/bg-2018-430, 2018.

Rocci, K. S., Lavallee, J. M., Stewart, C. E. and Cotrufo, M. F.: Soil organic carbon response to global environmental change depends on its distribution between mineral-associated and particulate organic matter: A meta-analysis, Sci. Total Environ., 793, 148569, doi:10.1016/j.scitotenv.2021.148569, 2021.

Rogers, A., Medlyn, B. E. and Dukes, J. S.: Improving representation of photosynthesis in Earth System Models, New Phytol., 204(1), 12–14, doi:10.1111/nph.12972, 2014.

Rogers, A., Medlyn, B. E., Dukes, J. S., Bonan, G., von Caemmerer, S., Dietze, M. C., Kattge, J., Leakey, A. D. B., Mercado, L. M., Niinemets, Ü., Prentice, I. C., Serbin, S. P., Sitch, S., Way, D. A. and Zaehle, S.: A roadmap for improving the representation of photosynthesis in Earth system models, New Phytol., 213(1), 22–42, doi:10.1111/nph.14283, 2017.

Sagan, C., Toon, O. B. and Pollack, J. B.: Anthropogenic Albedo Changes and the Earth's Climate, Science (80-.)., 206(4425), 1363–1368, 1979.

Sato, N., Sellers, P. J., Randall, D. A., Schneider, E. K., Shukla, J., Kinter III, J. L., Hou, Y.-T. and Albertazzi, E.: Effects of Implementing the Simple Biosphere Model in a General Circulation Model, J. Atmos. Sci., 46(18), 2757–2782, 1989.

Schimel, J.: Biogeochemical Models: Implicit versus Explicit Microbiology, Glob. Biogeochem. Cycles Clim. Syst., 177–183, doi:10.1016/B978-012631260-7/50015-7, 2001.

Schimel, J.: Microbes and global carbon, Nat. Clim. Chang., 3(10), 867–868, doi:10.1038/nclimate2015, 2013.

Schimel, J. P. and Schaeffer, S. M.: Microbial control over carbon cycling in soil, Front. Microbiol., 3(September), 1–11, doi:10.3389/fmicb.2012.00348, 2012.

Schimel, J. P. and Weintraub, M. N.: The implications of exoenzyme activity on microbial carbon and nitrogen limitation in soil: a theoretical model, Soil Biol. Biochem., 35(4), 549–563, doi:Doi 10.1016/S0038-0717(03)00015-4, 2003.

Schmidt, M. W. I., Torn, M. S., Abiven, S., Dittmar, T., Guggenberger, G., Janssens, I. A., Kleber, M., Kögel-Knabner, I., Lehmann, J., Manning, D. A. C., Nannipieri, P., Rasse, D. P., Weiner, S. and Trumbore, S. E.: Persistence of soil organic matter as an ecosystem property., Nature, 478(7367), 49–56, doi:10.1038/nature10386, 2011.

Schneider, S. H., & Dickinson, R. E.: Climate modeling, Rev. Geophys., 12(3), 447 [online] Available from: https://doi.org/10.1029/rg012i003p00447, 1974.

Schulten, H. R. and Schnitzer, M.: The chemistry of soil organic nitrogen: A review, Biol. Fertil. Soils, 26(1), 1–15, doi:10.1007/s003740050335, 1997.

Scott, H. G. and Smith, N. G.: A model of C4 photosynthetic acclimation based on least-cost optimality theory suitable for Earth System Model incorporation, J. Adv. Model. Earth Syst., 2021.

Sellers, P. J., Mintz, Y., Sud, Y. C. and Dalcher, A.: A Simple Biosphere Model (SiB) for Use within General Circulation Models, J. Atmos. Sci., 43(6), 505–531, 1986.

Sellers, P. J., Randall, D. A., Collatz, G. J., Berry, J. A., Field, C. B., Dazlich, D. A., Zhang, C., Collelo, G. D. and Bounoua, L.: A revised land surface parameterization (SiB2) for atmospheric GCMs. Part I: Model formulation, J. Clim., 9(4), 676–705, doi:10.1175/1520-0442(1996)009<0676:ARLSPF>2.0.CO;2, 1996.

Shi, Z., Crowell, S., Luo, Y. and Moore, B.: Model structures amplify uncertainty in predicted soil carbon responses to climate change, Nat. Commun., 9(1), 1–11, doi:10.1038/s41467-018-04526-9, 2018.

Shukla, J. and Mintz, Y.: Influence of Land-Surface Evapotranspiration on the Earth's Climate, Science (80-.)., 215(4539), 1498–1501, 1982.

Sinsabaugh, R. L., Hill, B. H. and Follstad Shah, J. J.: Ecoenzymatic stoichiometry of microbial organic nutrient acquisition in soil and sediment, Nature, 462(7274), 795–798, doi:10.1038/nature08632, 2009.

Sinsabaugh, R. L., Manzoni, S., Moorhead, D. L. and Richter, A.: Carbon use efficiency of microbial communities: stoichiometry, methodology and modelling., Ecol. Lett., 16(7), 930–9, doi:10.1111/ele.12113, 2013.

Sistla, S. A., Rastetter, E. B. and Schimel, J. P.: Responses of a tundra system to warming using SCAMPS: A stoichiometrically coupled, acclimating microbeplantsoil model, Ecol. Monogr., 84(1), 151–170, doi:10.1890/12-2119.1, 2014.

Sitch, S., Smith, B., Prentice, I. C., Arneth, A., Bondeau, A., Cramer, W., Kaplan, J. O., Levis, S., Lucht, W., Sykes, M. T., Thonicke, K. and Venevsky, S.: Evaluation of ecosystem dynamics, plant geography and terrestrial carbon cycling in the LPJ dynamic global vegetation model, Glob. Chang. Biol., 9(2), 161–185, doi:10.1046/j.1365-2486.2003.00569.x, 2003.

Six, J., Frey, S. D., Thiet, R. K. and Batten, K. M.: Bacterial and Fungal Contributions to Carbon Sequestration in Agroecosystems, Soil Sci. Soc. Am. J., 70(2), 555, doi:10.2136/sssaj2004.0347, 2006.

Smith, N. G. and Dukes, J. S.: Plant respiration and photosynthesis in global-scale models: Incorporating acclimation to temperature and CO2, Glob. Chang. Biol., 19(1), 45–63, doi:10.1111/j.1365-2486.2012.02797.x, 2013.

Smith, N. G. and Keenan, T. F.: Mechanisms underlying leaf photosynthetic acclimation to warming and elevated CO2 as inferred from least-cost optimality theory, Glob. Chang. Biol., 26(9), 5202–5216, doi:10.1111/gcb.15212, 2020.

Smith, N. G., Malyshev, S. L., Shevliakova, E., Kattge, J. and Dukes, J. S.: Foliar temperature acclimation reduces simulated carbon sensitivity to climate, Nat. Clim. Chang., (December), doi:10.1038/nclimate2878, 2015.

Smith, N. G., Malyshev, S. L., Shevliakova, E., Kattge, J. and Dukes, J. S.: Foliar temperature acclimation reduces simulated carbon sensitivity to climate, Nat. Clim. Chang., 6(4), 407–411, doi:10.1038/nclimate2878, 2016.

Smith, N. G., Keenan, T. F., Colin Prentice, I., Wang, H., Wright, I. J., Niinemets, Ü., Crous, K. Y., Domingues, T. F., Guerrieri, R., Yoko Ishida, F., Kattge, J., Kruger, E. L., Maire, V., Rogers, A., Serbin, S. P., Tarvainen, L., Togashi, H. F., Townsend, P. A., Wang, M., Weerasinghe, L. K. and Zhou, S. X.: Global photosynthetic capacity is optimized to the environment, Ecol. Lett., 22(3), 506–517, doi:10.1111/ele.13210, 2019.

Smith, P., Smith, J. U., Powlson, D. S., McGill, W. B., Arah, J. R. M., Chertov, O. G., Coleman, K., Franko, U., Frolking, S., Jenkinson, D. S., Jensen, L. S., Kelly, R. H., Klein-Gunnewiek, H., Komarov, A. S., Li, C., Molina, J. A. E., Mueller, T., Parton, W. J., Thornley, J. H. M. and Whitmore, A. P.: A comparison of the performance of nine soil organic matter models using datasets from seven long-term experiments, Geoderma, 81(1–2), 153–225, doi:10.1016/S0016-7061(97)00087-6, 1997.

Smithwick, E. a. H., Mack, M. C., Turner, M. G., Chapin, F. S., Zhu, J. and Balser, T. C.: Spatial heterogeneity and soil nitrogen dynamics in a burned black spruce forest stand: Distinct controls at different scales, Biogeochemistry, 76(3), 517–537, doi:10.1007/s10533-005-0031-y, 2005.

Soetaert, K. and Herman, P. M. J.: A Practical Guide to Ecological Modelling. Using R as a Simulation Platform., , doi:10.1016/s0075-9511(09)00026-7, 2009.

Sokolov, A. P., Kicklighter, D. W., Melillo, J. M., Felzer, B. S., Schlosser, C. A. and Cronin, T. W.: Consequences of considering carbon-nitrogen interactions on the feedbacks between climate and the terrestrial carbon cycle, J. Clim., 21(15), 3776–3796, doi:10.1175/2008JCLI2038.1, 2008.

Soong, J. L., Fuchslueger, L., Marañon-Jimenez, S., Torn, M. S., Janssens, I. A., Penuelas, J. and Richter, A.: Microbial carbon limitation: The need for integrating microorganisms into our understanding of ecosystem carbon cycling, Glob. Chang. Biol., 26(4), 1953–1961, doi:10.1111/gcb.14962, 2020.

Soranno, P. A., Wagner, T., Collins, S. M., Lapierre, J., Lottig, N. R. and Oliver, S. K.: Spatial and temporal variation of ecosystem properties at macroscales, Ecol. Lett., 1587–1598, doi:10.1111/ele.13346, 2019.

Steinweg, J. M., Plante, A. F., Conant, R. T., Paul, E. A. and Tanaka, D. L.: Patterns of substrate utilization during long-term incubations at different temperatures, Soil Biol. Biochem., 40(11), 2722–2728, doi:10.1016/j.soilbio.2008.07.002, 2008.

Stewart, C. E., Paustian, K., Conant, R. T., Plante, A. F. and Six, J.: Soil carbon saturation: Implications for measurable carbon pool dynamics in long-term incubations, Soil Biol. Biochem., 41(2), 357–366, doi:10.1016/j.soilbio.2008.11.011, 2009.

Sulman, B. N., Phillips, R. P., Oishi, A. C., Shevliakova, E. and Pacala, S. W.: Microbe-driven turnover offsets mineral-mediated storage of soil carbon under elevated CO2, Nat. Clim. Chang., 4(December), 1099–1102, doi:10.1038/NCLIMATE2436, 2014.

Sulman, B. N., Brzostek, E. R., Medici, C., Shevliakova, E., Menge, D. N. L. and Phillips, R. P.: Feedbacks between plant N demand and rhizosphere priming depend on type of mycorrhizal association, Ecol. Lett., 20(8), 1043–1053, doi:10.1111/ele.12802, 2017.

Sulman, B. N., Moore, J. A. M., Abramoff, R. Z., Averill, C., Kivlin, S., Georgiou, K., Sridhar, B., Hartman, M., Wang, G., Wieder, W. R., Bradford, M. A., Luo, Y., Mayes, M. A., Morrison, E., Riley, W. J., Salazar, A., Schimel, J. P., Tang, J. and Classen, A. T.: Multiple models and experiments underscore large uncertainty in soil carbon dynamics, Ecol. Lett., 14, 109–123, doi:10.1007/s10533-018-0509-z, 2018.

Sulman, B. N., Shevliakova, E., Brzostek, E. R., Kivlin, S. N., Malyshev, S., Menge, D. N. L. and Zhang, X.: Diverse mycorrhizal associations enhance terrestrial C storage in a global model, Global Biogeochem. Cycles, 0–3, doi:10.1029/2018GB005973, 2019.

Terrer, C., Vicca, S., Stocker, B. D., Hungate, B. A., Phillips, R. P., Reich, P. B., Finzi, A. C. and Prentice, I. C.: Ecosystem responses to elevated CO2 governed by plant–soil interactions and the cost of nitrogen acquisition, New Phytol., 217(2), 507–522, doi:10.1111/nph.14872, 2018.

Thomas, R. Q., Zaehle, S., Templer, P. H. and Goodale, C. L.: Global patterns of nitrogen limitation: confronting two global biogeochemical models with observations., Glob. Chang. Biol., 19(10), 2986–98, doi:10.1111/gcb.12281, 2013.

Thomas, R. Q., Brookshire, E. N. J. and Gerber, S.: Nitrogen limitation on land: How can it occur in Earth system models?, Glob. Chang. Biol., 21(5), 1777–1793, doi:10.1111/gcb.12813, 2015.

Thornton, P. E., Lamarque, J.-F., Rosenbloom, N. a. and Mahowald, N. M.: Influence of carbon-nitrogen cycle coupling on land model response to CO 2 fertilization and climate variability, Global Biogeochem. Cycles, 21(4), n/a-n/a, doi:10.1029/2006GB002868, 2007.

Thornton, P. E., Doney, S. C., Lindsay, K., Moore, J. K., Mahowald, N., Randerson, J. T., Fung, I., Lamarque, J.-F., Feddema, J. J. and Lee, Y.-H.: Carbon-nitrogen interactions regulate climate-carbon cycle feedbacks: results from an atmosphere-ocean general circulation model, Biogeosciences Discuss., 6, 2099–2120, doi:10.5194/bgd-6-3303-2009, 2009.

Tian, C., Kasiborski, B., Koul, R., Lammers, P. J., Bucking, H. and Shachar-Hill, Y.: Regulation of the Nitrogen Transfer Pathway in the Arbuscular Mycorrhizal Symbiosis: Gene Characterization and the Coordination of Expression with Nitrogen Flux, Plant Physiol., 153(3), 1175–1187, doi:10.1104/pp.110.156430, 2010.

Tipping, E., Somerville, C. J. and Luster, J.: The C:N:P:S stoichiometry of soil organic matter, Biogeochemistry, 130(1–2), 117–131, doi:10.1007/s10533-016-0247-z, 2016.

Todd-Brown, K. E. O., Hopkins, F. M., Kivlin, S. N., Talbot, J. M. and Allison, S. D.: A framework for representing microbial decomposition in coupled climate models, Biogeochemistry, 109(1–3), 19–33, doi:10.1007/s10533-011-9635-6, 2012.

Todd-Brown, K. E. O., Randerson, J. T., Post, W. M., Hoffman, F. M., Tarnocai, C., Schuur, E. A. G. and Allison, S. D.: Causes of variation in soil carbon simulations from CMIP5 Earth system models and comparison with observations, Biogeosciences, 10(3), 1717–1736, doi:10.5194/bg-10-1717-2013, 2013.

Torn, M. S., Trumbore, S. E., Chadwick, O. A., Vitousek, P. M. and Hendricks, D. M.: Mineral control of soil organic carbon storage and turnover content were measured by horizon down to the depth at which, Nature, 389(1992), 3601–3603, doi:10.1038/38260, 1997.

Wagai, R., Kishimoto-Mo, A. W., Yonemura, S., Shirato, Y., Hiradate, S. and Yagasaki, Y.: Linking temperature sensitivity of soil organic matter decomposition to its molecular structure, accessibility, and microbial physiology, Glob. Chang. Biol., 19(4), 1114–1125, doi:10.1111/gcb.12112, 2013.

Wang, G., Post, W. M. and Mayes, M. A.: Development of microbial-enzyme-mediated decomposition model parameters through steady-state and dynamic analyses, Ecol. Appl., 23(1), 255–272, doi:10.1890/12-0681.1, 2013a.

Wang, G., Post, W. M., Mayes, M. A., Wang, G., Post, W. and Mayes, M. A.: Development of microbial-enzyme-mediated decomposition model parameters through steady-state and dynamic analyses, Ecol. Appl., 23(1), 255–272, 2013b.

Wang, G., Jagadamma, S., Mayes, M. A., Schadt, C. W., Steinweg, J. M., Gu, L. and Post, W. M.: Microbial dormancy improves development and experimental validation of ecosystem model, ISME J., 9(1), 226–237, doi:10.1038/ismej.2014.120, 2014a.

Wang, H., Prentice, I. C., Davis, T. W., Keenan, T. F., Wright, I. J. and Peng, C.: Photosynthetic responses to altitude: an explanation based on optimality principles, New Phytol., 213(3), 976–982, doi:10.1111/nph.14332, 2017a.

Wang, H., Atkin, O. K., Keenan, T. F., Smith, N. G., Wright, I. J., Bloomfield, K. J., Kattge, J., Reich, P. B. and Prentice, I. C.: Acclimation of leaf respiration consistent with optimal photosynthetic capacity, Glob. Chang. Biol., 26(4), 2573–2583, doi:10.1111/gcb.14980, 2020.

Wang, K., Peng, C., Zhu, Q., Zhou, X., Wang, M., Zhang, K. and Wang, G.: Modeling Global Soil Carbon and Soil Microbial Carbon by Integrating Microbial Processes into the Ecosystem Process Model TRIPLEX-GHG, J. Adv. Model. Earth Syst., 9(6), 2368–2384, doi:10.1002/2017MS000920, 2017b.

Wang, Y. P., Law, R. M. and Pak, B.: A global model of carbon, nitrogen and phosphorus cycles for the terrestrial biosphere, Biogeosciences Discuss., 6(5), 9891–9944, doi:10.5194/bgd-6-9891-2009, 2009.

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

Wang, Y. P., Chen, B. C., Wieder, W. R., Leite, M., Medlyn, B. E., Rasmussen, M., Smith, M. J., Agusto, F. B., Hoffman, F. and Luo, Y. Q.: Oscillatory behavior of two nonlinear microbial models of soil carbon decomposition, Biogeosciences, 11(7), 1817–1831, doi:10.5194/bg-11-1817-2014, 2014b.

Wang, Y. P., Jiang, J., Chen-Charpentier, B., Agusto, F. B., Hastings, A., Hoffman, F., Rasmussen, M., Smith, M. J., Todd-Brown, K., Wang, Y., Xu, X. and Luo, Y. Q.: Responses of two nonlinear microbial models to warming and increased carbon input, Biogeosciences, 13(4), 887–902, doi:10.5194/bg-13-887-2016, 2016.

Wang, Y. P., Zhang, H., Ciais, P., Goll, D., Huang, Y., Wood, J. D., Ollinger, S. V., Tang, X. and Prescher, A. K.: Microbial Activity and Root Carbon Inputs Are More Important than Soil Carbon Diffusion in Simulating Soil Carbon Profiles, J. Geophys. Res. Biogeosciences, 126(4), 1–19, doi:10.1029/2020JG006205, 2021.

Washington, W. M., Buja, L. and Craig, A.: The computational future for climate and Earth system models: On the path to petaflop and beyond, Philos. Trans. R. Soc. A Math. Phys. Eng. Sci., 367(1890), 833–846, doi:10.1098/rsta.2008.0219, 2009.

Weil, R. R. and Brady, N. C.: The Nature and Properties of Soils, 15th ed., Pearson., 2016.

Weng, E. S., Malyshev, S., Lichstein, J. W., Farrior, C. E., Dybzinski, R., Zhang, T., Shevliakova, E. and Pacala, S. W.: Scaling from individual trees to forests in an Earth system modeling framework using a mathematically tractable model of height-structured competition, Biogeosciences, 12(9), 2655–2694, doi:10.5194/bg-12-2655-2015, 2015.

Wickham, H.: Reshaping Data with the reshape Package, J. Stat. Softw., 21(12), 1–20 [online] Available from: http://www.jstatsoft.org/v21/i12/, 2007.

Wickham, H.: ggplot2: Elegant Graphics for Data Analysis., [online] Available from: https://cran.r-project.org/web/packages/ggplot2/citation.html, 2016.

Wickham, H.: scales: Scale Functions for Visualization. R package version 1.0.0., [online] Available from: https://cran.r-project.org/package=scales, 2018.

Wickings, K., Grandy, a S., Reed, S. C. and Cleveland, C. C.: The origin of litter chemical complexity during decomposition., Ecol. Lett., 15(10), 1180–8, doi:10.1111/j.1461-0248.2012.01837.x, 2012.

Wieder, W. R., Bonan, G. B. and Allison, S. D.: Global soil carbon projections are improved by modelling microbial processes, Nat. Clim. Chang., 3(10), 909–912, doi:10.1038/nclimate1951, 2013.

Wieder, W. R., Grandy, A. S., Kallenbach, C. M. and Bonan, G. B.: Integrating microbial physiology and physio-chemical principles in soils with the MIcrobial-MIneral Carbon Stabilization (MIMICS) model, Biogeosciences, 11(14), 3899–3917, doi:10.5194/bg-11-3899-

2014, 2014.

Wieder, W. R., Cleveland, C. C., Lawrence, D. M. and Bonan, G. B.: Effects of model structural uncertainty on carbon cycle projections: biological nitrogen fixation as a case study, Environ. Res. Lett., 10(4), 044016, doi:10.1088/1748-9326/10/4/044016, 2015a.

Wieder, W. R., Allison, S. D., Davidson, E. A., Georgiou, K., Hararuk, O., He, Y., Hopkins, F., Luo, Y., Smith, M. J., Sulman, B. N., Todd-Brown, K., Wang, Y.-P., Xia, J. and Xu, X.: Explicitly representing soil microbial processes in Earth system models, Global Biogeochem. Cycles, 29, 1782–1800, doi:10.1002/2015GB005188, 2015b.

Wieder, W. R., Cleveland, C. C., Smith, W. K. and Todd-Brown, K.: Future productivity and carbon storage limited by terrestrial nutrient availability, Nat. Geosci., (April), 1–5, doi:10.1038/ngeo2413, 2015c.

Wieder, W. R., Grandy, A. S., Kallenbach, C. M., Taylor, P. G. and Bonan, G. B.: Representing life in the Earth system with soil microbial functional traits in the MIMICS model, Geosci. Model Dev., 8(6), 1789–1808, doi:10.5194/gmd-8-1789-2015, 2015d.

Wieder, W. R., Hartman, M. D., Sulman, B. N., Wang, Y. P., Koven, C. D. and Bonan, G. B.: Carbon cycle confidence and uncertainty: Exploring variation among soil biogeochemical models, Glob. Chang. Biol., 24(4), 1563–1579, doi:10.1111/gcb.13979, 2018.

Wieder, W. R., Sulman, B. N., Hartman, M. D., Koven, C. D. and Bradford, M. A.: Arctic Soil Governs Whether Climate Change Drives Global Losses or Gains in Soil Carbon, Geophys. Res. Lett., 46(24), 14486–14495, doi:10.1029/2019GL085543, 2019.

Wieder, W. R., Pierson, D., Earl, S., Lajtha, K., Baer, S. G., Ballantyne, F., Asefaw Berhe, A., Billings, S. A., Brigham, L. M., Chacon, S. S., Fraterrigo, J., Frey, S. D., Georgiou, K., De Graaff, M. A., Stuart Grandy, A., Hartman, M. D., Hobbie, S. E., Johnson, C., Kaye, J., Kyker-Snowman, E., Litvak, M. E., Mack, M. C., Malhotra, A., Moore, J. A. M., Nadelhoffer, K., Rasmussen, C., Silver, W. L., Sulman, B. N., Walker, X. and Weintraub, S.: SoDaH: The SOils DAta Harmonization database, an open-source synthesis of soil data from research networks, version 1.0, Earth Syst. Sci. Data, 13(5), 1843–1854, doi:10.5194/essd-13-1843-2021, 2021.

Wilke, C. O.: cowplot: Streamlined Plot Theme and Plot Annotations for “ggplot2”. R package version 0.7.0, URL https//CRAN.R-project.org/package=cowplot. Acessed 18 Oct. 2018, 2016.

Wutzler, T. and Reichstein, M.: Priming and substrate quality interactions in soil organic matter models, Biogeosciences, 10(3), 2089–2103, doi:10.5194/bg-10-2089-2013, 2013.

Xu, X., Thornton, P. E. and Post, W. M.: A global analysis of soil microbial biomass carbon, nitrogen and phosphorus in terrestrial ecosystems, Glob. Ecol. Biogeogr., 22(6), 737–749, doi:10.1111/geb.12029, 2013.

Yang, X., Thornton, P. E., Ricciuto, D. M. and Post, W. M.: The role of phosphorus dynamics in tropical forests - A modeling study using CLM-CNP, Biogeosciences, 11(6), 1667–1681, doi:10.5194/bg-11-1667-2014, 2014.

Zaehle, S. and Dalmonech, D.: Carbon-nitrogen interactions on land at global scales: Current understanding in modelling climate biosphere feedbacks, Curr. Opin. Environ. Sustain., 3(5), 311–320, doi:10.1016/j.cosust.2011.08.008, 2011.

Zaehle, S. and Friend, A. D.: Carbon and nitrogen cycle dynamics in the O-CN land surface model: 1. Model description, site-scale evaluation, and sensitivity to parameter estimates, Global Biogeochem. Cycles, 24(1), 1–13, doi:10.1029/2009GB003521, 2010.

Zaehle, S., Medlyn, B. E., De Kauwe, M. G., Walker, A. P., Dietze, M. C., Hickler, T., Luo, Y., Wang, Y. P., El-Masri, B., Thornton, P., Jain, A., Wang, S., Warlind, D., Weng, E., Parton, W., Iversen, C. M., Gallet-Budynek, A., Mccarthy, H. R., Finzi, A., Hanson, P. J., Prentice, I. C., Oren, R. and Norby, R. J.: Evaluation of 11 terrestrial carbon-nitrogen cycle models against observations from two temperate Free-Air CO2 Enrichment studies, New Phytol., 202(3), 803–822, doi:10.1111/nph.12697, 2014.

Zaehle, S., Jones, C. D., Houlton, B., Lamarque, J. F. and Robertson, E.: Nitrogen availability reduces CMIP5 projections of twenty-first-century land carbon uptake, J. Clim., 28(6), 2494–2511, doi:10.1175/JCLI-D-13-00776.1, 2015.

Zak, D. R., Tilman, D., Parmenter, R. R., Rice, C. W., Fisher, F. M., Vose, J., Milchunas, D. and Martin, C. W.: Plant Production and Soil Microorganisms in Late-Succesional Ecosystems: A Continental-Scale Study, Ecology, 75(8), 2333–2347, 1994.

Zambrano-Bigiarini, M.: hydroGOF: Goodness-of-fit functions for comparison of simulated and observed hydrological time seriesR package version 0.3-10, , doi:10.5281/zenodo.840087, 2017.

Zhang, H., Goll, D. S., Wang, Y. P., Ciais, P., Wieder, W. R., Abramoff, R., Huang, Y., Guenet, B., Prescher, A. K., Viscarra Rossel, R. A., Barré, P., Chenu, C., Zhou, G. and Tang, X.: Microbial dynamics and soil physicochemical properties explain large-scale variations in soil organic carbon, Glob. Chang. Biol., 26(4), 2668–2685, doi:10.1111/gcb.14994, 2020.

Ziehn, T., Kattge, J., Knorr, W. and Scholze, M.: Improving the predictability of global CO2 assimilation rates under climate change, Geophys. Res. Lett., 38(10), 1–5, doi:10.1029/2011GL047182, 2011.