JGR Biogeosciences

RESEARCH ARTICLE
10.1029/2023JG007625

Key Points:

- Simplified Monte Carlo method uses constrained parameter sets and extensive dataset to parameterize decomposition in a microbial model
- The new litter decomposition parameters improved carbon and nitrogen decomposition metrics in a process-based microbial model
- Applying the new parameters to bioenergy systems altered modeled soil carbon with variation by plant traits, management, and litter inputs

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

S. M. Juice,
stephanie.juice@mail.wvu.edu

Citation:

Juice, S. M., Ridgeway, J. R., Hartman, M. D., Parton, W. J., Berardi, D. M., Sulman, B. N., et al. (2024). Reparameterizing litter decomposition using a simplified Monte Carlo method improves litter decay simulated by a microbial model and alters bioenergy soil carbon estimates. *Journal of Geophysical Research: Biogeosciences*, *129*, e2023JG007625. https://doi.org/10.1029/2023JG007625

Received 16 JUN 2023
Accepted 24 FEB 2024

Reparameterizing Litter Decomposition Using a Simplified Monte Carlo Method Improves Litter Decay Simulated by a Microbial Model and Alters Bioenergy Soil Carbon Estimates

S. M. Juice[1,2], J. R. Ridgeway[1,2], M. D. Hartman[2,3], W. J. Parton[2,3], D. M. Berardi[2,4], B. N. Sulman[5], K. E. Allen[6], and E. R. Brzostek[1,2]

[1]Department of Biology, West Virginia University, Morgantown, WV, USA, [2]Center for Advanced Bioenergy and Bioproducts Innovation, University of Illinois Urbana-Champaign, Urbana, IL, USA, [3]Natural Resource Ecology Laboratory, Colorado State University, Fort Collins, CO, USA, [4]Department of Forest, Rangeland and Fire Sciences, University of Idaho, Moscow, ID, USA, [5]Climate Change Science Institute and Environmental Sciences Division, Oak Ridge National Laboratory, Oak Ridge, TN, USA, [6]Manaaki Whenua Landcare Research, Lincoln, New Zealand

Abstract Litter decomposition determines soil organic matter (SOM) formation and plant-available nutrient cycles. Therefore, accurate model representation of litter decomposition is critical to improving soil carbon (C) projections of bioenergy feedstocks. Soil C models that simulate microbial physiology (i.e., microbial models) are new to bioenergy agriculture, and their parameterization is often based on small datasets or manual calibration to reach benchmarks. Here, we reparameterized litter decomposition in a microbial soil C model (CORPSE - Carbon, Organisms, Rhizosphere, and Protection in the Soil Environment) using the continental-scale Long-term Inter-site Decomposition Experiment Team (LIDET) dataset which documents decomposition across a range of litter qualities over a decade. We conducted a simplified Monte Carlo simulation that constrained parameter values to reduce computational costs. The LIDET-derived parameters improved modeled C and nitrogen (N) remaining, decomposition rates, and litter mean residence times as compared to Baseline parameters. We applied the LIDET litter decomposition parameters to a microbial bioenergy model (Fixation and Uptake of Nitrogen – Bioenergy Carbon, Rhizosphere, Organisms, and Protection) to examine soil C estimates generated by Baseline and LIDET parameters. LIDET parameters increased estimated soil C in bioenergy feedstocks, with even greater increases under elevated plant inputs (i.e., by increasing residue, N fertilization). This was due to the integrated effects of plant litter quantity, quality, and agricultural practices (tillage, fertilization). Collectively, we developed a simple framework for using large-scale datasets to inform the parameterization of microbial models that impacts projections of soil C for bioenergy feedstocks.

Plain Language Summary Decomposition breaks down organic matter like leaves and roots, creating soil organic material and releasing essential nutrients for plant and microbial growth. Soil creation and nutrient release are processes that affect how much carbon is stored in soil. Soil carbon storage in bioenergy agriculture may help create a favorable carbon balance for biofuels, ultimately reducing the rate of climate change. However, environmental decision makers need reliable information about how different bioenergy plants change soil carbon stocks to predict long-term outcomes of present-day decisions. These predictions are generated by computer models that mathematically represent ecological processes using observations from field studies. However, some models that include microbial decomposition lack a robust observational and mathematical basis for their representation of decomposition. We used a large-scale litter decomposition dataset and simplified a statistical simulation that is typically complex and time-consuming to improve the mathematical basis for litter decomposition in a soil carbon model. We used the improved decomposition representation in a different model that calculates soil carbon in bioenergy agriculture, and found the new representation increased predicted soil carbon in bioenergy feedstocks. Our statistical, data-based framework can be adopted to help make model predictions more accurate, and environmental management decisions more effective.

1. Introduction

Microbial decomposition of plant litter partitions carbon (C) between the atmosphere and longer-term storage in soil organic matter (SOM) and is the first, rate-limiting step of nutrient cycling (Cotrufo et al., 2013; Wallenstein & Weintraub, 2008). As a microbially-driven process, litter decomposition varies across environmental gradients that affect microbial activity and populations, including climate variables and litter quality (Meentemeyer, 1978; Zhang et al., 2008), with the relative impact of climate or litter quality varying geographically (Dyer et al., 1990). Due to its central role in SOM formation and production of plant-available nutrients, correctly modeling litter decomposition in process-based models is critical to generating realistic projections of soil C. This is particularly relevant to soils under cultivation for bioenergy, where accurate simulation of litter decomposition is essential for management strategies (e.g., residue removal, feedstock choice) that seek to achieve reduced reliance on fossil fuels in addition to soil C accrual and nitrogen (N) retention. While decomposition has traditionally been modeled as a first-order process (i.e., decay rates are determined as a function of the pool size and a decay constant, k) that varies with climate and litter quality, there is growing evidence that simulating the microbial physiology driving the process may improve model outcomes (Todd-Brown et al., 2012; Wieder et al., 2013), particularly under changing environmental conditions (Lawrence et al., 2009; Wieder et al., 2014). The recent development of microbially-explicit ecosystem models designed for use in agroecosystems such as bioenergy agriculture (Juice et al., 2022) presents a novel opportunity to explore the ability of bioenergy to achieve the double benefit of increasing soil C while reducing use of fossil fuels.

Accurate models of litter decomposition are necessary to predict if bioenergy can deliver this double benefit of low C fuel and enhanced soil C storage. One challenge to the models is accurate parameterization of litter decomposition to produce sound results. Often, parameters for litter decomposition in soil models are based on datasets with limited representation of litter chemistry, climatic conditions, and ecosystem types. As such, the accuracy of model outcomes could be limited to sites and species like those used for parameterization, thus hindering widespread utility of the model. To address this issue, soil decomposition models that rely on first-order decomposition dynamics have used large datasets to parameterize litter decomposition (Bonan et al., 2013).

By contrast, most of the new generation of decomposition models that explicitly represent microbial physiology (herein microbial models) have yet to employ large datasets for their parameterization. Instead, they typically parameterize litter decomposition using data from only a single site, lab experiment, or have calibrated litter decomposition parameters to ensure the model meets other benchmarks (e.g., Millennial model: Abramoff et al., 2018; CORPSE: Sulman et al., 2014a; MEND: Wang et al., 2013; but see Woolf & Lehmann, 2019). In comparison, comprehensive efforts to parameterize the MIMICS microbial model improved its ability to perform across a large diversity of environmental conditions and scales (Kyker-Snowman et al., 2020; Pierson et al., 2022; Wang et al., 2021; Zhang et al., 2020) and demonstrated the importance of data assimilation to refine parameters and reduce uncertainty (Shi et al., 2018). While these efforts provide an effective road map for improving the ability of microbial models to project litter decomposition, they have yet to be widely applied to microbial bioenergy models due to microbial models only being recently introduced into the bioenergy space (Berardi et al., 2020; Juice et al., 2022).

Projecting soil C in bioenergy agriculture and other agroecosystems represents one arena where meeting this challenge of improving the parameterization of litter decomposition is critical to the performance of microbial models. For example, litter decomposition rates determine the degree to which leaving more feedstock residue on the field after harvest enhances soil C stocks (Carvalho et al., 2017). In addition, the suite of potential bioenergy feedstocks differs widely in litter chemistry and management strategies (i.e., till vs. no till, N fertilization rates) that alter litter decomposition (Johnson et al., 2007; Lupwayi et al., 2004; Zang et al., 2016) and lead to differences in soil C stocks. Thus, projecting the success of different management strategies and feedstocks is highly dependent on accurate representations of litter decomposition.

Until recently, soil decomposition models that project soil C in bioenergy crop systems have relied on first-order decay functions (Berardi et al., 2020; Qin et al., 2016). Given the critiques against first-order models (e.g., microbes are represented as a C pool with no function, lack of microbial necromass production that preferentially forms stable soil C), there have been efforts to expand the use of microbial models in bioenergy agriculture such as the development of the FUN-BioCROP model (Fixation and Uptake of Nitrogen-Bioenergy Carbon, Rhizosphere, Organisms, and Protection; Juice et al., 2022). FUN-BioCROP is the bioenergy version of the FUN-CORPSE model (Fixation and Uptake of Nitrogen-Carbon, Organisms, Rhizosphere and Protection in the Soil

Environment; Sulman et al., 2017), a microbially-explicit SOM model that was developed for forest soils. Similar to other microbial models, however, litter decomposition in FUN-BioCROP, which stems from the original CORPSE model (Sulman et al., 2014a), relied on field-scale data and laboratory analyses for its parameterization of litter decomposition (Sulman et al., 2014a). As such, there remains uncertainty in projections of microbially-mediated litter decomposition in bioenergy crop systems. Improving projections of litter decomposition is critical to identify management techniques and bioenergy crops that restore soil C and promote N retention in lands that have a legacy of intensive row crop agriculture.

The challenge in using large datasets to improve the parameterization of litter decomposition in microbial models is that it involves inherent tradeoffs between the time and computational cost of the parameterization versus the ability of the models to be generalizable across a wide range of sites that vary in litter quality, climatic conditions, and biomes. One robust method for parameterization of models is to run a Monte Carlo simulation, in which probability distributions for each parameter are randomly sampled and the model is run iteratively, often thousands of times, with the different parameter sets (Luengo et al., 2020; Metropolis & Ulam, 1949). However, Monte Carlo simulations require substantial time and computational power, which can make them overly burdensome and difficult to operationalize, especially if the model is parameterized separately for differing sites or ecosystems. Despite the successful development of these Monte Carlo simulations, most studies still manually calibrate decomposition parameters through trial and error. While the calibration technique has been successfully employed in first order (Bonan et al., 2013) and microbial models (Abramoff et al., 2018; Sulman et al., 2014a; Wang et al., 2013), this practice may hinder the ability of these models to perform across a wide range of environmental gradients. Much of this reliance on manual calibration could be resolved by simplifying and reducing the effort required for Monte Carlo simulations. Therefore, a challenge remains in forging a middle ground between calibration techniques and employing computationally prohibitive Monte Carlo simulations.

In this study, our objective was to improve the parameterization of litter C loss during decomposition in a microbial model that predicts soil C and N cycling in bioenergy agriculture. We base our parameterization on the trajectory of C loss during decomposition because of the importance of soil C to the overall climate implications of bioenergy agriculture. To achieve our objective, we developed a modified Monte Carlo approach for parameterization that reduces the burden of a typical Monte Carlo simulation by using a priori knowledge to constrain the distributions of parameter values included in the analysis. We conducted the modified Monte Carlo simulation using the CORPSE model (Carbon, Organisms, Rhizosphere, and Protection in the Soil Environment; Sulman et al., 2017; Sulman et al., 2014a), which mathematically describes microbial physiology to simulate soil C and N cycling. First, we identified the best parameter set by comparing modeled leaf litter decomposition against measured decomposition in the LIDET dataset (Long-term Inter-site Decomposition Experiment Team; Harmon, 2013; LIDET, 1995), which documents decomposition across a spectrum of litter qualities, climates, and biomes over a 10-year period. We selected the LIDET dataset for this parameterization exercise because it covers a range of litter qualities, and approximates the variability in litter quality of bioenergy feedstocks. Furthermore, studies that compared decomposition in natural versus agricultural systems attributed the differences to litter characteristics (Bakker et al., 2011), which are well represented in the LIDET dataset. Second, we compared the ability of the new and old litter decomposition parameters to capture known patterns of N immobilization during litter decay (Parton et al., 2007) and the resulting impacts on downstream C and N pools. Finally, we assessed the impact of the litter decomposition reparameterization on the biogeochemistry of bioenergy agriculture by applying the new parameters in the FUN-BioCROP model (Juice et al., 2022). We examined the ability of the new parameters to simulate the impacts of differences in residue removal and variability between feedstocks in litter chemistry, litter quantity, and management strategies. The methods we outline here provide a framework for using large-scale datasets to better parameterize microbial models, increasing their applicability to a wide variety of sites, and improving the accuracy of their projections.

2. Materials and Methods

To reparameterize and improve modeled leaf litter C loss during decomposition in a widely applied soil microbial model (CORPSE; Sulman et al., 2017; Sulman et al., 2018; Sulman et al., 2014a; Sulman et al., 2019) we implemented a multi-step parameterization, validation, and analysis process (Figure 1). First, we ran a modified Monte Carlo approach using data from a global, long-term litter decomposition study (LIDET, 1995) to select the litter decomposition parameter set that minimized model error (herein, "LIDET parameters"). We assessed model improvement by comparing leaf litter C loss, decomposition rate (k), and leaf litter mean residence time (MRT)

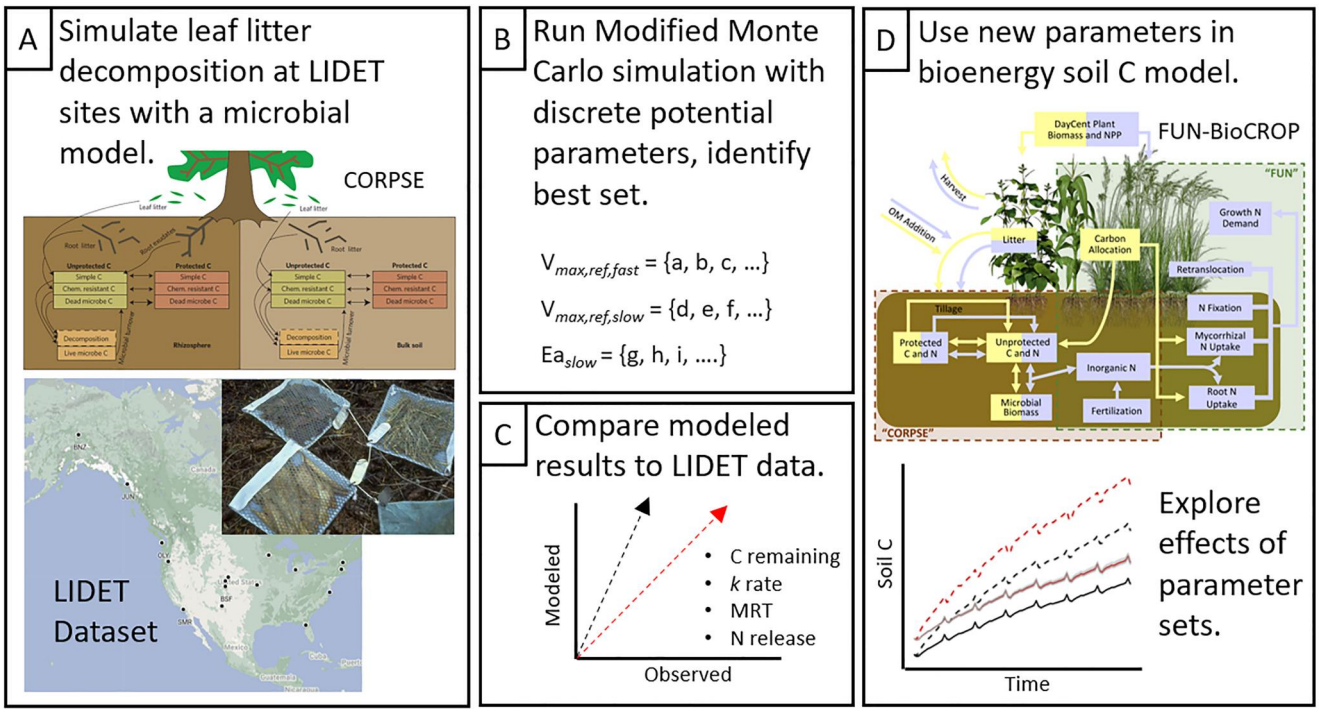

Figure 1. Schematic outline of methods used in this study. (a) We used a subset of the LIDET dataset (Harmon, 2013) to reparameterize litter decomposition in the CORPSE model (Sulman et al., 2014a). (b) We developed a modified Monte Carlo technique that utilized discrete parameter sets instead of parameter distributions. (c) We then compared the Baseline and new LIDET parameter model results to the LIDET field data using various decomposition metrics. (d) Finally, we used the new LIDET litter decomposition parameters in the FUN-BioCROP model (Juice et al., 2022) to simulate bioenergy soil C under different feedstock and management scenarios. C: carbon; CORPSE: Carbon, Organisms, Rhizosphere, and Protection in the Soil Environment; FUN-BioCROP: Fixation and Uptake of Nitrogen-Bioenergy Carbon, Rhizosphere, Organisms, and Protection; LIDET: Long-term Intersite Decomposition Experiment Team; MRT: mean residence time; N: nitrogen.

from LIDET field observations with model-simulated values using both Baseline and reparameterized LIDET decomposition parameters. We also evaluated the effect of the LIDET parameters on modeled N immobilization and release during decomposition. We then updated the leaf litter decomposition parameters in the FUN-BioCROP model (Juice et al., 2022) using the reparameterized LIDET decomposition parameters, which was possible because FUN-BioCROP uses CORPSE litter decomposition functions and parameters. To examine the difference in model estimates due to the reparameterization, we used both the baseline version of FUN-BioCROP and the reparameterized version to simulate soil C stocks under four bioenergy feedstock scenarios (corn-corn-soybean, fertilized *Miscanthus*, unfertilized *Miscanthus*, and switchgrass) and compared the results generated by the different parameter sets for each feedstock. Because the difference in soil C produced by the two parameter sets varied with plant traits (i.e., annual vs. perennial), we ran further model sensitivity tests to determine the relative importance of changing plant residue, litter quantity, and litter quality on soil C estimates generated by the Baseline and LIDET parameter sets. All analyses were conducted in the R statistical computing environment (R Core Team, 2023).

2.1. LIDET Data Set

We selected the LIDET dataset (Harmon, 2013) to reparameterize litter C loss during decomposition in the CORPSE model because it documents leaf litter mass loss over a 10-year period and spans a range of litter qualities, climatic conditions, and biomes (Tables S1 and S2 in Supporting Information S1; Gholz et al., 2000; LIDET, 1995). LIDET was a reciprocal litterbag study with litter from 25 plant species deployed in 27 sites across a range of environmental conditions in North and Central America (Gholz et al., 2000). Each site received the same six standard species of leaf litter that were chosen to cover a range of N and lignin contents, plus an additional wildcard litter that was deployed at a subset of sites (Table S2 in Supporting Information S1). Litterbags (20 cm^2, 1 mm nylon mesh top and 55 μm Dacron cloth bottom) contained 10 g of air-dried leaf litter, with four replicate bags for each species, site, and time. Litterbags were placed on the ground in each site in the autumn

of 1990, and collected annually for 10 years, except in the subtropical and tropical sites where they were collected every 3–6 months to account for the faster decomposition rate (LIDET, 1995). At each collection time, ash-free mass of the litterbags was determined and replicate bags were averaged to calculate the mean mass remaining of each species (Gholz et al., 2000; Long-term Intersite Decomposition Experiment Team (LIDET), 1995).

In this study, we used a subset of LIDET data from 16 sites representing 9 biomes (Table S1 in Supporting Information S1): agriculture (1 site), arid grassland (2 sites), humid grassland (2 sites), chaparral (1 site), boreal forest (1 site), conifer forest (5 sites), deciduous forest (2 sites), saltmarsh (1 site), and alpine tundra (1 site). Sites were selected according to the availability of data to run the CORPSE model (i.e., site-specific soil temperature, soil moisture, and litter production). For the reparameterization, we included all species decomposed at each of the 20 sites (standard litters and wildcards, Table S2 in Supporting Information S1). This included six grass species, seven broadleaf species, four conifers, and two shrubs. To analyze model performance between the Baseline and improved parameters, we focused on the six standard litter species (indicated in Table S2 in Supporting Information S1).

2.2. Model Description

We reparameterized C loss during litter decomposition in the CORPSE model (full description and model equations in Sulman et al., 2017; Sulman et al., 2014a), which simulates decomposition using microbial activity, rhizosphere processes, and physical protection of SOM. CORPSE has three compartments that together comprise the whole soil profile: the litter compartment that corresponds to the leaf litter layer on the soil surface, rhizosphere soil that is closely associated with roots, and bulk soil. Each compartment contains its own microbial biomass and organic matter (OM) pools. The OM pools are divided into three chemical classes —simple, chemically resistant, and microbial necromass— as well as living microbial C and N. The chemical classes have unique parameters for microbial enzymatic decomposition and microbial C and N uptake efficiencies. Additionally, each chemical class has physically protected and unprotected pools, except for in the litter compartment which only contains unprotected OM. Each chemical class of unprotected OM gets converted into protected OM at a different rate. Microbial necromass has the highest protection rate in the model, causing more protected OM formation to result from decomposition of simple OM that fosters greater microbial growth and necromass production. Likewise, simple OM addition results in relatively higher microbial activity than other chemical classes, and can lead to decomposition priming effects by accelerating microbial activity and growth. Decomposition in CORPSE occurs only in the unprotected OM pools, with the rate determined by microbial biomass, soil temperature, and soil moisture. In turn, the microbial biomass grows through decomposition. The C:N of the microbial biomass pool relative to the OM being decomposed determines whether net N mineralization or immobilization occurs, with insufficient C leading to net N mineralization and insufficient N leading to N immobilization. Decomposition and microbial biomass turnover also result in N mineralization, with inorganic N simulated as a single pool without differentiation into ammonium and nitrate.

Parameters governing litter decomposition in the original CORPSE model were calibrated based on a laboratory incubation and values from the literature (Sulman et al., 2014b). Specifically, the relative maximum enzymatic decomposition rate ($V_{max,ref}$), the Michaelis-Menten half-saturation constant (k), microbial turnover time (T_{mic}), and the C efficiency of microbial turnover (e_t) were calibrated for the model to match microbial biomass and substrate mineralization observed in a laboratory incubation study that added isotopically labeled glucose to grassland soils and tracked microbial biomass over 20 days (Sulman et al., 2014b; Wu et al., 1993). The remaining litter decomposition parameters include activation energy (E_a) and microbial C uptake efficiency (ε_i) for decomposition of each C class, which were based on published values (Table 2 in Sulman et al., 2014b). Importantly, the original CORPSE parameterization was the same for all three soil compartments (litter, rhizosphere, and bulk), and validation of decomposition parameters focused on reproducing microbial and SOM dynamics as opposed to litter decomposition trajectories. The only parameter calibrated for litter decomposition itself was the plant parameter defining the fast-decomposing fraction of the leaf litter (i.e., fast fraction). Two fast fractions were calibrated to match decomposition curves from the Duke FACE experiment (Sulman et al., 2014b): one for loblolly pine, and one for an average of four deciduous tree species. In contrast, our study uses species-specific fast fractions that are calculated based on the lignin:N ratio of the tissue, and we aim to reparameterize decomposition for leaf litter specifically in the litter compartment using a large-scale litter decomposition dataset.

Data streams required to run CORPSE include above and belowground litter production, fine root turnover, and soil temperature and moisture. Litter quality is simulated through the C:N ratio of leaf and root litter, as well as the fast fraction of the litter as defined by the lignin:N ratio of the tissues (see caption of Table S2 in Supporting Information S1 for equation; Parton et al., 1987). Leaf litter inputs are partitioned between the simple and chemically resistant C and N pools in the litter compartment according to the leaf litter fast fraction. Root litter inputs resulting from root turnover enter the rhizosphere and bulk soil compartments. They are fractionated between the rhizosphere and bulk soil based on the rhizosphere volume, and the rhizosphere was assumed to be a constant at 15% of the total soil volume for the reparameterization. Partitioning root inputs between the rhizosphere and bulk compartments, rather than exclusively into the rhizosphere, is designed to account for the dynamic nature of the rhizosphere which can change location within the soil profile, as well as maintain consistency on a volume basis between the two compartments. Root inputs are additionally divided between simple or chemically resistant C and N pools within both the rhizosphere and bulk soil compartments according to the root fast fraction.

The FUN-BioCROP model (full description in Juice et al., 2022) calculates plant C investment in different N uptake pathways, and the impacts of that C partitioning on OM decomposition and protection in the top 30 cm of the soil profile in agricultural systems. FUN-BioCROP is based on the framework of the coupled FUN-CORPSE model (Sulman et al., 2017), and uses the same decomposition functions and parameters as CORPSE. FUN-BioCROP is additionally coupled to the DayCent-CABBI model (offline coupling; Berardi et al., 2020; Hartman et al., 2022; Moore et al., 2020; Parton et al., 1998), using above- and belowground plant biomass and net primary productivity data calculated by DayCent-CABBI. Agricultural practices simulated by FUN-BioCROP include tillage, fertilization, OM addition, and harvest. Tillage in FUN-BioCROP occurs within tilled rhizosphere and tilled bulk soil compartments, the size of which are determined by a tillage disturbance parameter (set to 15% of the soil profile for the current study). This results in FUN-BioCROP having five compartments that together comprise the whole soil profile (litter that corresponds to the leaf litter layer on the soil surface, tilled rhizosphere, tilled bulk, rhizosphere, and bulk). When the soil is tilled, it triggers the transfer of a set percentage of OM from protected pools into unprotected pools within the tilled soil compartments, making that OM newly available for microbial decomposition. This mimics the in situ effects of soil tillage, which disrupts soil aggregate turnover and formation (Six et al., 1999), releasing previously protected particulate organic matter and making it susceptible to decomposition (Jastrow & Miller, 1997). The increase in unprotected OM can lead to priming of decomposition, both in agricultural fields (Kan et al., 2020; Mo et al., 2021) and in the CORPSE model (Sulman et al., 2014a), accelerating the decomposition of the unprotected OM pools. Fertilization occurs as an addition to the inorganic N pool, and amended OM is partitioned equally between the litter compartment simple and chemically resistant C pools, and into the simple and chemically resistant N pools according to the C:N ratio of the organic matter. Harvests in FUN-BioCROP reduce the aboveground litter production proportionally to the extent of the removal.

As in CORPSE, leaf and root litter chemistry in FUN-BioCROP is represented as the fast-decomposing fraction of the tissue for each species as determined by its lignin:N ratio (Parton et al., 1987). We averaged the lignin:N of young and old leaves (not including stems) to determine the fast fraction of the leaf litter, and the root fast fraction was determined by the lignin:N of fine roots. FUN-BioCROP additionally has plant-specific root morphology parameters — root diameter and specific root length— which are used along with belowground biomass to calculate the rhizosphere volume, assuming the rhizosphere extends 1 mm cylindrically from the root surface. As in CORPSE, leaf litter inputs enter the litter compartment and are partitioned between simple and chemically resistant OM pools according to the fast fraction. At each time step, a portion of the litter compartment mixes into the tilled and untilled rhizosphere and bulk compartments according to both the rhizosphere size (to determine the proportion deposited into rhizosphere vs. bulk) and the tillage disturbance parameter (to determine the proportion deposited into tilled vs. untilled). Root litter is similarly partitioned between the tilled and untilled rhizosphere and bulk soil compartments according to the rhizosphere size and the tillage disturbance parameter, and between simple and chemically resistant OM pools according to the root fast fraction.

Within FUN-BioCROP, the FUN model (Fixation and Uptake of Nitrogen; detailed description and model equations can be found in Brzostek et al., 2014; Fisher et al., 2010; Shi et al., 2016) optimizes plant C investment in N acquisition to meet plant N demand and maximize net primary productivity. The FUN resistance framework calculates simultaneous N uptake across five possible pathways: (a) biological N fixation; (b) retranslocation; (c) non-mycorrhizal root uptake; (d) mycorrhizal root uptake; and (e) N from storage. To minimize the total C spent

by the plant on N uptake, the C cost of each pathway (i.e., its resistance) determines its uptake rate to fulfill the plant's calculated N demand. In the coupled FUN-BioCROP model (as in FUN-CORPSE), C spent on root-facilitated N uptake in FUN is transferred into CORPSE soil C pools. Specifically, C spent on non-mycorrhizal N uptake represents root exudates and as such is deposited into the simple C pools in the tilled and untilled rhizosphere compartments according to the tillage disturbance parameter. Conversely, C spent on mycorrhizal N uptake is added to both the rhizosphere and bulk soil compartments to better simulate the range of mycorrhizal hyphae throughout the soil profile. The C addition is partitioned between rhizosphere and bulk by the rhizosphere size, and between tilled and untilled by the tillage disturbance parameter. Finally, the calculated pool of inorganic N in FUN is used to update the CORPSE inorganic N pool.

2.3. Reparameterization of Litter C Loss During Decomposition

We used site-specific and species-specific data to run CORPSE at each LIDET site (Juice et al., 2023a). We used soil temperature data and a soil moisture scalar for each LIDET site that had been previously calculated for use in DayCent LIDET simulations (Bonan et al., 2013). The soil moisture scalar influences decomposition by scaling the base decomposition rate by a value that ranges from 0 to 1 depending on moisture conditions. For both soil temperature and the soil moisture scalar, mean monthly values were linearly interpolated to produce 1 year of mean daily values that was used for the simulations. We used litter production data from each site (Table S3 in Supporting Information S1) to represent litter production, root production, and fine root turnover in CORPSE due to the lack of those individual data streams. Lastly, we represented species litter quality differences in the model with LIDET values of the C:N ratio and lignin:N ratio (Harmon, 2013), the latter of which was used to determine the species-specific fast fraction for CORPSE (Table S2 in Supporting Information S1).

To approximate litterbag field conditions, we modified the CORPSE model to include a litterbag compartment in which we simulated decomposition of an added mass of litter C that did not mix with the other soil compartments, much like a litterbag set on the ground. As opposed to the litter compartment, which mixes into the bulk and rhizosphere compartments, all C lost from the litterbag compartment was due to decomposition alone. The litterbag compartment contained the same C and N pools as the litter compartment (unprotected simple, unprotected chemically resistant, and unprotected necromass), with the same functions and moisture effects on decomposition as the other soil compartments. In the simulations with the Baseline parameters, all the model compartments (litter, litterbag, rhizosphere, and bulk) used the Baseline decomposition parameter values. In the modified Monte Carlo analysis, we changed the decomposition parameters (as described below) in the litter and litterbag compartments, and maintained the Baseline decomposition parameters in the rhizosphere and bulk soil compartments.

We spun up the CORPSE model for each site individually to reach equilibrium soil C prior to the addition of the litterbag compartment. Following spin up, we simulated the addition of a litterbag containing 100 g of leaf litter C (as in Bonan et al., 2013) for each LIDET site and plant species. The initial mass of each SOM chemical class in the litterbag reflected the chemistry as reported in the LIDET dataset (Harmon, 2013), with the unprotected simple C pool represented by the fast fraction of the litter, and the unprotected chemically resistant C pool containing the remainder of the litter. The initial N pools were calculated based on the C pools using the C:N ratio of each litter, which was constant across the fast and slow pools. Finally, we assumed a small initial mass of living microbial C (0.0001 g) on the leaf litter following preparation of the litterbags. Carbon remaining of litter from each species at each site was then simulated using the CORPSE model for the same length of time as in the field experiment (generally 10 years with some variability by site, see Section 2.1). All litterbag simulations began on 1 October to reflect the fall initiation of the LIDET study.

Our modified Monte Carlo approach to reparameterize litter C loss during decomposition in CORPSE consisted of identifying a set of potential values for each of the most influential decomposition parameters (Table 1), then randomly selecting from those values to find the combination that produced results closest to the LIDET field data across sites and species. Using a priori knowledge of the sensitivity of the model to the different litter decomposition parameters, we selected the following to include in the modified Monte Carlo: the relative maximum enzymatic decomposition rate for simple and chemically resistant SOM chemical classes ($V_{max,ref,simple/chemically\ resistant}$), activation energy for chemically resistant SOM ($Ea_{\text{chemically resistant}}$), C uptake efficiency for chemically resistant SOM ($\varepsilon_{\text{chemically resistant}}$), and the Michaelis-Menten constant for all three SOM chemical classes

Table 1
Baseline and Reparameterized Litter Decomposition Parameters in the CORPSE Model

Parameter	Description and units	Discrete parameter options	CORPSE baseline value	FUN-BioCROP baseline value	LIDET value
$V_{max,\ ref,\ fast}$	Relative maximum enzymatic decomposition rate, simple chemical class, ($year^{-1}$)	0.5, 1, 2, 2.5, 2.75, 3,3.25, 3.5, 4, 4.5, 5, 7	33.0	9.0	2.5
$V_{max,\ ref,\ slow}$	Relative maximum enzymatic decomposition rate, chemically resistant chemical class, ($year^{-1}$)	0.1, 0.2, 0.25, 0.3,0.35, 0.4, 0.5, 0.6	0.6	0.25	0.35
Ea_{slow}	Activation energy, chemically resistant chemical class, (J/mol)	2.5e4, 3e4, 3.25e4,3.5e4, 3.75e4, 4e4, 4.5e4	3.0e4	3.0e4	3.25e4
ε_{slow}	Carbon uptake efficiency, chemically resistant chemical class, (Dimensionless)	0.001, 0.01, 0.02, 0.04,0.6, 0.08, 0.10, 0.12	0.1	0.1	0.1
$k_{C,\ fast}$	Michealis-Menten parameter, simple chemical class (Dimensionless)	0.001, 0.005, 0.007, 0.009,0.01, 0.011, 0.013, 0.015, 0.02	0.01	0.01	0.007
$k_{C,\ slow}$	Michealis-Menten parameter, chemically resistant chemical class, (Dimensionless)	0.001, 0.005, 0.007, 0.009,0.01, 0.011, 0.013, 0.015, 0.02	0.01	0.01	0.007
$k_{C,\ necro}$	Michealis-Menten parameter, necromass chemical class (Dimensionless)	0.007, 0.009,0.01, 0.011, 0.013	0.01	0.01	0.009

Note. Reparameterized values were derived through a modified Monte Carlo approach and the LIDET dataset. Discrete parameter options were included in the modified Monte Carlo, the CORPSE and FUN-BioCROP baseline values were the parameters in those models prior to reparameterization, and the LIDET value was the parameter selected through reparameterization. CORPSE: Carbon, Organisms, Rhizosphere, and Protection in the Soil Environment; LIDET: Long-Term Inter-site Decomposition Experiment Team; FUN-BioCROP: Fixation and Uptake of Nitrogen-Bioenergy Carbon, Rhizosphere, Organisms and Protection.

($k_{C,\ simple,\ chemically\ resistant,\ necro}$). If a priori knowledge of model sensitivity to different parameters is unknown, a simple sensitivity analysis can help identify key parameters to include in the modified Monte Carlo simulation.

In the model, the maximum enzymatic decomposition rate ($V_{max,i}$) represents the fastest possible speed of decomposition of unprotected C and N for the three OM chemical classes (*i;* simple, chemically resistant, and microbial necromass). $V_{max,i}$ is determined by the Arrhenius equation, in which the relative maximum enzymatic decomposition rate for each OM chemical class ($V_{max,ref,i}$, included in the modified Monte Carlo for simple and chemically resistant OM) is modified by the activation energy for each OM chemical class ($E_{a,i}$, included in the modified Monte Carlo for chemically resistant OM) and temperature (T):

$$V_{max,i}(T) = V_{max,ref,i} \times \exp\left(-\frac{E_{a,i}}{RT}\right) \quad (1)$$

where R is the ideal gas constant (8.31 J K^{-1} mol^{-1}). In Equation 1, increasing the $E_{a,i}$ decreases $V_{max,i}$, while increasing T increases the $V_{max,i}$, and thus the speed of decomposition.

The decomposition flux ($D_{C,i}$ and $D_{N,i}$) is determined by $V_{max,i}$ (from Equation 1), temperature (T), volumetric soil water content (θ), and the ratio of microbial biomass carbon (C_M) to total substrate carbon ($C_{U,i}$), modified by the Michaelis-Menten constant ($k_{C,i}$ included in the modified Monte Carlo for all OM chemical classes):

$$D_{C,i} = V_{max,i}(T) \cdot \left(\frac{\theta}{\theta_{sat}}\right)^3 \left(1 - \frac{\theta}{\theta_{sat}}\right)^{2.5} \cdot C_{U,i} \frac{C_M/C_{U,i}}{\frac{C_M}{C_{U,i}} + k_{C,i}} \quad (2)$$

$$D_{N,i} = V_{max,i}(T) \cdot \left(\frac{\theta}{\theta_{sat}}\right)^3 \left(1 - \frac{\theta}{\theta_{sat}}\right)^{2.5} \cdot N_{U,i} \frac{C_M/C_{U,i}}{\frac{C_M}{C_{U,i}} + k_{C,i}} \quad (3)$$

where θ_{sat} is the saturation level of θ. As $k_{C,i}$ increases (i.e., enzyme affinity decreases), greater amounts of substrate are needed to achieve rapid enzymatic decomposition.

Decomposed organic matter is taken up by microbes in the model leading to microbial growth (G_M). The amount of $D_{C,i}$ that becomes microbial biomass is modified by microbial C uptake efficiency ($\epsilon_{C,i}$, included in the

modified Monte Carlo for chemically resistant OM), with higher $\epsilon_{C,i}$ leading to more accumulation of microbial biomass:

$$G_M = \sum_i (\epsilon_{C,i} D_{C,i}) \quad (4)$$

The complete description of model equations and parameters is detailed by Sulman et al. (2014a, 2014b) and is available on GitHub (Juice et al., 2023a).

Rather than assigning each parameter a distribution from which to randomly sample, as in a typical Monte Carlo simulation, we defined a set of discrete values for each parameter (Table 1). To determine the potential parameter space, we first conducted individual parameter analyses in which we ran the model changing only one parameter at a time. For each parameter included in the analysis, we ran an even sequence of 20 possible values that ranged ±300% of the Baseline parameter value. We then compared modeled C loss to field observations to identify the optimal individual parameter value, which we used as the center value for the Monte Carlo multi-parameter analysis. These methods are consistent with Monte Carlo approaches which assume normality of model residuals, not of parameter distributions themselves. Parameters with more variable impact on the model were given a broader range of potential values, such that the parameter spaces differed in number for each parameter (Table 1). We then randomly sampled from the defined sets of discrete values for each parameter to make 140 different combinations of parameter values. We simulated each LIDET litterbag (i.e., site-species decomposition curve) with each of the 140 parameter combinations. To identify the best parameter set and to ensure that certain sites did not have undue weight on the parameterization, we then performed 50 simulations where we randomly selected 70% of the data and identified which parameter set performed best. We assessed each parameter set according to its deviation from the line of equality (i.e., the root mean square error, RMSE) between modeled and observed mass loss for each species at each site. This effort revealed nine total parameter sets that improved model performance (Table S4 in Supporting Information S1), and from those nine we selected the one set of parameters that most frequently increased model accuracy (i.e., decreased RMSE) across each randomly selected dataset (Table 1). Here, we refer to that set as the "LIDET" parameters.

The model runs for this analysis were executed using Amazon Elastic Compute Cloud (Amazon EC2, zone us-east-1a, Amazon Web Services 2020). A memory optimized on-demand instance with 8 processors was selected where one processor allowed data to be read or written and seven processors were dedicated to running the models.

2.4. Comparing LIDET Parameters Versus Baseline Parameters

We compared model performance using LIDET parameters versus Baseline parameters for all sites and species by assessing several metrics used to characterize decomposition: point-in-time estimates of percent C remaining, decomposition rate (k) of each site-species litter combination, and the mean residence time (MRT) of the litter. Examination of the LIDET data revealed that the percent C of the litter was conserved as the mass declined, such that mass remaining LIDET data was analogous to model C remaining. We then examined the individual C and N pools in the litterbag layer to understand the varying model behavior under the different parameter sets. Finally, we compared the modeled N release patterns during decomposition C loss to the empirical data to evaluate the ability of the model to capture documented trends in N immobilization and release (as in Bonan et al., 2013; Kyker-Snowman et al., 2020; Parton et al., 2007). For all comparisons, we included only the six standard species that were decomposed at all LIDET sites, not the wildcard species that were decomposed at a subset of sites.

First, we compared modeled estimates of C remaining to field observations to assess the model's ability to capture individual points in time. We calculated the total litter C modeled with both Baseline and LIDET parameters as the sum of all the C pools (unprotected simple, chemically resistant and microbial necromass, and microbial biomass C) in the litter compartment. Then, we performed a linear least squares regression of C remaining at each collection time as estimated by CORPSE using both Baseline parameters and LIDET parameters against observed C remaining. We then compared the two fits using the coefficient of determination (R^2) of each regression line in addition to the RMSE and bias calculated by the Metrics package in R (Hamner & Frasco, 2018). We also grouped the C remaining data by biome and used the RMSE, R^2, and bias to evaluate if model performance varied by environmental characteristics that differ across biomes, as has been previously found (Bonan et al., 2013). Finally,

we examined the improvement in modeled C remaining of drypetes (*Drypetes glauca,* DRGL) across biomes because it is the LIDET species with the closest leaf fast fraction (0.779, Table S2 in Supporting Information S1) to the bioenergy corn-corn-soybean rotation (corn leaf fast fraction: 0.790, soybean leaf fast fraction: 0.815, Table S5 in Supporting Information S1) and *Miscanthus* (0.808, Table S5 in Supporting Information S1) and can thus help interpretation of the bioenergy soil C results.

We also examined the ability of the Baseline and LIDET parameters to replicate the decomposition curves observed in the LIDET data for each site-species combination. Here, we calculated the k of each site-species curve for the field observations, Baseline model estimates, and LIDET model estimates based on the negative exponential decomposition equation (Olson, 1963):

$$\ln\left(\frac{M_t}{M_0}\right) = -kt \tag{5}$$

where M_0 is the initial mass, M_t is the mass at time t, and k is calculated as the slope of the regression between $ln\left(\frac{M_t}{M_0}\right)$ and time. We then performed linear least squares regression of the modeled k versus the observed k for both the Baseline and LIDET estimates. We compared the fits using the R^2 of each regression line, the RMSE, and bias of the predicted versus the observed ks for both sets of parameters (using the R Metrics package, Hamner & Frasco, 2018).

Additionally, we compared the ability of the Baseline parameters versus LIDET parameters to estimate the average amount of time litter endured in the litterbag to that observed in the LIDET data. To do this, we first fit each site-species decomposition curve with a Weibull model that characterizes litter decomposition as a distribution of residence times described by shape and scale parameters, with the decomposition rate changing continuously (Cornwell & Weedon, 2014; Gill et al., 2021; Weibull, 1951). We then calculated mean residence time (MRT) with parameters derived from the Weibull fit:

$$MRT = \beta\Gamma\left(1 + \frac{1}{\alpha}\right) \tag{6}$$

where β describes the scale of the Weibull curve, Γ is the gamma function, and α describes the shape (Gill et al., 2021). We used the R code developed by Gill et al. (2020) for both the Weibull curve fitting and the MRT calculation. Lastly, we performed linear least squares regression of the modeled MRT versus the observed MRT for both the Baseline and LIDET estimates and compared the fits using the R^2 of each line. We also calculated the RMSE and bias for the modeled values predicted by the Baseline and LIDET parameters versus the field observations of mass remaining (with the R Metrics package, Hamner & Frasco, 2018).

Finally, to evaluate the ability of the model to document known patterns of N release during decomposition mass loss, we plotted the fraction of N remaining versus the percent C remaining for the data modeled by both parameter sets and compared the patterns to the LIDET field data. We examined the relationship between initial leaf N content, the level of N immobilization estimated by each parameter set, and the percent mass lost when overall N release began (Parton et al., 2007).

2.5. Evaluating Impacts of LIDET Parameters on Soil C Pools in Bioenergy Systems

To examine the impact of the improved litter decomposition parameters on estimates of soil C stocks, we used the LIDET parameters in the litter layer of the FUN-BioCROP model to simulate soil C and N in bioenergy agriculture (Juice et al., 2023b). We selected four feedstock scenarios that were previously simulated by FUN-BioCROP at the bioenergy field experiment at the University of Illinois at Urbana-Champaign (UIUC) Energy Farm located in Urbana, IL, USA (see Juice et al., 2022 for full model description and parameters). The feedstocks in that study included a corn-corn-soybean rotation (*Zea mays* and *Glycine max*), *Miscanthus* (*Miscanthus x giganteus*), and switchgrass (*Panicum virgatum*). Feedstocks were managed in accordance with regional standard practice (corn-corn-soybean) or best-known management practices (*Miscanthus* and switchgrass; Zeri et al., 2011). This included annual addition of N fertilizer to all feedstocks, although *Miscanthus* was additionally grown and modeled without fertilizer application, resulting in four feedstock model scenarios: corn-corn-

soybean, fertilized *Miscanthus*, unfertilized *Miscanthus*, and switchgrass. In the model, *Miscanthus* (fertilized and unfertilized) is replanted every 13 years as standard practice due to age related yield declines. Tillage is only simulated for *Miscanthus* prior to replanting, while corn-corn-soybean is tilled annually, and switchgrass is only tilled prior to its initial planting. Because switchgrass did not exhibit age related declines in yield, it was not replanted in the model. Harvests occur annually in the model, and remove 85% of perennial litter production and 90% of corn-corn-soybean.

For the current study, we spun up FUN-BioCROP using both the Baseline and LIDET parameters under a tall grass prairie plant community (following the description of the Konza Praire plant community by Bark, 1987) for 3,000 years (Figure S3 in Supporting Information S1; Juice et al., 2022). After spin up, we simulated a 160-year historical trajectory of agricultural development in the Great Plains under both parameter sets (Figure S3 in Supporting Information S1). In that simulation, crop varieties, yields, and management techniques (e.g., fertilization rates, tillage types) varied in accordance with the historical record (Hartman et al., 2011). The modeled SOC loss due to conversion of prairie to agricultural land fell within the range of empirical estimates (Juice et al., 2022). Soil C stocks in 2008 and the ratio of protected to unprotected soil C estimated by the model also fall within the range of the validation data from the UIUC Energy Farm bioenergy experimental plots (Figure S3 in Supporting Information S1, Juice et al., 2022; Ridgeway et al., 2022; Ilsa Kantola, unpublished data). Full details of the FUN-BioCROP spin up, historical simulation, and validation are described by Juice et al. (2022).

Models of soil C in bioenergy agriculture were then run using the Baseline FUN-BioCROP parameters, the LIDET parameters, and the eight other parameter sets that improved model performance as identified in the Monte Carlo analysis (Table S4 in Supporting Information S1). Bioenergy simulations were run from 2008 to 2100 using feedstock-specific parameters (Table S5 in Supporting Information S1), and data and agricultural management schedules from the UIUC Energy Farm. We held soil temperature and moisture constant by using the average of the field measurements from 2008 to 2016, thereby allowing us to explore the effect of bioenergy feedstocks on soil C without interference due to micrometeorological differences. Estimates of above and belowground plant biomass and productivity used to run FUN-BioCROP for the four bioenergy crop scenarios were generated using the DayCent-CABBI model (Berardi et al., 2020; Hartman et al., 2022; Moore et al., 2020; Parton et al., 1998). We evaluated the difference in estimates of soil C under the four bioenergy crops that were generated by the LIDET parameter set, the eight additional parameter sets from the Monte Carlo simulation, and the Baseline decomposition parameters in FUN-BioCROP originating from the CORPSE model (Table 1; Juice et al., 2022; Sulman et al., 2014a; Sulman et al., 2019; Wieder et al., 2019).

2.6. Sensitivity of Soil C Pools to Litter Quality and Quantity

We conducted subsequent bioenergy soil C simulations (2008–2100) in a sensitivity analysis that varied the amount of plant residue left after harvest, leaf litter quality, and leaf litter quantity to examine the cause of the differences in soil C estimated by the Baseline versus LIDET litter decomposition parameters. We tested the effect of two residue levels (15% and 50%) on soil C stocks in all four bioenergy feedstock scenarios. Because all perennial feedstocks (fertilized and unfertilized *Miscanthus*, and switchgrass) performed similarly to each other under both parameter sets (and differently from the corn-corn-soybean rotation), we selected only fertilized *Miscanthus* and corn-corn-soybean for the sensitivity analysis of litter quantity and quality. To test how the different parameter sets interacted with differences in litter quality, we varied the fast fraction of decomposing leaf litter (Table S5 in Supporting Information S1) by ±20% for both fertilized *Miscanthus* and corn-corn-soybean. Similarly, to test the effect of changing the quantity of litter inputs under both Baseline and LIDET parameters, we varied the aboveground litter production ±20% for both fertilized *Miscanthus* and corn-corn-soybean.

3. Results

Reparameterizing litter decomposition in FUN-BioCROP using a modified Monte Carlo approach improved model estimates of litter C remaining, decomposition rates (k), and litter mean residence time (MRT; Figure 2). Using Baseline parameters, the modeled output of litter C remaining captured 45% of the variability in the field data (Figure 2a). This increased to 54% with the improved version, which also improved the RMSE and bias as compared to the Baseline parameters (Figure 2b). Similarly, decomposition rates calculated from Baseline model output captured 41% of the variability in the field data (Figure 2c), which increased to 48% by using parameters

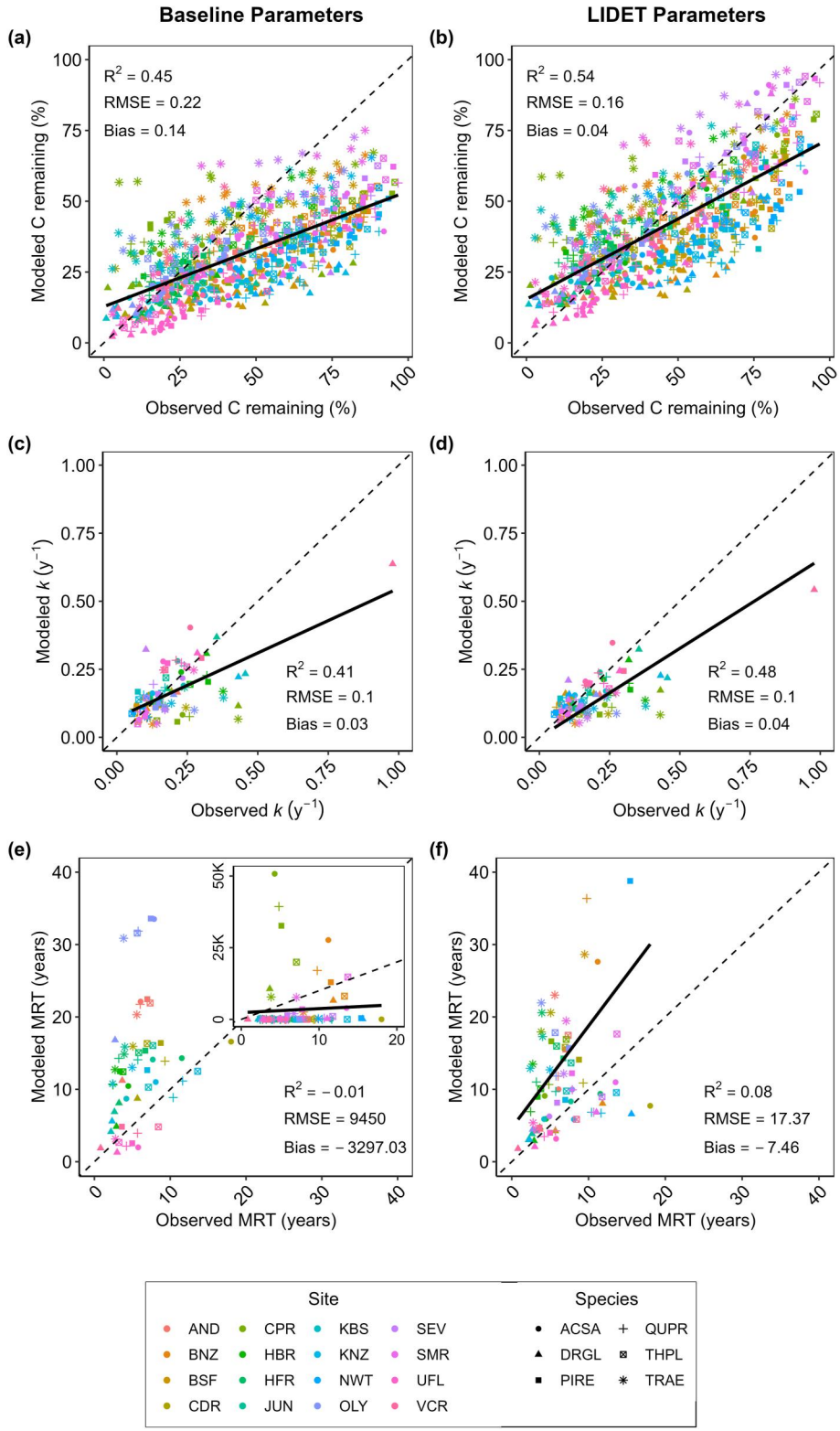

Figure 2.

derived from LIDET data (Figure 2d). Finally, Baseline model estimates of litter MRT effectively didn't capture the variability in the field data ($R^2 = -0.01$; Figure 2e). Baseline MRTs included multiple extreme values (Figure 2e inset) due to overestimation of early decomposition and underestimation of late-stage decomposition that led to very shallow decomposition curves and extremely high MRT estimates. Following reparameterization, calculated MRTs were improved and captured 8% of the variability in the field data with drastically reduced RMSE and bias (Figure 2f). Using the LIDET parameters reduced, but did not eliminate, the pattern of overestimating early decomposition and underestimating late decomposition, resulting in a smaller improvement in R^2 for MRT than that observed in the other decomposition metrics (i.e., C remaining and k).

Examination of the litterbag SOM pools revealed differences in microbial activity levels that altered patterns of C and N decomposition under the Baseline versus LIDET parameter sets. Because all sites exhibited comparable trends, we present detailed data from one representative site (CPR, Central Plains Experimental Range) that illustrates the differences between the parameter sets (Figure 3). In the Baseline CORPSE model, the higher decomposition parameters (see Table 1) caused the microbial biomass to peak early in decomposition followed by a population crash. The simple C and N pools were quickly depleted, and decomposition of the chemically resistant C and N pools was too fast. Using the LIDET parameters reduced and lengthened the peak of microbial biomass growth, extended the time required to deplete the simple C and N pools, and slowed decomposition of the chemically resistant pools. Collectively, these changes brought total litterbag C pools more in line with LIDET observations.

Patterns of N immobilization and release varied under the Baseline versus LIDET parameters (Figure 4). Using the LIDET parameters, very little N immobilization occurred in the highest concentration N litter, and net N release occurred earlier during mass loss (Figure 4, DRGL, 1.97% N, <50% mass lost at time of N release). The most N immobilization occurred in the lowest concentration N litters (Figure 4, TRAE, 0.38% N; PIRE, 0.59% N), which also experienced later net N release (>60% C loss at time of net N release, Figure 4). The remaining litters (QUPR, 1.03% N; ACSA, 0.81% N; and THPL, 0.62% N) displayed intermediate levels of N immobilization and timing of N release relative to mass loss. Conversely, the Baseline parameters generally overestimated N immobilization, and predicted late N release in the highest concentration N litter (DRGL).

The level of improvement in modeled mass remaining estimates using the LIDET parameters versus the Baseline parameters varied across biomes (Figure S1 in Supporting Information S1). As assessed by RMSE and bias, LIDET parameters improved model estimates in all biomes except deciduous forests. In deciduous forests, the LIDET parameters resulted in the same RMSE and slightly higher bias than the Baseline parameters. As assessed by R^2, LIDET parameters improved model estimates in all biomes to varying degrees. The largest improvement in modeled mass remaining estimates due to the use of the LIDET litter decomposition parameters was observed in arid grasslands (0.43 increase in R^2) and chaparral (0.34 increase in R^2). These sites had among the lowest R^2 before the reparameterization (arid grassland: 0.19, chaparral: 0.19). The smallest improvements in modeled mass remaining estimates as assessed by R^2 were observed at the saltmarsh biome (0.01 increase in R^2) and in the agricultural biome (0.03 increase in R^2). As the LIDET litter with the leaf fast fraction closest to that of bioenergy feedstocks corn-corn-soybean and *Miscanthus*, modeled C remaining of drypetes (*Drypetes glauca*) improved across all biomes as assessed by RMSE and bias, with variable response in the R^2 value (Figure S2 in Supporting Information S1).

The impact of applying the LIDET litter decomposition parameters in FUN-BioCROP to soil C simulations of bioenergy crops at the UIUC Energy Farm was consistent across feedstocks (Figure 5). The saw-tooth pattern of modeled soil C observed in *Miscanthus* was caused by its 13-year planting rotation, which caused a large, periodic influx of belowground litter and a spike in total estimated soil carbon. At the end of the 93-year simulation, model estimates of total mean annual soil C (i.e., the sum of all C pools in all model compartments–litter, tilled and untilled rhizosphere and bulk soil) were greater from the model run with LIDET parameters as compared to output generated with the Baseline parameters (corn-corn-soybean: 7.8%, fertilized *Miscanthus:* 8.3%, unfertilized

Figure 2. Comparison of model estimates of litter C remaining (a) and (b), decomposition rate (k; c and d) and litter C mean residence time (e) and (f) calculated using the Baseline parameters (left column) and the LIDET parameters (right column) that were selected through a modified Monte Carlo parameterization of the CORPSE model using the temporally and spatially extensive LIDET dataset. Model estimates of MRT generated with Baseline parameters included multiple extreme values as shown in the inset of panel (e). The dashed line shows the 1:1 line. C: carbon; CORPSE: Carbon, Organisms, Rhizosphere, and Protection in the Soil Environment; LIDET: Long-term Inter-site Decomposition Experiment Team; MRT: mean residence time.

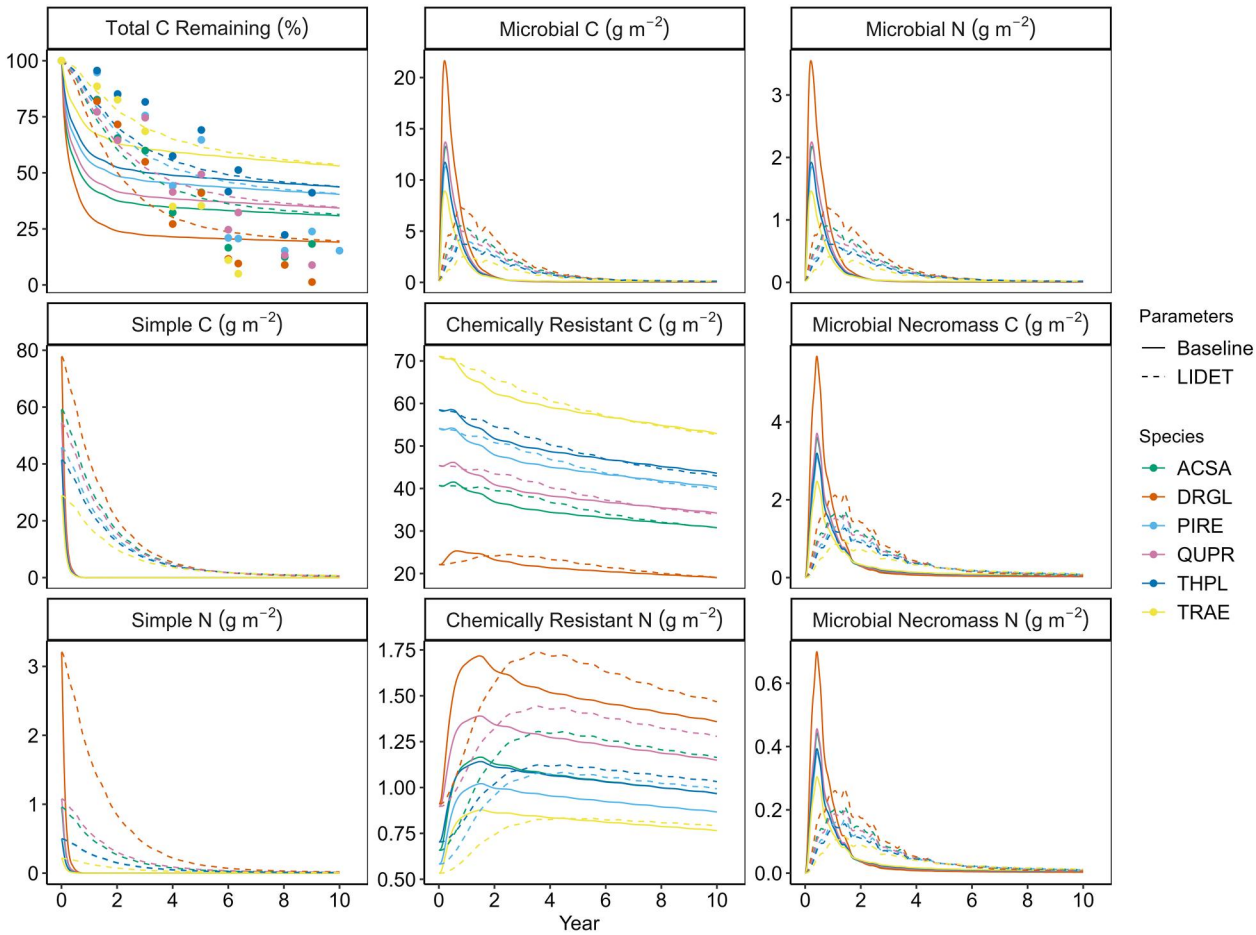

Figure 3. Total C remaining and individual C and N model pools over the course of decomposition estimated with the Baseline (solid lines) and LIDET parameters (dashed lines) at the Central Plains Experimental Range site in the LIDET experiment. The data points in the total C remaining panel show the LIDET field data, all other panels show modeled data. C: carbon; LIDET: Long-term Inter-site Decomposition Experiment Team.

Miscanthus: 8.2%, switchgrass: 7.8%). Additionally, the LIDET parameters and the eight other parameter sets selected in the Monte Carlo analysis all resulted in similar model outcomes (Figure 5, gray and red lines).

Increasing the modeled amount of residue left on the field after harvest from 15% to 50% greatly increased soil C stocks in the perennial feedstocks, but not the annual corn-corn-soybean rotation. For both fertilized and unfertilized M*iscanthus*, using the LIDET parameters and the high residue scenario resulted in higher estimated soil C stocks than with the Baseline parameters (fertilized *Miscanthus:* 27% vs. 21%; unfertilized *Miscanthus:* 24% vs. 18%). Conversely, in the switchgrass high residue scenario, the two parameter sets resulted in similar increases in SOC stocks (34% with Baseline vs. 36% with LIDET parameters).

In the litter quantity and quality sensitivity analysis, model estimates of soil C for corn-corn-soybean were slightly more sensitive to changing the quality of litter inputs (i.e., the fast fraction) than changing the quantity of those inputs (Figures S4a and S4b in Supporting Information S1). Conversely, modeled soil C stocks in fertilized *Miscanthus* were more sensitive to changing litter quantity than quality (Figures S4a and S4b in Supporting Information S1). For both feedstocks, litter inputs correlated positively with the size of the total soil C stocks at the end of the simulation, and litter quality was inversely related to total soil C (i.e., a higher proportion of fast-decomposing litter led to lower soil C stocks).

Fertilized *Miscanthus* and the corn-corn-soybean rotation responded similarly to changes in leaf litter chemistry (Figure S4a in Supporting Information S1), perhaps due to the similarity in their leaf litter fast fractions (Table S5 in Supporting Information S1). Increasing the fast fraction of the litter by 20% decreased total soil C for both feedstocks over the 93-year simulation (corn-corn-soybean: −1.6%, fertilized *Miscanthus*: −1.5%). Conversely,

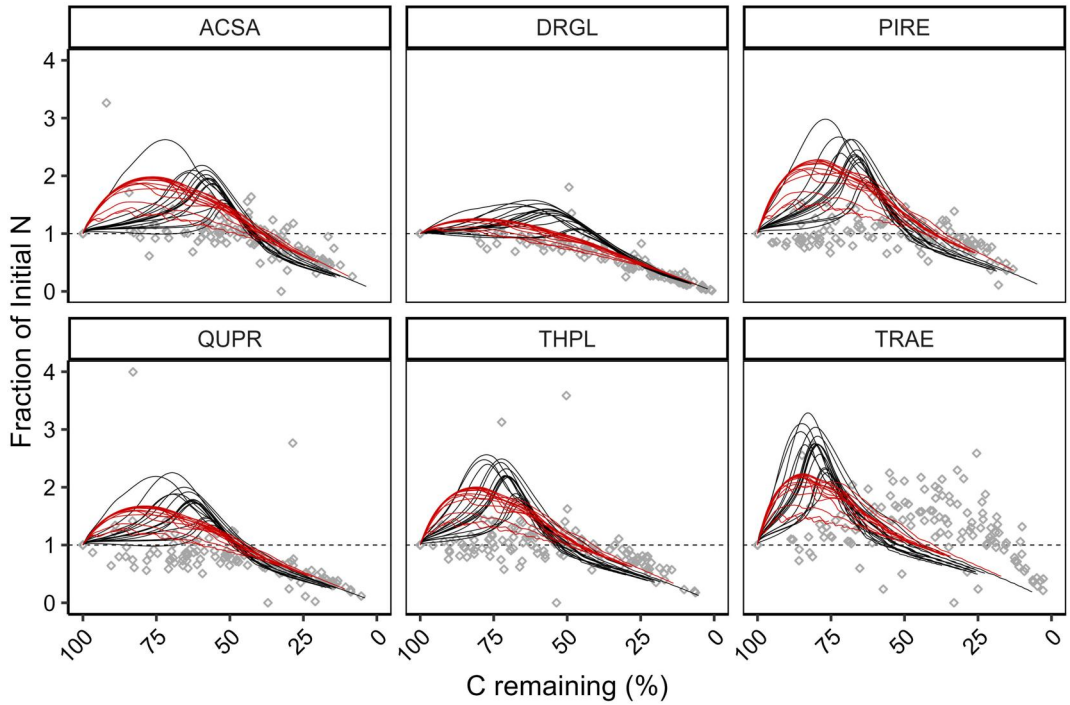

Figure 4. Fraction of initial litter N remaining as a function of percent C remaining during decomposition of the six standard litters decomposed at the LIDET sites. All points greater than the dashed horizontal lines indicate net N immobilization. The Baseline parameters (black lines) generally overestimated the magnitude and timing of N immobilization as compared to the LIDET parameters (red lines). The gray points show the LIDET observations. C: carbon; LIDET: Long-term Intersite Decomposition Experiment Team; N: nitrogen.

decreasing the fast fraction of litter increased total soil C over the 93-year period (corn-corn-soybean: +1.6%, fertilized *Miscanthus*: +1.3%).

Fertilized *Miscanthus* was more responsive to changing the quantity of inputs to the soil than was corn-corn-soybean (Figure S4b in Supporting Information S1). For fertilized *Miscanthus*, increasing leaf litter C inputs by 20% resulted in a 2.3% increase in soil C stocks at the end of the simulation. Reducing leaf litter C inputs to fertilized *Miscanthus* reduced the total soil C stocks by 2.5% by the end of the simulation. The effect was smaller in corn-corn soybean, where increasing leaf litter C input to the soil resulted in a 1.2% increase in soil C at the end of the simulation. Over the same period in corn-corn-soybean, reducing leaf litter C input by 20% decreased soil C stocks by 1.5%.

4. Discussion

Our results indicate that large-scale datasets can be used to systematically parameterize microbial models while avoiding the difficulties of full Monte Carlo approaches or the pitfalls of manual calibration. There have been multiple efforts to use data assimilation to improve the parameterization of process-based models (De Kauwe et al., 2014; Keenan et al., 2013; Ma et al., 2022). While the ecosystem modeling community has recognized that these data assimilation techniques and approaches are robust, the computational time and technical expertise needed to implement these approaches has been a barrier to their widespread adoption. As such, much of the community still relies on calibration techniques where parameters are adjusted manually until the model output of choice matches a certain benchmark. By forging a middle ground between these approaches, our framework provides a robust parameterization that is simple to implement and improves predictions of litter decomposition in the CORPSE model across three different commonly used decomposition metrics (Figure 2).

The improvement in CORPSE model performance with the new parameters reflects the integrated impact of decreasing the V_{max} parameters for the decomposition of simple litter components (Table 1; $V_{max,ref,simple}$; Baseline: 33 years^{-1}, LIDET: 2.5 years^{-1}) and chemically resistant litter components ($V_{max,ref,chemically\ resistant}$;

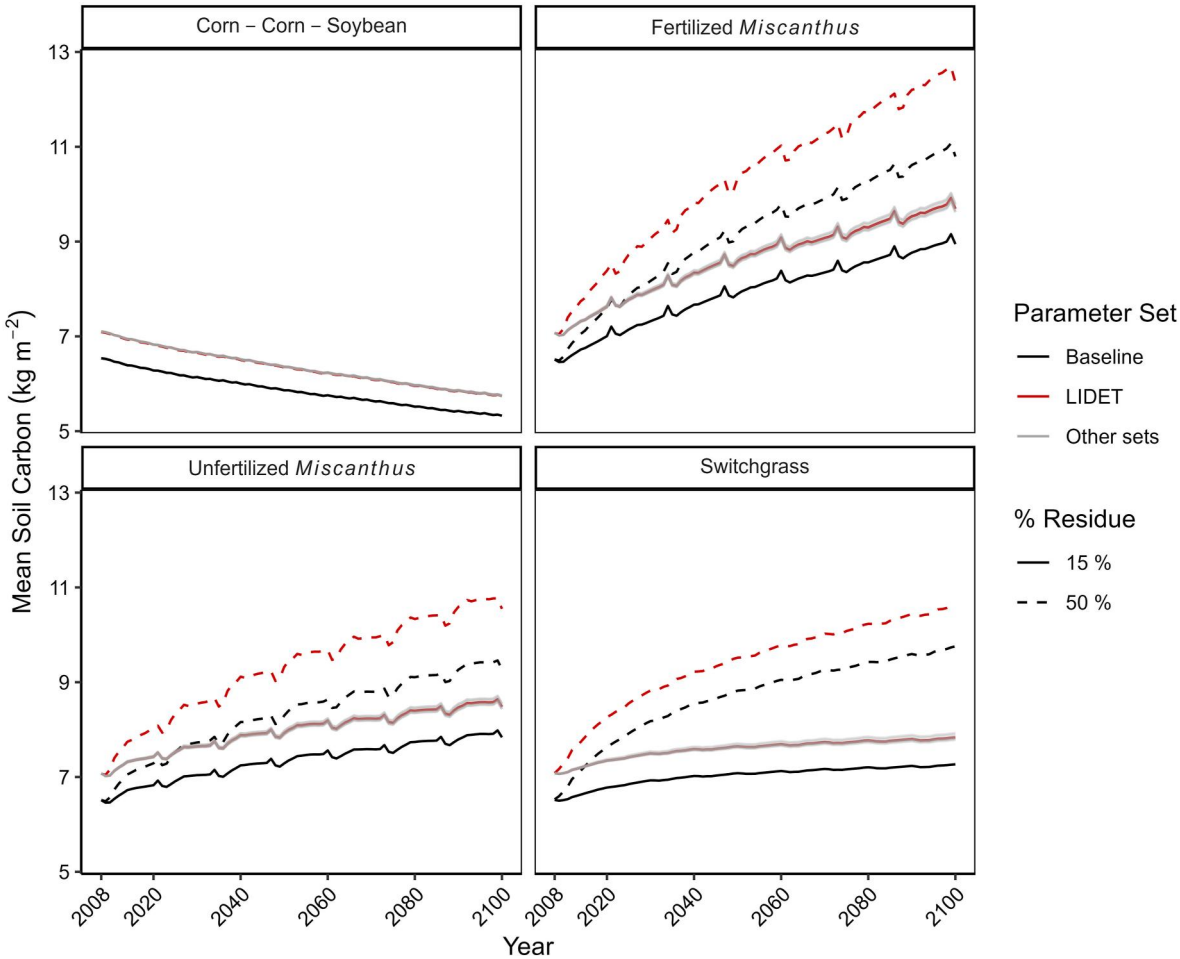

Figure 5. Model estimates of soil C generated for four bioenergy feedstock scenarios by the FUN-BioCROP model using Baseline parameters (black lines), the new LIDET parameters (red lines), and the eight best parameter sets (gray lines) that were selected through a modified Monte Carlo parameterization of the CORPSE model using the temporally and spatially extensive LIDET dataset. FUN-BioCROP simulates the top 30 cm of the soil profile. Residue levels were varied in the model between 15% (solid lines) and 50% (dashed lines) to explore the interaction between model parameterizations and litter inputs. The same model simulations were run for all four feedstock scenarios. In the corn-corn-soybean panel, the 50% residue addition did not increase soil C stocks so the dashed lines for the 50% residue simulations overlap with the 15% residue simulations for each parameter set. C: carbon; CORPSE: Carbon, Organisms, Rhizosphere, and Protection in the Soil Environment; FUN-BioCROP: Fixation and Uptake of Nitrogen-Bioenergy Carbon, Rhizosphere, Organisms and Protection; LIDET: Long-term Inter-site Decomposition Experiment Team.

Baseline: 0.6 years^{-1}, LIDET: 0.35 years^{-1}). By significantly slowing down the initial rate of litter decomposition and, to a lesser extent, later stage decomposition, the new model was better able to capture the trajectory of mass loss over time. For example, at the Central Plains Experimental Range (CPR), the Baseline model clearly overpredicted the rate of litter decomposition, particularly in the early stages of litter decay (Figure 3). With the shift in the V_{max} rates and other parameters (i.e., activation energy of chemically resistant C, uptake efficiency of chemically resistant C, and the Michaelis-Menten parameter for all C classes), the overall trajectory of the curve was better captured by the model run with the LIDET parameters. Examination of the different C and N pools in the litterbag compartment (Figure 3) revealed that the LIDET parameters corrected a boom-bust cycle in the microbial biomass that resulted in overestimates of litter decomposition using the Baseline parameters. Under the new parameters, the simple C and N pools were depleted more slowly, and the chemically resistant C pool underwent slower decomposition in initial stages. But this examination of the model behavior in the individual C and N pools makes clear that the LIDET parameters correctly addressed the problem inherent to the Baseline parameter set: namely, too fast of decomposition rate parameters leading to rapid depletion of simple and chemically resistant C pools and too fast C loss estimates. However, since we parameterized a subset of the

decomposition parameters in the Monte Carlo analysis, it is possible that including all the parameters would have resulted in selection of different optimal values (i.e., an equifinality issue).

Notably, the LIDET parameters also changed the shape of the curve of chemically resistant N decomposition (Figure 3) with implications for N loss dynamics during decomposition C loss (Figure 4). With the new parameters, the results better reflect previously described patterns of N immobilization in the LIDET field data (Parton et al., 2007). Namely, leaf litter high in initial N exhibited little to no immobilization of N during decomposition. As initial N levels decreased, N immobilization increased, and net N release occurred at later stages of mass loss (Parton et al., 2007). With the new parameters, CORPSE approximated these patterns better than with the Baseline parameters, which generally overestimated immobilization in all the litters (Figure 4). Our results coupled with work from other microbial models (Kyker-Snowman et al., 2020) show that microbial models can capture these N immobilization patterns observed in the LIDET data and previously simulated by first-order models (Bonan et al., 2013).

While the reparameterized model performed well across the entire dataset, there were a limited number of biomes where the model did not capture the trajectory of litter decomposition (Figure S1 in Supporting Information S1). For example, despite the marked improvement in simulated decomposition at CPR, the model still underpredicted late-stage decomposition rates at that site (Figure 3). This underestimation of late-stage decomposition is common in arid sites where exposure to UV radiation can enhance decomposition rates beyond what is predicted by litter quality and climate factors alone (Adair et al., 2008). Modeled C loss in such arid ecosystems can be significantly improved by accounting for mass loss due to photodegradation (Chen et al., 2016), which is not currently simulated by the CORPSE model. Photodegradation can even accelerate mass loss in humid grasslands (Brandt et al., 2010), perhaps explaining the underprediction of decomposition by CORPSE in that biome as well (Figure S1 in Supporting Information S1). Additionally, in the Boreal Forest site the new parameters improved model outcomes but still underpredicted early-stage decomposition as compared to field data (Figure S1 in Supporting Information S1). The inability of the new parameters to represent mass loss in this biome may reflect its position as an endmember in the overall climate space captured in the LIDET dataset. These sites are at the edge of the range of soil moisture and suggest a clear need to improve the representation of soil moisture impacts on decomposition in the CORPSE model. Overall, the parameterization exercise improved the ability of the model to simulate the trajectory of litter decomposition in the majority of sites, but its limitations also diagnosed key areas where other model structures need to be improved.

The higher soil C stocks estimated by FUN-BioCROP using the LIDET litter decomposition parameters as compared to its Baseline parameters reflect overall slower decomposition caused by a reduction in multiple parameter values and their resulting impact on microbial decomposition (Table 1). The reduction in $V_{max,ref,\ simple}$ was particularly impactful because it decreased the decomposition rate of the simple (labile) C pool, which dominates the composition of the litter C inputs. Across all the bioenergy feedstocks, the fast-decomposing fraction (i.e., fast fraction) of their leaves comprised over 50% of the litter C in the model with the overall average fast fraction of feedstock leaves at 75% (Table S5 in Supporting Information S1). By contrast, the comparatively smaller increase in the $V_{max,\ ref,\ chemically\ resistant}$ parameter compared to the FUN-BioCROP Baseline value increased the decomposition rate of a much smaller proportion of the litter (the chemically resistant pool), an effect which was overshadowed by the much larger decrease in the decomposition rate of the significantly larger simple pool. There were also feedbacks of the new parameters on microbial biomass. For example, in corn-corn-soybean litter, the reduction in $V_{max,\ ref,\ simple}$ decreased the size of the microbial biomass pool, ultimately leading to lower respiratory losses of C and contributing to the increase in soil C calculated using the LIDET litter decomposition parameters (Figure S5 in Supporting Information S1). Finally, at each timestep the model mixes a set proportion of the litter compartment into the rhizosphere and bulk soil compartments to represent bioturbation and frost heaves. With the decreased decomposition occurring in the litter compartment, more of that undecomposed C was therefore transferred into the bulk and rhizosphere soil. Overall, the new parameters combined with plant chemistry altered microbial biomass and the magnitude of modeled C fluxes to affect soil C stores.

As regenerative agricultural practices and carbon farming seek to increase soil C storage in agroecosystems, with the target of increasing soil C by 4 per mil per year to offset anthropogenic greenhouse gas emissions (Torquebiau, 2017), it is notable that simply improving model parameterization led to an average estimated difference of 0.9 per mil per year in soil C as compared to the Baseline model. This represents nearly a quarter of the targeted

soil C increase, and highlights the need for improved C accounting through data-based parameterization of litter decomposition in microbial models.

The updated decomposition dynamics represented by the LIDET parameters interacted with plant traits and management activities to more strongly alter soil C estimates in perennial feedstocks as compared to the annual corn-corn-soybean rotation (Figure 5). Under default model conditions with the same residue removal (85% for all feedstocks) and similar harvest levels (85% in perennials and 90% in annuals as determined by regional practices), the effect of the LIDET parameters on increasing simulated soil C stocks was similar across feedstocks (~8% more soil C projected by LIDET vs. Baseline parameters in all four feedstocks). However, the diverging effect of the parameter sets on perennial versus annual feedstocks was illustrated by the simulations that increased litter inputs either by increasing residue left after harvest, or increasing aboveground litter overall. For example, both Baseline and LIDET parameters estimated significant increases in soil C for perennials but not annuals under elevated residue addition, but the increases predicted by the new LIDET parameters were greater than the Baseline parameters (~6% higher soil C projected by LIDET vs. Baseline parameters for fertilized and unfertilized *Miscanthus* and 2% higher for switchgrass). Leaving more crop residues after harvest is one regenerative farming method that seeks to increase soil C (Lal et al., 2004). Our results highlight that the ability to predict the outcome of implementing such a method relies on accurate parameterization of litter decomposition.

Model tests that increased plant productivity provided further evidence that the different parameter sets interacted with the quantity of plant inputs to influence projected soil C levels. First, N fertilization of *Miscanthus* in the model enhanced plant productivity which increased soil C stocks under both parameter sets, but with greater increases under the new LIDET parameters (Figures 5, 1% greater with LIDET vs. Baseline parameters). Similarly, in the sensitivity tests that increased leaf litter input quantity in fertilized *Miscanthus* and corn-corn-soybean (Figure S4 in Supporting Information S1), fertilized *Miscanthus* had twice the percent increase in soil C stocks as observed in corn-corn-soybean. Part of this greater estimated accumulation of soil C is due to *Miscanthus* having generally greater litter inputs to soils in the model owing to its higher net primary productivity compared to corn-corn-soybean. It can also be attributed to the lack of tillage in *Miscanthus*. Indeed, our result of increased residue having negligible effect on the corn-corn-soybean soil C coincides with empirical findings that residue that is incorporated into the soil, such as by tillage, is more quickly degraded than that which is left on the surface (Reicosky et al., 1997; Sandhu et al., 2022). Because the LIDET parameters slowed decomposition, these greater litter inputs and lack of tillage in *Miscanthus* interacted with the LIDET parameters to project more soil C by the model. In support, when we reduced litter inputs to the soils in the *Miscanthus* plots, it removed a portion of the increase we saw in soil C over time under the new parameters (Figure S4 in Supporting Information S1). By contrast, the corn-corn-soybean system was less sensitive to changes in litter magnitude due to overriding effects of tillage (Figure S4 in Supporting Information S1). Future model experiments should evaluate the relative effects of reducing tillage versus increasing litter inputs on soil C stocks to help identify best management practices for soil C accumulation in bioenergy agriculture (e.g., He et al., 2021).

While our simplified Monte-Carlo reparameterization improved the accuracy of litter C loss during decomposition estimated by the CORPSE model, it is worth noting several assumptions that shaped this modeling exercise. First, to simplify the Monte Carlo approach and make it more feasible for parameterization of microbial models, we constrained the parameter space rather than using full distributions for each parameter. Bounding potential parameter values in this way relies on prior knowledge of both ecosystem processes and model structures, which is feasible in most situations given modelers' familiarity with their study systems. However, incorrect parameterization could result if the prior information necessary to select an appropriate set of potential parameters is lacking, or if incorrect assumptions are made regarding reasonable parameter ranges. Our simplified Monte Carlo approach may therefore be most appropriate in cases, such as that presented here, where there is sufficient knowledge regarding parameter values derived from observations as well as of the key parameters that control model outputs to reasonably define a range of possible parameter values. Second, the LIDET experiment measured leaf decomposition in litterbags, which have some well-documented artifacts (e.g., invertebrate exclusion, microclimate and photodegradation effects, lack of litter fragmentation) that affect their decomposition rates (Bokhorst & Wardle, 2013; Cotrufo et al., 2015; De Santo et al., 1993; Kurz-Besson et al., 2005; Slade & Riutta, 2012). They also lack the ability to undergo tillage, a practice which greatly affects soil processes in agricultural systems. Third, our reparameterization of litter decomposition focused on leaf litter, leaving root litter estimates to be calculated with the Baseline model parameters. Because a large proportion of soil C is derived

from roots (Rasse et al., 2005), using the LIDET root decomposition data to update root litter decomposition parameters could further improve modeled estimates of decomposition. Finally, the improvement we show in projections of litter decomposition here are slightly less than previous efforts using the MIMICS model (Kyker-Snowman et al., 2020; Wieder et al., 2015). Addressing the limitations outlined here (e.g., soil moisture, root decomposition parameterization) may help close the gap in model performance.

5. Conclusion

The reparameterization methods outlined here can help improve model estimates of litter decomposition in microbial-based soil C models, with important implications for soil C projections. Operationalizing agro-ecosystem soil C storage as a natural climate solution to offset greenhouse gas emissions requires accurate soil C accounting from models that rely on data-based parameterizations such as the one we show here. By changing the rate of decomposition both in initial stages and end stages, the new LIDET parameters better approximated field observations of C loss and N release over the course of decomposition, and projected greater soil C accumulation over time in bioenergy soils (0.9 per mil per year more than Baseline parameters). The analysis also revealed future avenues of model development, as lack of photodegradation and soil moisture extremes characteristic of certain biomes limited the improvement in decomposition estimates resulting from the reparameterization. Applying the LIDET parameters to simulations of bioenergy feedstock soil C revealed interactions between the parameters, plant life history (annual vs. perennial), and management practices that related most strongly to the quantity of inputs, either as residue or increased productivity, and tillage in annuals that prevents accumulation of soil C over time. Overall, the reparameterization framework presented here provides proof of concept of decomposition model parameter improvement derived from a simplified Monte Carlo simulation and a dataset that spans a range of litter qualities, climatic conditions, and biomes.

Conflict of Interest

The authors declare no conflicts of interest relevant to this study.

Data Availability Statement

The data and software used in this research are publicly available. The LIDET dataset (Harmon, 2013) is available at https://andlter.forestry.oregonstate.edu/data/abstract.aspx?dbcode=TD023. The model and input data to run CORPSE at each LIDET site (Juice et al., 2023a) is preserved at https://doi.org/10.5281/zenodo.10214017 and available openly on GitHub (https://github.com/BrzostekEcologyLab/CORPSE-LIDET). The CORPSE-LIDET model can be run with the Baseline parameters, LIDET parameters, or any of the other of the eight best parameter sets identified in the modified Monte Carlo simulation. The FUN-BioCROP model (Juice et al., 2023b) with Baseline and LIDET litter decomposition parameters, as well as the eight other parameter sets, is preserved at https://www.doi.org/10.5281/zenodo.10162303 and openly available on GitHub at https://github.com/BrzostekEcologyLab/FUN-BioCROP-LIDET.

Acknowledgments
We thank Chris Walter for many helpful conversations about this data analysis. This work was funded by the DOE Center for Advanced Bioenergy and Bioproducts Innovation (U.S. Department of Energy, Office of Science, Biological and Environmental Research Program under Award Number DE-SC0018420). Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the views of the U.S. Department of Energy. This material is based upon work supported by the National Science Foundation Graduate Research Fellowship Program under Grant DGE-1102689. This research was supported in part by the Intramural Research Program of the U.S. Department of Agriculture, National Institute of Food and Agriculture, Predoctoral Fellowship Program (Grant 2021-67034-35046). The findings and conclusions in this publication have not been formally disseminated by the U.S. Department of Agriculture and should not be construed to represent any agency determination or policy. Data were provided by the HJ Andrews Experimental Forest research program, funded by the National Science Foundation's Long-Term Ecological Research Program (DEB 1440409), US Forest Service Pacific Northwest Research Station, and Oregon State University.

References

Abramoff, R., Xu, X., Hartman, M., O'Brien, S., Feng, W., Davidson, E., et al. (2018). The Millennial model: In search of measurable pools and transformations for modeling soil carbon in the new century. *Biogeochemistry*, *137*(1), 51–71. https://doi.org/10.1007/s10533-017-0409-7

Adair, E. C., Parton, W. J., Del Grosso, S. J., Silver, W. L., Harmon, M. E., Hall, S. A., et al. (2008). Simple three-pool model accurately describes patterns of long-term litter decomposition in diverse climates. *Global Change Biology*, *14*(11), 2636–2660. https://doi.org/10.1111/j.1365-2486.2008.01674.x

Bakker, M. A., Carreño-Rocabado, G., & Poorter, L. (2011). Leaf economics traits predict litter decomposition of tropical plants and differ among land use types. *Functional Ecology*, *25*(3), 473–483. https://doi.org/10.1111/j.1365-2435.2010.01802.x

Bark, L. D. (1987). *Konza prairie research natural area, Kansas*. In D. Greenwood (Ed.), *The climates of the long-term ecological research sites* (pp. 45–50). University of Colorado: Institute for Arctic and Alpine research. occasional paper 44.

Berardi, D., Brzostek, E., Blanc-Betes, E., Davison, B., DeLucia, E. H., Hartman, M. D., et al. (2020). 21st-century biogeochemical modeling: Challenges for Century-based models and where do we go from here? *GCB Bioenergy*, *12*(10), 774–788. https://doi.org/10.1111/gcbb.12730

Bokhorst, S., & Wardle, D. A. (2013). Microclimate within litter bags of different mesh size: Implications for the 'arthropod effect' on litter decomposition. *Soil Biology and Biochemistry*, *58*, 147–152. https://doi.org/10.1016/j.soilbio.2012.12.001

Bonan, G. B., Hartman, M. D., Parton, W. J., & Wieder, W. R. (2013). Evaluating litter decomposition in earth system models with long-term litterbag experiments: An example using the Community Land Model version 4 (CLM4). *Global Change Biology*, *19*(3), 957–974. https://doi.org/10.1111/gcb.12031

Brandt, L. A., King, J. Y., Hobbie, S. E., Milchunas, D. G., & Sinsabaugh, R. L. (2010). The role of photodegradation in surface litter decomposition across a grassland ecosystem precipitation gradient. *Ecosystems*, *13*(5), 765–781. https://doi.org/10.1007/s10021-010-9353-2

Brzostek, E. R., Fisher, J. B., & Phillips, R. P. (2014). Modeling the carbon cost of plant nitrogen acquisition: Mycorrhizal trade-offs and multipath resistance uptake improve predictions of retranslocation. *Journal of Geophysical Research: Biogeosciences*, *119*(8), 1684–1697. https://doi.org/10.1002/2014jg002660

Carvalho, J. L. N., Hudiburg, T. W., Franco, H. C. J., & DeLucia, E. H. (2017). Contribution of above- and belowground bioenergy crop residues to soil carbon. *GCB Bioenergy*, *9*(8), 1333–1343. https://doi.org/10.1111/gcbb.12411

Chen, M., Parton, W. J., Adair, E. C., Asao, S., Hartman, M. D., & Gao, W. (2016). Simulation of the effects of photodecay on long-term litter decay using DayCent. *Ecosphere*, *7*(12), e01631. https://doi.org/10.1002/ecs2.1631

Cornwell, W. K., & Weedon, J. T. (2014). Decomposition trajectories of diverse litter types: A model selection analysis. *Methods in Ecology and Evolution*, *5*(2), 173–182. https://doi.org/10.1111/2041-210X.12138

Cotrufo, M. F., Soong, J. L., Horton, A. J., Campbell, E. E., Haddix, M. L., Wall, D. H., & Parton, W. J. (2015). Formation of soil organic matter via biochemical and physical pathways of litter mass loss. *Nature Geoscience*, *8*(10), 776–779. https://doi.org/10.1038/ngeo2520

Cotrufo, M. F., Wallenstein, M. D., Boot, C. M., Denef, K., & Paul, E. (2013). 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? *Global Change Biology*, *19*(4), 988–995. https://doi.org/10.1111/gcb.12113

De Kauwe, M. G., Medlyn, B. E., Zaehle, S., Walker, A. P., Dietze, M. C., Wang, Y.-P., et al. (2014). Where does the carbon go? A model–data intercomparison of vegetation carbon allocation and turnover processes at two temperate forest free-air CO_2 enrichment sites. *New Phytologist*, *203*(3), 883–899. https://doi.org/10.1111/nph.12847

De Santo, A. V., Berg, B., Rutigliano, F. A., Alfani, A., & Floretto, A. (1993). Factors regulating early-stage decomposition of needle litters in five different coniferous forests. *Soil Biology and Biochemistry*, *25*(10), 1423–1433. https://doi.org/10.1016/0038-0717(93)90057-I

Dyer, M. L., Meentemeyer, V., & Berg, B. (1990). Apparent controls of mass loss rate of leaf litter on a regional scale. *Scandinavian Journal of Forest Research*, *5*(1–4), 311–323. https://doi.org/10.1080/02827589009382615

Fisher, J. B., Sitch, S., Malhi, Y., Fisher, R. A., Huntingford, C., & Tan, S.-Y. (2010). Carbon cost of plant nitrogen acquisition: A mechanistic, globally applicable model of plant nitrogen uptake, retranslocation, and fixation. *Global Biogeochemical Cycles*, *24*(1). https://doi.org/10.1029/2009gb003621

Gholz, H. L., Wedin, D. A., Smitherman, S. M., Harmon, M. E., & Parton, W. J. (2000). Long-term dynamics of pine and hardwood litter in contrasting environments: Toward a global model of decomposition. *Global Change Biology*, *6*(7), 751–765. https://doi.org/10.1046/j.1365-2486.2000.00349.x

Gill, A. L., Schilling, J., & Hobbie, S. E. (2020). Litter decomposition meta-analysis [Source Code]. Zenodo. Retrieved from https://github.com/gill20a/LitterDecompMetaAnalysis.

Gill, A. L., Schilling, J., & Hobbie, S. E. (2021). Experimental nitrogen fertilisation globally accelerates, then slows decomposition of leaf litter. *Ecology Letters*, *24*(4), 802–811. https://doi.org/10.1111/ele.13700

Hamner, B., & Frasco, M. (2018). Metrics: Evaluation metrics for machine learning. R package version 0.1.4 [Software]. Retrieved from https://CRAN.R-project.org/package=Metrics

Harmon, M. (2013). LTER intersite fine litter decomposition experiment (LIDET), 1990 to 2002 [Dataset]. Long-Term Ecological Research. Forest Science Data Bank, Corvallis, OR. https://doi.org/10.6073/pasta/f35f56bea52d78b6a1ecf1952b4889c5

Hartman, M. D., Burnham, M., Parton, W. J., Finzi, A., Delucia, E. H., & Yang, W. H. (2022). In silico evaluation of plant nitrification suppression effects on agroecosystem nitrogen loss. *Ecosphere*, *13*(12), e4292. https://doi.org/10.1002/ecs2.4292

Hartman, M. D., Merchant, E. R., Parton, W. J., Gutmann, M. P., Lutz, S. M., & Williams, S. A. (2011). Impact of historical land-use changes on greenhouse gas exchange in the U.S. Great Plains, 1883–2003. *Ecological Applications*, *21*(4), 1105–1119. https://doi.org/10.1890/10-0036.1

He, W., Grant, B. B., Jing, Q., Lemke, R., St. Luce, M., Jiang, R., et al. (2021). Measuring and modeling soil carbon sequestration under diverse cropping systems in the semiarid prairies of western Canada. *Journal of Cleaner Production*, *328*, 129614. https://doi.org/10.1016/j.jclepro.2021.129614

Jastrow, J. D., & Miller, R. M. (1997). Soil aggregate stabilization and carbon sequestration: Feedbacks through organomineral associations. In R. Lal, J. M. Kimble, R. F. Follett, & B. A. Stewart (Eds.), *Soil processes and the carbon cycle* (pp. 207–223). CRC press.

Johnson, J. M.-F., Barbour, N. W., & Weyers, S. L. (2007). Chemical composition of crop biomass impacts its decomposition. *Soil Science Society of America Journal*, *71*(1), 155–162. https://doi.org/10.2136/sssaj2005.0419

Juice, S. M., Ridgeway, J. R., Hartman, M. D., Parton, W. J., Berardi, D. M., Sulman, B. N., et al. (2023a). CORPSE model with litter decomposition parameters derived from the LIDET dataset (v2.0.0) [Source Code]. https://doi.org/10.5281/zenodo.10214017

Juice, S. M., Ridgeway, J. R., Hartman, M. D., Parton, W. J., Berardi, D. M., Sulman, B. N., et al. (2023b). FUN-BioCROP model with litter decomposition parameters derived from the LIDET dataset (v1.0.1) [Source Code]. https://doi.org/10.5281/zenodo.10162303

Juice, S. M., Walter, C. A., Allen, K. E., Berardi, D. M., Hudiburg, T. W., Sulman, B. N., & Brzostek, E. R. (2022). A new bioenergy model that simulates the impacts of plant-microbial interactions, soil carbon protection, and mechanistic tillage on soil carbon cycling. *GCB Bioenergy*, *14*(3), 346–363. https://doi.org/10.1111/gcbb.12914

Kan, Z.-R., Virk, A. L., Wu, G., Qi, J.-Y., Ma, S.-T., Wang, X., et al. (2020). Priming effect intensity of soil organic carbon mineralization under no-till and residue retention. *Applied Soil Ecology*, *147*, 103445. https://doi.org/10.1016/j.apsoil.2019.103445

Keenan, T. F., Davidson, E. A., Munger, J. W., & Richardson, A. D. (2013). Rate my data: Quantifying the value of ecological data for the development of models of the terrestrial carbon cycle. *Ecological Applications*, *23*(1), 273–286. https://doi.org/10.1890/12-0747.1

Kurz-Besson, C., Coûteaux, M.-M., Thiéry, J. M., Berg, B., & Remacle, J. (2005). A comparison of litterbag and direct observation methods of Scots pine needle decomposition measurement. *Soil Biology and Biochemistry*, *37*(12), 2315–2318. https://doi.org/10.1016/j.soilbio.2005.03.022

Kyker-Snowman, E., Wieder, W. R., Frey, S. D., & Grandy, A. S. (2020). Stoichiometrically coupled carbon and nitrogen cycling in the MIcrobial-MIneral Carbon Stabilization model version 1.0 (MIMICS-CN v1.0). *Geoscientific Model Development*, *13*(9), 4413–4434. https://doi.org/10.5194/gmd-13-4413-2020

Lal, R., Griffin, M., Apt, J., Lave, L., & Morgan, M. G. (2004). Managing soil carbon. *Science*, *304*(5669), 393. https://doi.org/10.1126/science.1093079

Lawrence, C. R., Neff, J. C., & Schimel, J. P. (2009). Does adding microbial mechanisms of decomposition improve soil organic matter models? A comparison of four models using data from a pulsed rewetting experiment. *Soil Biology and Biochemistry*, *41*(9), 1923–1934. https://doi.org/10.1016/j.soilbio.2009.06.016

Long-term Intersite Decomposition Experiment Team (LIDET). (1995). *Meeting the challenges of long-term, broad scale ecological experiments (Publication No. 19)* (p. 23). LTER Network Office.

Luengo, D., Martino, L., Bugallo, M., Elvira, V., & Särkkä, S. (2020). A survey of Monte Carlo methods for parameter estimation. *EURASIP Journal on Advances in Signal Processing*, *2020*(1), 25. https://doi.org/10.1186/s13634-020-00675-6

Lupwayi, N. Z., Clayton, G. W., O'Donovan, J. T., Harker, K. N., Turkington, T. K., & Rice, W. A. (2004). Decomposition of crop residues under conventional and zero tillage. *Canadian Journal of Soil Science*, *84*(4), 403–410. https://doi.org/10.4141/s03-082

Ma, S., Jiang, L., Wilson, R. M., Chanton, J. P., Bridgham, S., Niu, S., et al. (2022). Evaluating alternative ebullition models for predicting peatland methane emission and its pathways via data–model fusion. *Biogeosciences*, *19*(8), 2245–2262. https://doi.org/10.5194/bg-19-2245-2022

Meentemeyer, V. (1978). Macroclimate and lignin control of litter decomposition rates. *Ecology*, *59*(3), 465–472. https://doi.org/10.2307/1936576

Metropolis, N., & Ulam, S. (1949). The Monte Carlo method. *Journal of the American Statistical Association*, *44*(247), 335–341. https://doi.org/10.1080/01621459.1949.10483310

Mo, F., Zhang, Y.-Y., Liu, Y., & Liao, Y.-C. (2021). Microbial carbon-use efficiency and straw-induced priming effect within soil aggregates are regulated by tillage history and balanced nutrient supply. *Biology and Fertility of Soils*, *57*(3), 409–420. https://doi.org/10.1007/s00374-021-01540-w

Moore, C. E., Berardi, D. M., Blanc-Betes, E., Dracup, E. C., Egenriether, S., Gomez-Casanovas, N., et al. (2020). The carbon and nitrogen cycle impacts of reverting perennial bioenergy switchgrass to an annual maize crop rotation. *GCB Bioenergy*, *12*(11), 941–954. https://doi.org/10.1111/gcbb.12743

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

Parton, W., Schimel, D. S., Cole, C., & Ojima, D. (1987). Analysis of factors controlling soil organic-matter levels in Great-Plains grasslands. *Soil Science Society of America Journal*, *51*(5), 1173–1179. https://doi.org/10.2136/sssaj1987.03615995005100050015x

Parton, W., Silver, W. L., Burke, I. C., Grassens, L., Harmon, M. E., Currie, W. S., et al. (2007). Global-scale similarities in nitrogen release patterns during long-term decomposition. *Science*, *315*(5810), 361–364. https://doi.org/10.1126/science.1134853

Parton, W. J., Hartan, M., Ojima, D., & Schimel, D. (1998). DAYCENT and its land surface submodel: Description and testing. *Global and Planetary Change*, *19*(1–4), 35–48. https://doi.org/10.1016/s0921-8181(98)00040-x

Pierson, D., Lohse, K. A., Wieder, W. R., Patton, N. R., Facer, J., de Graaff, M.-A., et al. (2022). Optimizing process-based models to predict current and future soil organic carbon stocks at high-resolution. *Scientific Reports*, *12*(1), 10824. https://doi.org/10.1038/s41598-022-14224-8

Qin, Z., Dunn, J. B., Kwon, H., Mueller, S., & Wander, M. M. (2016). Influence of spatially dependent, modeled soil carbon emission factors on life-cycle greenhouse gas emissions of corn and cellulosic ethanol. *GCB Bioenergy*, *8*(6), 1136–1149. https://doi.org/10.1111/gcbb.12333

Rasse, D. P., Rumpel, C., & Dignac, M.-F. (2005). Is soil carbon mostly root carbon? Mechanisms for a specific stabilisation. *Plant and Soil*, *269*(1), 341–356. https://doi.org/10.1007/s11104-004-0907-y

R Core Team. (2023). R: A language and environment for statistical computing [Software]. Retrieved from https://www.R-project.org/

Reicosky, D., Dugas, W., & Torbert, H. (1997). Tillage-induced soil carbon dioxide loss from different cropping systems. *Soil and Tillage Research*, *41*(1–2), 105–118. https://doi.org/10.1016/s0167-1987(96)01080-x

Ridgeway, J. R., Morrissey, E. M., & Brzostek, E. R. (2022). Plant litter traits control microbial decomposition and drive soil carbon stabilization. *Soil Biology and Biochemistry*, *175*, 108857. https://doi.org/10.1016/j.soilbio.2022.108857

Sandhu, O. S., Jat, M. L., Gupta, R. K., Thind, H. S., Sidhu, H. S., & Singh, Y. (2022). Influence of residue type and method of placement on dynamics of decomposition and nitrogen release in maize-wheat-mungbean cropping on permanent raised beds: A litterbag study. *Sustainability*, *14*(2), 864. https://doi.org/10.3390/su14020864

Shi, M., Fisher, J. B., Brzostek, E. R., & Phillips, R. P. (2016). Carbon cost of plant nitrogen acquisition: Global carbon cycle impact from an improved plant nitrogen cycle in the community land model. *Global Change Biology*, *22*(3), 1299–1314. https://doi.org/10.1111/gcb.13131

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

Six, J., Elliott, E. T., & Paustian, K. (1999). Aggregate and soil organic matter dynamics under conventional and no-tillage systems. *Soil Science Society of America Journal*, *63*(5), 1350–1358. https://doi.org/10.2136/sssaj1999.6351350x

Slade, E. M., & Riutta, T. (2012). Interacting effects of leaf litter species and macrofauna on decomposition in different litter environments. *Basic and Applied Ecology*, *13*(5), 423–431. https://doi.org/10.1016/j.baae.2012.06.008

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

Sulman, B. N., Moore, J. A. M., Abramoff, R., Averill, C., Kivlin, S., Georgiou, K., et al. (2018). Multiple models and experiments underscore large uncertainty in soil carbon dynamics. *Biogeochemistry*, *141*(2), 109–123. https://doi.org/10.1007/s10533-018-0509-z

Sulman, B. N., Phillips, R. P., Oishi, A. C., Shevliakova, E., & Pacala, S. W. (2014a). Microbe-driven turnover offsets mineral-mediated storage of soil carbon under elevated CO_2. *Nature Climate Change*, *4*(12), 1099–1102. https://doi.org/10.1038/nclimate2436

Sulman, B. N., Phillips, R. P., Oishi, A. C., Shevliakova, E., & Pacala, S. W. (2014b). Supplementary information: Microbe-driven turnover offsets mineral-mediated storage of soil carbon under elevated CO_2. *Nature Climate Change*, *4*(12), 1099–1102. https://doi.org/10.1038/NCLIMATE2436

Sulman, B. N., Shevliakova, E., Brzostek, E. R., Kivlin, S. N., Malyshev, S., Menge, D. N. L., & Zhang, X. (2019). Diverse mycorrhizal associations enhance terrestrial C storage in a global model. *Global Biogeochemical Cycles*, *33*(4), 501–523. https://doi.org/10.1029/2018GB005973

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

Torquebiau, E. (2017). Climate change, a research challenge: The example of the 4‰ hypothesis. *OCL*, *24*(1), D108. https://doi.org/10.1051/ocl/2016054

Wallenstein, M. D., & Weintraub, M. N. (2008). Emerging tools for measuring and modeling the in situ activity of soil extracellular enzymes. *Soil Biology and Biochemistry*, *40*(9), 2098–2106. https://doi.org/10.1016/j.soilbio.2008.01.024

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

Wang, Y.-P., Zhang, H., Ciais, P., Goll, D., Huang, Y., Wood, J. D., et al. (2021). Microbial activity and root carbon inputs are more important than soil carbon diffusion in simulating soil carbon profiles. *Journal of Geophysical Research: Biogeosciences*, *126*(4), e2020JG006205. https://doi.org/10.1029/2020JG006205

Weibull, W. (1951). A statistical distribution function of wide applicability. *Journal of Applied Mechanics*, *18*(3), 293–297. https://doi.org/10.1115/1.4010337

Wieder, W. R., Bonan, G. B., & Allison, S. D. (2013). Global soil carbon projections are improved by modelling microbial processes. *Nature Climate Change*, *3*(10), 909–912. https://doi.org/10.1038/nclimate1951

Wieder, W. R., Grandy, A. S., Kallenbach, C. M., & Bonan, G. B. (2014). Integrating microbial physiology and physio-chemical principles in soils with the MIcrobial-MIneral Carbon Stabilization (MIMICS) model. *Biogeosciences*, *11*(14), 3899–3917. https://doi.org/10.5194/bg-11-3899-2014

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

Wieder, W. R., Sulman, B. N., Hartman, M. D., Koven, C. D., & Bradford, M. A. (2019). Arctic soil governs whether climate change drives global losses or gains in soil carbon. *Geophysical Research Letters*, *46*(24), 14486–14495. https://doi.org/10.1029/2019GL085543

Woolf, D., & Lehmann, J. (2019). Microbial models with minimal mineral protection can explain long-term soil organic carbon persistence. *Scientific Reports*, *9*(1), 6522. https://doi.org/10.1038/s41598-019-43026-8

Wu, J., Brookes, P. C., & Jenkinson, D. S. (1993). Formation and destruction of microbial biomass during the decomposition of glucose and ryegrass in soil. *Soil Biology and Biochemistry*, *25*(10), 1435–1441. https://doi.org/10.1016/0038-0717(93)90058-J

Zang, H., Wang, J., & Kuzyakov, Y. (2016). N fertilization decreases soil organic matter decomposition in the rhizosphere. *Applied Soil Ecology*, *108*, 47–53. https://doi.org/10.1016/j.apsoil.2016.07.021

Zeri, M., Anderson-Teixeira, K., Hickman, G., Masters, M., DeLucia, E., & Bernacchi, C. J. (2011). Carbon exchange by establishing biofuel crops in Central Illinois. *Agriculture, Ecosystems and Environment*, *144*(1), 319–329. https://doi.org/10.1016/j.agee.2011.09.006

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

Zhang, H., Goll, D. S., Wang, Y.-P., Ciais, P., Wieder, W. R., Abramoff, R., et al. (2020). Microbial dynamics and soil physicochemical properties explain large-scale variations in soil organic carbon. *Global Change Biology*, *26*(4), 2668–2685. https://doi.org/10.1111/gcb.14994