Generalizing Microbial Parameters in Soil Biogeochemical Models: Insights From a Multi-Site Incubation Experiment

Siyang Jian[1,2] , Jianwei Li[2] , Gangsheng Wang[3], Jizhong Zhou[1,4,5], Christopher W. Schadt[6] , and Melanie A. Mayes[7]

[1]Institute for Environmental Genomics and Department of Microbiology & Plant Biology, University of Oklahoma, Norman, OK, USA, [2]Department of Agricultural and Environmental Sciences, Tennessee State University, Nashville, TN, USA, [3]State Key Laboratory of Water Resources Engineering and Management, Institute for Water-Carbon Cycles and Carbon Neutrality, Wuhan University, Wuhan, China, [4]School of Civil Engineering and Environmental Sciences, University of Oklahoma, Norman, OK, USA, [5]Earth and Environmental Sciences, Lawrence Berkeley National Laboratory, Berkeley, CA, USA, [6]Oak Ridge National Laboratory, Biosciences Division & Climate Change Science Institute, Oak Ridge, TN, USA, [7]Oak Ridge National Laboratory, Environmental Sciences Division & Climate Change Science Institute, Oak Ridge, TN, USA

Abstract Incorporating microbial processes into soil biogeochemical models has received growing interest. However, determining the parameters that govern microbially driven biogeochemical processes typically requires case-specific model calibration in various soil and ecosystem types. Here each case refers to an independent and individual experimental unit subjected to repeated measurements. Using the Microbial-ENzyme Decomposition model, this study aimed to test whether a common set of microbially-relevant parameters (i.e., generalized parameters) could be obtained across multiple cases based on a two-year incubation experiment in which soil samples of four distinct soil series (i.e., Coland, Kesswick, Westmoreland, and Etowah) collected from forest and grassland were subjected to cellulose or no cellulose amendment. Results showed that a common set of parameters controlling microbial growth and maintenance as well as extracellular enzyme production and turnover could be generalized at the soil series level but not land cover type. This indicates that microbial model developments need to prioritize soil series type over plant functional types when implemented across various sites. This study also suggests that, in addition to heterotrophic respiration and microbial biomass data, extracellular enzyme data sets are needed to achieve reliable microbial-relevant parameters for large-scale soil model projections.

Plain Language Summary Incorporating soil microbial processes can improve soil model projections, and achieving a common set of microbial parameters across sites remains less studied. Based on a two-year soil incubation data set, this study showed that key microbial parameters could be generalized at the soil series level (four distinct soil series) but not land cover type (forest vs. grassland). The common set of parameters includes those processes controlling microbial growth and maintenance as well as extracellular enzyme production and turnover. This study informs that future microbial model developments prioritize soil series type over plant functional types when implemented across various sites. Besides the heterotrophic respiration and microbial biomass data, soil extracellular enzyme data sets are particularly needed to achieve reliable microbial-relevant parameters for large-scale soil model projections.

1. Introduction

Incorporating soil microbial processes into an Earth system model (ESM) improved soil carbon (C) projection and reduced uncertainty of climate-carbon feedbacks (Wieder et al., 2013). Further improvements are sought by explicit inclusion of microbial processes given rigorous model calibration and validation (Luo et al., 2016; Schimel, 2023). In particular, microbial traits such as growth and maintenance (German et al., 2012; Wang et al., 2013), microbial dormancy (Wang et al., 2015), acclimation (Allison et al., 2010), turnover (Fan et al., 2021; Sulman et al., 2014) and community level interaction (Georgiou et al., 2017) impose key controls on the soil C decomposition dynamics. There are also growing interests in integrating omics information into microbial-explicit models to link microbial enzyme activity and ecosystem functions (Chen & Sinsabaugh, 2021; Guo et al., 2020; Wang et al., 2022). However, microbial model parameterization faces significant challenges due to the ambiguous definition of microbial parameters (Schimel et al., 2022) and the lack of in situ quantification capacity (Jian et al., 2020; Wieder et al., 2015).

Software: Gangsheng Wang
Supervision: Jianwei Li, Melanie A. Mayes
Validation: Siyang Jian, Gangsheng Wang, Christopher W. Schadt
Writing – original draft: Siyang Jian, Jianwei Li
Writing – review & editing: Jianwei Li, Gangsheng Wang, Jizhong Zhou, Christopher W. Schadt, Melanie A. Mayes

Microbial-explicit models generally have multiple sub-pools of soil organic C (SOC), such as dissolved organic carbon (DOC), microbial biomass carbon (MBC) and enzyme (ENZ) pools. Microbes release enzymes to decompose substrates and acquire DOC. Then microbes either release the C primarily through respiration or incorporate the C for growth or maintenance (Todd-Brown et al., 2012; Wieder et al., 2015). Nonlinear microbial models consist of substantially more parameters compared to the first-order decomposition models (e.g., CENTURY) and face a critical challenge of parameterization (Li et al., 2014). Field and laboratory experiments provide valuable data for model parameterization and validation, including measurements such as heterotrophic respiration, SOC, and MBC in response to environmental changes (Li et al., 2019; Wang, Huang, et al., 2019).

Microbial model parameterization can be achieved by either single-case calibration or multiple-case calibration. Single-case calibration refers to constraining parameters based on observational data collected from a single site under a specific treatment over a particular timeframe. For example, data sets collected from the control and warmed sites at the Harvard Forest were used independently to constrain parameters resulting in two sets of parameters (Li et al., 2019). On the other hand, multiple-case calibration is to combine observational data from multiple studies covering different soil types or vegetation to produce a common set of model parameters that can be used across varied environments. For instance, Zhang et al. (2020) calibrated the MIcrobial-Mineral Carbon Stabilization model (MIMICS) based on observations from 206 forest sites and found better prediction of SOC compared to the CENTURY model. The choice of calibration approaches depends on the modeling questions to be addressed. Single-case calibration is often implemented to provide mechanistic understanding of environmental perturbations on microbial community functional traits (e.g., decomposition kinetics or temperature sensitivity). In contrast, multiple-case calibration is preferred for answering large-scale questions through model extrapolation. However, it has never been evaluated to what extent key microbial parameters can be generalized across treatments, soil series, and land covers without compromising overall model performance.

With increasing mechanistic representation of microbial processes in models (Schimel, 2023; Sulman et al., 2018), it is desired to implement one common set of parameters instead of treatment-dependent parameters to capture microbial activities under various environmental changes. For instance, microbial heterotrophic respiration rates were simulated by one set of parameters for soils subjected to different nutrient manipulations (Blagodatsky et al., 2010; Wang et al., 2022). However, implementing the common set of parameters at ecosystem scale must take into account the heterogeneity of soils across different environments (Chakrawal et al., 2020; Wieder et al., 2015), as microbial parameters have been shown to vary across different environmental and edaphic gradients (Sinsabaugh et al., 2013; Wang et al., 2015). This suggests that application of microbial models at the large scale could require extensive observational data for model parameterization. However, it is common to extrapolate microbial parameters derived from the site-specific data to the biome scale to simulate global soil C storage (Wang et al., 2017). Furthermore, synthesized data sets have enabled the derivation of microbial parameters based on empirical relationships with environmental factors (Bond-Lamberty & Thomson, 2010). For example, Ye et al. (2019) found microbial carbon use efficiency (CUE), a key parameter defined as the fraction of microbial C uptake allocated to growth, is positively correlated with mean annual temperature (MAT) globally. This implies that some key microbial parameters can be generalized based on empirical relationships, but it is important to test their reliability in simulating decomposition dynamics.

Ideally, microbial parameters should be generalized across models aiming to capture the same microbial processes. Contemporarily, microbial models still depict similar microbial processes while differing in structure (Li et al., 2014; Sulman et al., 2018). The Microbial-ENzyme Decomposition (MEND) model has a relatively complex structure, explicitly representing microbial physiology and soil organic matter (SOM) decomposition catalyzed by oxidative or hydrolytic enzymes. Its focus on enzymatic processes has prompted several data assimilation studies, resulting in refined predictions of microbial response to priming, experimental warming, and nitrogen amendment (Guo et al., 2020; Li et al., 2019; Tao et al., 2020, 2024; Wang et al., 2022). Therefore, this study employs the MEND model to demonstrate whether generalized microbial parameters can be achieved across different treatments, soil series, or land cover types. This study does not delve into the applicability of generalized parameters to other microbial models.

Here, we compiled a data set of 729-day soil incubation experiments that represented two treatments with and without substrate amendments (i.e., cellulose), four distinct soil series (i.e., Coland, Kesswick, Westmoreland, and Etowah), and two land covers (forest vs. grassland). We first performed 16 independent single-case calibrations to constrain four microbial kinetic parameters relevant to microbial uptake, maintenance, growth, and

Table 1
Site Characteristics, Soil Classification, and Chemistry of Soil Samples in the Incubation Study

State	Land cover	Location	Soil series	Soil taxonomy	pH	Carbon (%)	Nitrogen (%)
IA	Forest	41.79°N, 93.43°W	Coland	Cumulic *Endoaquolls*	6.56	3.17 (0.29)	0.27 (0.01)
	Grass				6.45	3.13 (0.62)	0.25 (0.01)
MO	Forest	38.74°N, 92.19°W	Kesswick	Aquertic Chromic *Hapludalfs*	5.89	4.42 (0.62)	0.22 (0.13)
	Grass				5.57	2.21 (0.08)	0.20 (0.01)
OH	Forest	39.32°N, 82.12°W	Westmoreland	Ultic *Hapludalfs*	5.5	4.14 (0.06)	0.31 (0.01)
	Grass				6.36	2.18 (0.08)	0.23 (0.01)
TN	Forest	35.93°N, 84.31°W	Etowah	Typic *Paleudults*	7.56	5.71 (0.22)	0.47 (0.07)
	Grass				7.29	3.08 (0.16)	0.33 (0.03)

Note. Standard deviations are listed in parenthesis based on four analytical replicates.

carbon use efficiency. Then, we performed multiple-case calibrations to constrain the above-mentioned microbial kinetic parameters as well as four enzymatic parameters (see Materials and Methods). The model performances based on multiple-case and single-case calibrations were compared to evaluate the effectiveness of generalized parameters associated with substrate treatments, soil series, and land cover. We hypothesized that a common set of microbial parameters would be achieved at the soil series level, but not at the land cover type given that soil biogeochemical characteristics associated with soil series are major drivers of soil microbial community function. It is expected that one or more soil features will be responsible for key microbial functions that are distinct among soil series. We also hypothesized that a common set of microbial parameters would be achieved for different treatments of substrate amendments.

2. Materials and Methods

2.1. Data Source

This modeling study utilized incubation data sets from a previously published 2-year incubation experiment (Jian et al., 2020; Kluber et al., 2020). The data plots can be found in the Figures S3–S6 in Jian et al. (2020), see also Tables S1–S3 in Supporting Information S1. With two replicates in each plot, eight soil samples were collected from four paired forest and grassland plots located in Iowa (IA), Missouri (MO), Ohio (OH) and Tennessee (TN), representing four different soil series of Coland (a Mollisol from IA), Kesswick (an Alfisol from MO), Westmoreland (an Alfisol from OH), and Etowah (an Ultisol from TN). The paired forest and grassland plots, representing different land covers, were located within 1 mile of each other and mapped to the same soil series, according to the U.S. Department of Agriculture Natural Resources Conservation Service (NRCS) soil surveys (Table 1). Forest sites were mature mixed eastern deciduous forests (>60 years old) and grassland sites were minimally managed grasslands with no recent fertilization or grazing, although all sites were mowed semi-annually or annually (Kluber et al., 2020).

The 729-day incubation experiment included two levels of substrate treatments (control vs. cellulose). There were 16 incubation cases in total (4 soil series × 2 land covers × 2 substrate treatments). During the 2-year period, heterotrophic soil respiration rate (R_h) was measured at 18 timepoints (i.e., 1, 2, 4, 6, 11, 20, 39, 64, 90, 120, 151, 229, 323, 390, 480, 571, 665, and 729 days) and soil MBC was measured at 8 timepoints (i.e., 1, 4, 20, 64, 151, 323, 480, and 729 days). The incubation data sets were used to perform the single-case and multiple-case calibrations to achieve generalized parameters based on replicate soil sample, soil series, and land cover, respectively (Table 2).

2.2. MEND Model and Its Multi-Case Version (MENDmult)

The MEND model describes the SOM decomposition processes by explicitly representing relevant microbial and enzymatic physiology (Wang et al., 2015). The SOM pool consists of two particulate organic matter (POM) pools and one mineral-associated organic matter (MOM) pool. The two POMs are decomposed by oxidative and hydrolytic enzymes, respectively, while the MOM is decomposed by a generic enzyme group associated with MOM (EM). Dissolved organic matter (DOM) produced by enzymatic decomposition is taken up by the microbes. The

Table 2
The 729-Day Incubation Data Set Re-Grouped for Single-Case and Multiple-Case Calibrations Based on Soil Replicate, Soil Series, and Land Cover Type

Grouping	Number of groups	Number of cases in each group	Number of observations in each calibration
Single case	16	1	26
Treatment	8	2	52
Soil series	4	4	104
Land cover	2	8	208

model's state variables, governing equations, component fluxes and parameters are described in Tables S1–S3 in Supporting Information S1, respectively. A schematic diagram of the MEND model describing the carbon pools and fluxes was presented in Figure 1.

The soil chemical measurements including SOC, POC, MOC, DOC and MBC were extracted from Tables 2a–2c in Kluber et al. (2020) to initialize the MEND model. Given that the incubation was carried out in a well-controlled laboratory and soil moisture was frequently monitored, the model forcing adopted a constant temperature at 22°C and a constant moisture at 30% gravimetric water content. Most of the model parameters were fixed across all incubation settings except for those parameters to be constrained described below (Table S3 in Supporting Information S1).

This study first focused on constraining four key microbial kinetic parameters, including the maximum specific growth rate (V_g), the ratio (α) of the maximum specific maintenance rate (V_{mt}) to ($V_g + V_{mt}$), the half-saturation constant for microbial uptake of DOM (K_D), and the intrinsic carbon use efficiency (Y_g) at reference temperature. These four microbial kinetic parameters were selected because they regulated microbial uptake, growth, maintenance and transformation between dormancy and activation (Jian et al., 2020; Wang et al., 2014). For single-case calibration, all 18 R_h and 8 MBC measurements within each case were used for model calibration to generate single-case parameters, as described in our previous study (Jian et al., 2020). For multiple-case calibration, we
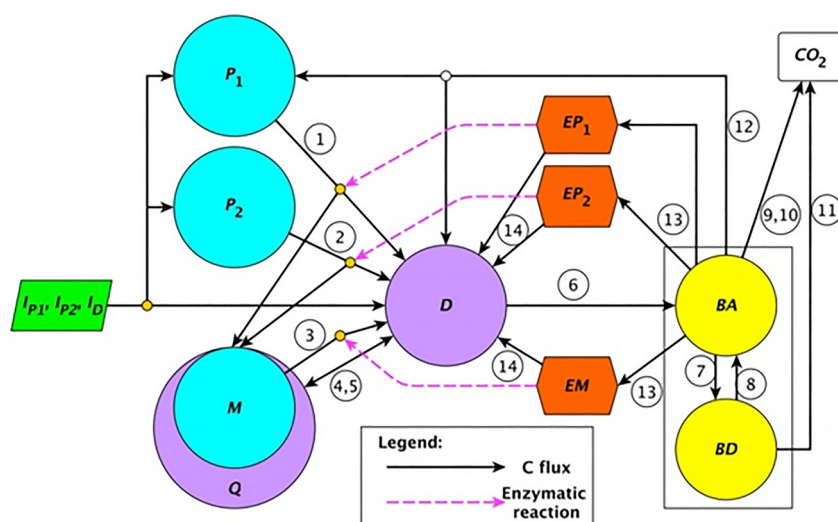

Figure 1. Diagram of the Microbial-ENzyme Decomposition (MEND) model adapted from Jian et al. (2020). Soil organic carbon pools include: (1) particulate organic C (POC), which can be further divided into two components POC$_1$ (P_1, containing POC that can be degraded by oxidative enzymes) and POC$_2$ (P_2, containing POC that can be degraded by hydrolytic enzymes); (2) mineral-associated organic C (MOC, M); (3) dissolved organic C (DOC, D); (4) adsorbed DOC (QOC, Q): an active layer of MOC that adsorbs and desorbs DOC; (5) active microbial biomass (BA) and dormant microbial biomass (BD); (6) enzyme pools containing POC-degraded enzymes (EP_1 and EP_2 that decompose POC$_1$ and POC$_2$, respectively) and MOC-degraded enzymes (EM). External inputs to the model can be separated into I$_{P1}$, I$_{P2}$ and I$_D$ denoting inputs to the pools of P_1, P_2, and DOC, respectively.

generated common sets of the same four microbial kinetic parameters (also called generalized parameters) based on the combined cases under the same soil substrate treatment, soil series, and land cover, respectively.

When implementing multiple-case calibration, we further constrained additional four enzymatic parameters in addition to the four abovementioned kinetic parameters. The new parameters included the fraction of enzyme depolymerized POM allocated to DOM (f_D), turnover rate of enzymes (r_E), production rate of enzyme specialized to decompose POM pool (p_{EP}), and a ratio ($f_{pEM} = p_{EM}/p_{EP}$) between the production rate (p_{EM}) of enzyme targeting MOM pool and p_{EP}. These four parameters were often selected in field-scale model-data integration studies for their important role in regulating enzyme pools and carbon flow of enzymatic reaction (Guo et al., 2020; Wang, Huang, et al., 2019; Wang et al., 2022).

We used a modified Shuffled Complex Evolution (SCE) algorithm embedded within the MEND model to calibrate selected model parameters by minimizing the overall objective function (J) as shown in Equations 1 and 2 (Wang et al., 2021). SCE is a stochastic optimization method that includes competitive evolution of a "complex" of points spanning the parameter space and the shuffling of complexes (Duan et al., 1992). SCE has been widely used in calibration of hydrological, environmental, and ecosystem models and proved to be efficient and robust (Jian et al., 2020; Wang et al., 2018, 2022; Zhang et al., 2009).

$$J_k = 0.5 \cdot J_1 + 0.5 \cdot J_2 \tag{1}$$

$$J = \sum_{k=1}^{N} \left(\frac{1}{N} \cdot J_k \right) \tag{2}$$

To facilitate model calibrations for multiple cases, MENDmult, the formerly developed multi-case version of the MEND model (Wang, 2015) was implemented in this study. It allows the simultaneous running of multiple cases, with each case having its own model settings and input data. MENDmult also enables the use of one set of parameters to fit the observations in multiple cases, achieved by minimizing an overall objective function that evaluates the modeling performances of all cases. Identical to the MEND model, MENDmult also used objective function value (J) to assess model performance, which was computed as the weighted average of objective function values of R_h (J_1) and MBC (J_2), where J_1 and J_2 were calculated as the ($1–R^2$) and the mean absolute relative error (*MARE*), respectively (Equations 3 and 4). A better model performance was evaluated by the goodness-of-fit represented by a higher R^2 or lower MARE. We determined acceptable model performance when both objective function values of R_h and MBC were smaller than 0.5. The performance of the generalized parameter set was assessed by the weighted average of the objective function value of each individual case (J_k, Equations 1 and 2).

$$R^2 = 1 - \frac{\sum_{i=1}^{n} [Y_{\text{sim}}(i) - Y_{\text{obs}}(i)]^2}{\sum_{i=1}^{n} [Y_{\text{obs}}(i) - \overline{Y_{\text{obs}}}]^2} \tag{3}$$

$$\text{MARE} = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{Y_{\text{sim}}(i) - Y_{\text{obs}}(i)}{Y_{\text{obs}}(i)} \right| \tag{4}$$

where Y_{sim} and Y_{obs} are the simulated and observed values of the response variable, respectively; $\overline{Y_{\text{obs}}}$ is the mean value for Y_{obs}; k denotes each individual case; and N denotes the number of cases.

2.3. Statistical Analysis

All statistical analyses were carried out using R software 4.0.3 (R Core Team, 2020). To compare the model performance based on multiple-case calibration against single-case calibration, the objective function values of simulating R_h and MBC were compared using the Wilcoxon test. The non-parametric Kruskal–Wallis test was employed to test whether the parameter uncertainty significantly differed among the four soil series. The uncertainties of model parameters were quantified by the UQ-COFI (Uncertainty Quantification by Critical Objective Function Index) method (Wang et al., 2022). The UQ-COFI method is based on a global stochastic

optimization technique (e.g., SCE in this study). It also accounts for model complexity (represented by the number of model parameters) and observational data availability (represented by the number of observations). The critical objective function index (COFI) (J_{cr}) is calculated following Equation 5.

$$J_{cr} = J_{opt} \cdot \left(1 + \frac{p}{n-p} F_{\alpha,p,n-p}\right) \tag{5}$$

where J_{cr} is the COFI that defines the parameter uncertainty range, J_{opt} is the optimum (minimum) objective function value achieved by calibration, n is the number of observational data points, p is the number of parameters, and $F_{\alpha,p,n-p}$ is the value of the F-distribution for α, p, and n-p.

In this study, given that R_h and MBC were used to calibrate the model, the relationship between observed R_h and MBC is the key to constraining model parameters. To explore how edaphic factors interacted with apparent kinetic rate (i.e., R_h/MBC), a linear mixed effect model was implemented based on the following formula:

$$R_h/MBC \sim pH + SOC + TN + Sand + Clay + (1|Timepoint) + (1 |Soil\ series : Land\ cover$$
$$: substrate\ addition)$$

Mass specific heterotrophic respiration (R_h/MBC), independent of microbial biomass size, was selected to investigate the apparent kinetic rate. Edaphic factors included were pH, SOC, soil total N (TN), sand, and clay contents. The linear mixed effect model accounts for the differences in treatment, soil series, or plant type by using the nested random factors *(1| Soil series*: *Land cover*: *substrate addition)*. The regression coefficients represent the directions and magnitudes of the fixed effects. Environmental variables were standardized by the *scale()* function in R when constructing the linear mixed effect model. And the standardized slope estimates ± standard error for each fixed effect were derived by the *summary()* function on the constructed linear mixed effect model. The Wald type II χ^2 tests were employed to calculate the p values from the linear mixed effect model using the *lme4* and *car* packages (Wu et al., 2022).

3. Results

3.1. Parameter Estimates Based on Single- and Multiple-Case Calibrations

When calibrating only four microbial kinetic parameters through multiple-case calibration, the generalized parameters had poorer model performance compared to single-case parameters because the objective function values ($J_1 = 1-R^2$) of R_h with the generalized parameters were all significantly higher than those with single-case parameters ($p < 0.001$; Figure 2a). However, the model performance of simulating MBC was less affected by the multiple-case calibration (Figure 2b). Thus, implementing multiple-case calibration to derive a common set of microbial kinetic parameters was not successful based on the performance of simulating R_h over combinations of treatments, soil series, or land covers.

After constraining an additional set of four enzymatic parameters along with the original four microbial kinetic parameters through multiple-case calibration, the model performance based on the generalized parameters were both improved over different treatments and soil series (Figures 2c and 2d). The averaged objective function values of R_h were reduced by 68.8% and 56.1% relative to those without constraining enzymatic parameters, respectively, and did not differ from those with single-case parameters (Figure 1c; $p > 0.05$). More importantly, the objective function values of both R_h and MBC were less than 0.5 except that of R_h across land cover (Figures 2c and 2d), meeting the criteria for an acceptable model calibration. However, as for the multiple-case calibration based on land cover, the objective function values of R_h were still significantly higher than those with single-case parameters ($p < 0.001$; Figure 2c).

3.2. Parameter Uncertainties Based on Multiple-Case Calibrations

At the soil series level, each of the two sets of microbial kinetic and enzymatic parameters was further examined by deriving their uncertainty ranges. Among the eight microbial kinetic and enzymatic parameters, each showed large variations of parameter estimate based on multiple-case calibrations at each of four distinct soil series
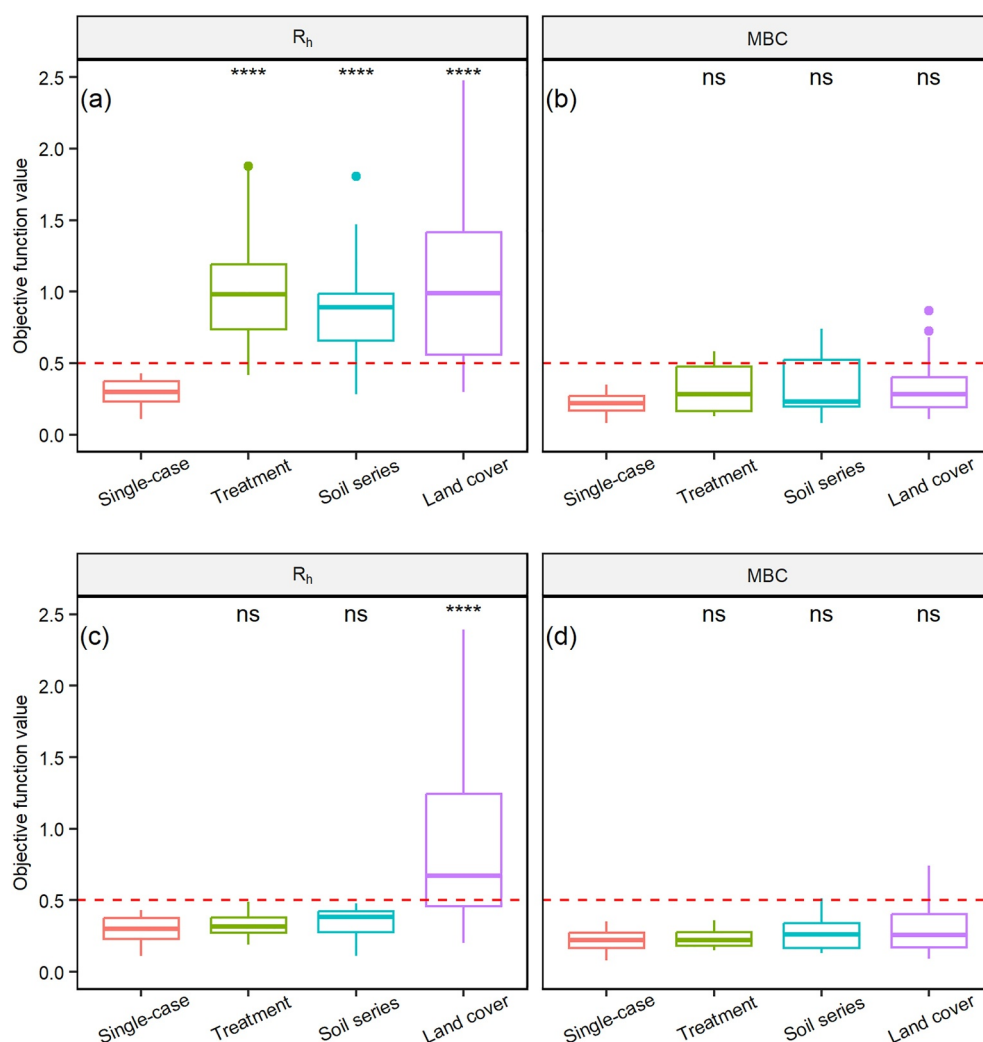
Figure 2. Objective function values for R_h and MBC derived from the single-case and multiple-case calibrations targeting only for four microbial kinetic parameters (V_g, α, K_D and Y_g) (a, b), and for four microbial kinetic parameters plus four enzymatic parameters (r_E, p_{EP}, fp_{EM}, and f_D) (c, d). Smaller objective function values denote better model performance. *, **, and *** denote significant difference between single-case and multiple-case calibrations based on Wilcoxon tests at p-value <0.05, <0.01, and <0.001, respectively. "ns" means not significant.

(Figure 3). For r_E, f_D, and Y_g, their uncertain ranges appeared narrower; overall, each of all eight parameters were significantly different among the four different soil series based on the K-W tests (Figure 3).

3.3. Edaphic Factors Interact With Microbial Kinetic Rates

The linear mixing effect model showed that the standardized slope estimates ±standard error for each fixed effects were −0.64 ± 0.13 (pH, $p < 0.001$), −0.23 ± 0.12 (soil C, $p = 0.06$), 0.53 ± 0.19 (TN, $p = 0.005$), 0.10 ± 0.10 (sand, $p = 0.34$), and −0.03 ± 0.08 (clay, $p = 0.66$). The linear mixed effect model revealed significant negative effects of pH and positive effects of soil nitrogen (N) content on R_h/MBC (Figure 4).

4. Discussion

Integrating microbial processes can improve the performance of ESMs (Wang, Peng, et al., 2019; Wieder et al., 2013), but estimation and implementation of microbial parameters included in microbial-explicit models are still controversial (Crowther et al., 2019; Luo et al., 2016; Wieder et al., 2015). When integrating microbial processes into ESMs, achieving one common set of parameters across conditions can vastly reduce the demand
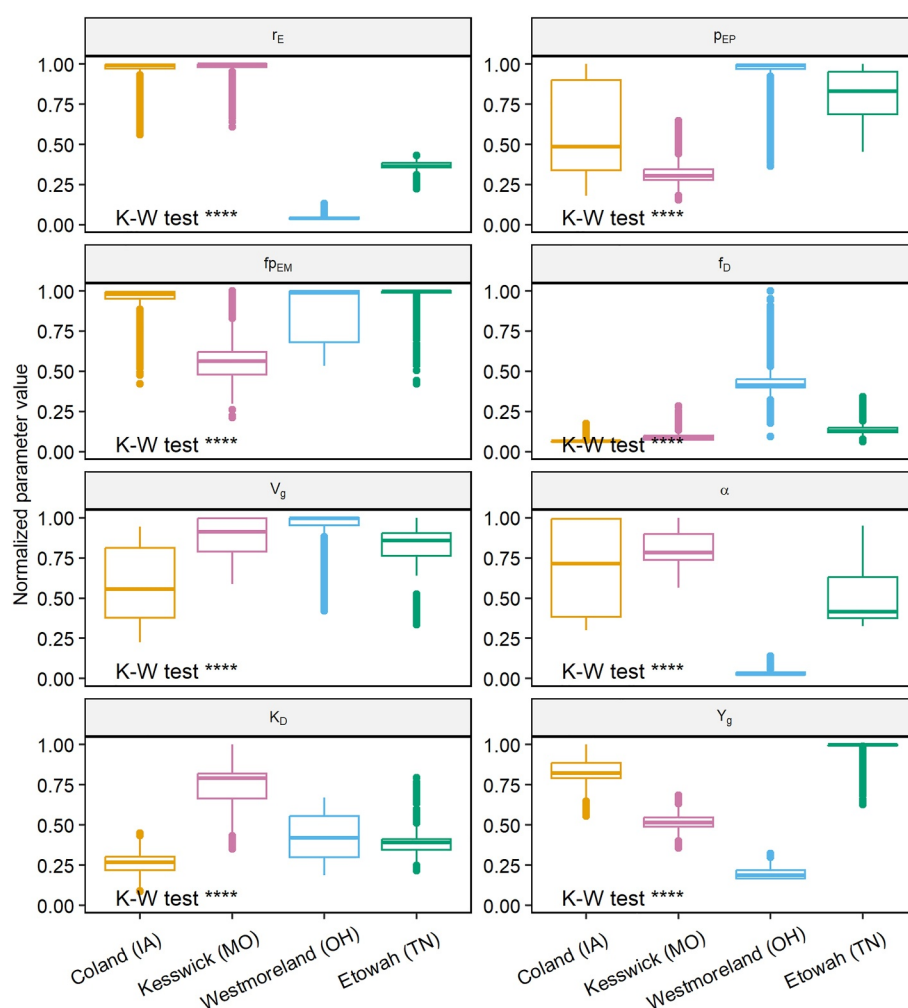
Figure 3. Box plots of normalized values of eight microbial kinetic and enzymatic parameters (V_g, α, K_D, Y_g, r_E, p_{EP}, fp_{EM}, and f_D) based on multiple-case calibrations at soil series level. The normalized value is derived against maximum value of each parameter and used to index uncertainty ranges of parameters. In each box, the bottom and top edges of the box denote the 25th and 75th percentile (the lower and upper quartiles), respectively. The line within a box is the 50th percentile (median). **** denotes significant differences of parameter uncertainty range among four soil series based on the non-parametric Kruskal-Wallis (K–W) at p-value < 0.05.

for expensive manipulative experiment cost and minimize computing power for model calibration and parameterization. In this study, we evaluated how microbial parameters can be generalized to a common set of parameters at different levels (e.g., amendment treatments, soil series, land cover) through model performance comparison using single- and multiple-case calibrations. This provides helpful insights for future studies to include microbial processes in ESMs and constrain microbial parameters.

4.1. Common Set of Parameters of Combined Microbial Kinetic and Enzymatic Processes

In our study, when we only constrained microbial kinetic parameters relevant to microbial uptake, growth and maintenance, the generalized parameter set showed unsatisfactory model performance compared to the parameter set constrained by the single-case calibration. However, by constraining the additional four enzymatic parameters that regulate enzyme production and turnover (r_E, p_{EP} and fp_{EM}) and the partitioning factor for enzyme depolymerized C (f_D), we obtained generalized parameter sets that showed sufficient model performance. With a conceptual framework testing of diffusion limitations of extracellular enzymes and soluble monomers on SOM decomposition, Tang and Riley (2019) found out that SOM depolymerization is limited by the abundance of enzyme binding sites supplied by polymer particles, requesting specific approximation of enzymatic kinetics for
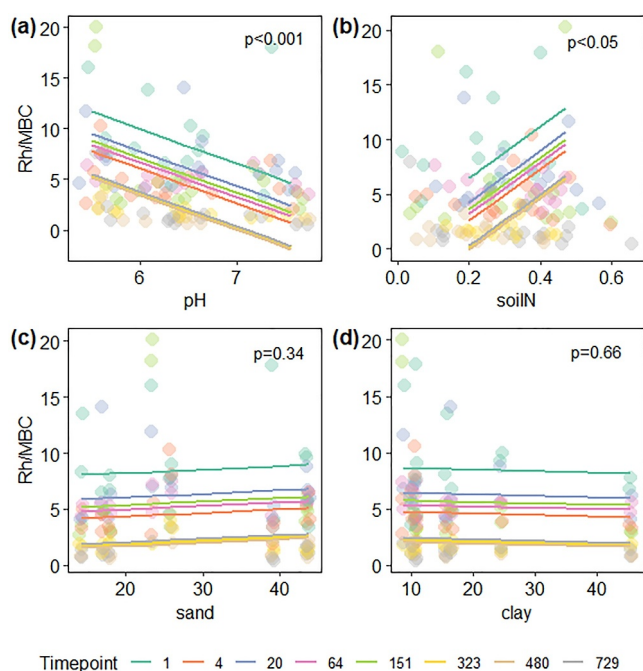
Figure 4. (a–d) Significant interaction of soil pH, soil total N content (soilN), sand, and clay with R_h/MBC for each collection day (i.e., Timepoint) revealed by linear mixed-effects model. The standardized slope estimates ± standard error for soil pH and N were -0.64 ± 0.13 ($p < 0.001$), and 0.53 ± 0.19 ($p = 0.005$), respectively. The data include all treatments (i.e., land cover, soil series, and substrate treatment) at each time point.

better simulation. Because of the competition between enzyme and substrate for mineral surface adsorption (Tang & Riley, 2015), the improved model performance by adding four enzyme parameters shed insights on the possible association and linkage of the enzymatic functions across different soil series. Furthermore, under strong nutrient limitation, relative competitiveness depends strongly on the competitor functional traits (affinity and nutrient carrier enzyme abundance) (Zhu et al., 2016).

The need to include the additional four enzymatic parameters in multiple-case calibration is therefore due to the variability observed in the enzymatic parameters between different soil series (Figure 3). Hence, by including enzymatic parameters in multiple-case calibration procedure, the acceptable model performance indicated the use of common sets of parameters across treatments, soil series, but not at land cover level. This result highlighted the need for exploring the effective incorporation of enzymatic data into microbial-explicit models, for example, through linking functional genes to ecosystem functioning (Chen & Sinsabaugh, 2021; Gao et al., 2020; Guo et al., 2020; Wang et al., 2022). However, the outcome may depend upon the specific model structure; only one microbial model (MENDmult) was tested in the current study.

4.2. Common Set of Parameters for Different Treatments of Substrate Amendments

Our study demonstrated that by including both microbial kinetic and enzymatic parameters, a common set of parameters could be achieved to successfully capture the R_h dynamics of soils under different substrate manipulations. Previous studies have also successfully used one set of parameters to simulate different nutrient manipulation treatments, such as Blagodatsky et al. (2010) and Wang et al. (2022). The implications of our study could extend beyond substrate amendment treatments, such as nutrient additions, warming, and water manipulations. Using a common set of parameters to simulate soils subjected to different substrate or nutrient addition treatments is likely valid only if the potential priming mechanisms (e.g., co-metabolism) or carbon-nitrogen coupled processes are embedded within model structure through explicit representation of microbial processes. In other words, applicability to other models beyond MEND may depend on the model representations of enzyme and microbial processes. Thus microbial-explicit models can simulate microbial responses to nutrient manipulations through intrinsic dynamic processes, such as microbial growth and maintenance, transition between active and dormant states, and varying microbial C/N ratio, without changing the intrinsic kinetic rates of microbial community activities.

Notably, other model-data integration studies using field-collected data sets found that certain microbial parameters (e.g., CUE, microbial turnover rate, and temperature sensitivities of microbial processes) differed significantly under control and warming treatments (Guo et al., 2020; Li et al., 2019), reflecting changes in the functions of the microbial community resulting from their physiological status or community structure (Garcia-Palacios et al., 2021; Zhou et al., 2012). It remains uncertain whether the microbial responses to warming or soil moisture manipulations can be represented with only one set of parameters as most of these microbial processes such as community adaptation are not well resolved at the scales of the ecosystem or community (Liang et al., 2018).

4.3. Common Set of Parameters Across Soil Series and Land Cover Level

Our results showed that the performance of a common set of parameters was acceptable when implementing multiple-case calibration at the soil series level, which consists of four single cases (two land covers × two substrate additions). However, we did not obtain common set of parameters at the land cover level, which contains eight single cases (four soil series × two substrate additions). The generalized parameters differed significantly among these four series (Figure 3), explaining why a common set of parameters cannot be achieved across four soil series within either forest or grassland land cover. In addition, the linear mixed effect model demonstrated substantial linkage among microbial apparent kinetic rate, soil pH and soil N. As soil pH and soil N were more

dependent on the soil series rather than land cover for all soils except for those in OH (Table 1), constraining parameters according to land cover could neglect the influence of these edaphic factors on parameter estimates. On the other hand, constraining parameters based on soil series could account for such influences, contributing to better performance of the common set of parameters.

Key soil environmental factors, such as temperature, moisture, oxygen and nutrient availability, and other factor such as plant root distribution vary substantially in space, which shape different microbial activities across scales from the smallest pores to landscapes and biomes (Fatichi et al., 2019; Li et al., 2018). Our further analysis revealed a strong control of soil pH and soil N content on biomass specific respiration (Figure 4). Integrating soil microscale processes into modeled microbial kinetics is still under debate (Manzoni & Porporato, 2009; Sierra & Müller, 2015; Wieder et al., 2015), though a few attempts have been made (Chakrawal et al., 2020; Ebrahimi & Or, 2016). Regarding upscaling, current approaches involving grid-based or biome-based parameterizations to integrate microbial processes into ESMs showed significant improvement in predicting SOC (Wang et al., 2017; Wang, Peng, et al., 2019; Wieder et al., 2013). Although this approach can reduce computational effort, it does not consider substantial differences in soil characteristics within individual grids or biomes. The direct use of microbial parameters in global soil models thus lacks sufficient justification. The empirical relationships between microbial apparent kinetic rates and edaphic factors revealed in this study suggested a possible solution—the soil series—to link microbial processes with large scale soil models enabling the fusion of microbial functions with global soil model predictions. However, the relationship between apparent kinetic rates and model parameters is still unclear. More data-model integration studies are needed to deepen our understanding of microbial model parameters varying with climate, vegetation, and edaphic properties.

In summary, this study found that microbial parameters can be generalized up to the soil series level (i.e., across different land covers and substrate additions), but generalization cannot be achieved at the land cover level (i.e., across soil series involving different substrate treatments). This study also highlights the importance of common parameter selections including not only microbial kinetic parameters but also enzyme production and turnover parameters for achieving better model performance. On the one hand, the obtained generalized parameters' values and ranges from this study may help benchmark the case-specific calibration in other studies. On the other hand, future studies should investigate whether model parameter variation can be explained by climate, vegetation, and edaphic properties on a larger scale, especially with field data. Critical processes such as mycorrhizal interactions, root respiration, and exudation will affect soil microbial C cycling processes, and the lack of representation in the incubations presented here and in the MEND model may explain the lack of generalization across different land cover types. Further, integrating enzymatic data sets with microbial models can lead to more robust parameterization of microbial enzymatic processes and should be a focus of future data-model integration efforts. With emerging soil microbial-ecosystem models of different model structures and complexity (Li et al., 2014), generalizing the key microbial parameters across these models is also imperative. Such an endeavor would involve exploring how microbial parameters vary across different model structures and identifying commonalities that transcend specific model architectures. By doing so, we can move towards developing standardized approaches for microbial parameterization that can be applied across multiple modeling frameworks.

Data Availability Statement

Data sets used for the modeling study are available online at ORNL TES-SFA portal (https://tes-sfa.ornl.gov/sites/default/files/Soil_Respiration_Microbial_Biomass_From_Soil_Incubations_20170519.csv). The data guide is available https://tes-sfa.ornl.gov/sites/default/files/Soil_Respiration_Microbial_Biomass_From_Soil_Incubations_20200610.pdf. The modeling results are available in Zenodo at https://doi.org/10.5281/zenodo.10576295. The MENDmult model code used in this study is publicly accessible at https://github.com/wanggangsheng/MEND_mult.git.

References

Allison, S. D., Wallenstein, M. D., & Bradford, M. A. (2010). Soil-carbon response to warming dependent on microbial physiology. *Nature Geoscience*, *3*(5), 336–340. https://doi.org/10.1038/ngeo846

Blagodatsky, S., Blagodatskaya, E., Yuyukina, T., & Kuzyakov, Y. (2010). Model of apparent and real priming effects: Linking microbial activity with soil organic matter decomposition. *Soil Biology and Biochemistry*, *42*(8), 1275–1283. https://doi.org/10.1016/j.soilbio.2010.04.005

Bond-Lamberty, B., & Thomson, A. (2010). A global database of soil respiration data. *Biogeosciences*, *7*(6), 1915–1926. https://doi.org/10.5194/bg-7-1915-2010

Chakrawal, A., Herrmann, A., Koestel, J., Jarsjo, J., Nunan, N., Kätterer, T., & Manzoni, S. (2020). Dynamic upscaling of decomposition kinetics for carbon cycling models. *Geoscientific Model Development*, *13*(3), 1399–1429. https://doi.org/10.5194/gmd-13-1399-2020

Chen, J., & Sinsabaugh, R. L. (2021). Linking microbial functional gene abundance and soil extracellular enzyme activity: Implications for soil carbon dynamics. *Global Change Biology*, *27*(7), 1322–1325. https://doi.org/10.1111/gcb.15506

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

Duan, Q., Sorooshian, S., & Gupta, V. (1992). Effective and efficient global optimization for conceptual rainfall-runoff models. *Water Resources Research*, *28*(4), 1015–1031. https://doi.org/10.1029/91wr02985

Ebrahimi, A., & Or, D. (2016). Microbial community dynamics in soil aggregates shape biogeochemical gas fluxes from soil profiles–upscaling an aggregate biophysical model. *Global Change Biology*, *22*(9), 3141–3156. https://doi.org/10.1111/gcb.13345

Fan, X., Gao, D., Zhao, C., Wang, C., Qu, Y., Zhang, J., & Bai, E. (2021). Improved model simulation of soil carbon cycling by representing the microbially derived organic carbon pool. *The ISME Journal*, *15*(8), 2248–2263. https://doi.org/10.1038/s41396-021-00914-0

Fatichi, S., Manzoni, S., Or, D., & Paschalis, A. (2019). A mechanistic model of microbially mediated soil biogeochemical processes: A reality check. *Global Biogeochemical Cycles*, *33*(6), 620–648. https://doi.org/10.1029/2018GB006077

Gao, Q., Wang, G., Xue, K., Yang, Y., Xie, J., Yu, H., et al. (2020). Stimulation of soil respiration by elevated CO_2 is enhanced under nitrogen limitation in a decade-long grassland study. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(52), 33317–33324. https://doi.org/10.1073/pnas.2002780117

Garcia-Palacios, P., Crowther, T. W., Dacal, M., Hartley, L. P., Reinsch, S., Rinna, R., et al. (2021). Evidence for large microbial-mediated losses of soil carbon under anthropogenic warming. *Nature Reviews Earth & Environment*, *2*(7), 507–517. https://doi.org/10.1038/s43017-021-00178-4

Georgiou, K., Abramoff, R. Z., Harte, J., Riley, W. J., & Torn, M. S. (2017). Microbial community-level regulation explains soil carbon responses to long-term litter manipulations. *Nature Communications*, *8*(1), 1223. https://doi.org/10.1038/S41467-017-01116-Z

German, D. P., Marcelo, K. R. B., Stone, M. M., & Allison, S. D. (2012). The Michaelis-Menten kinetics of soil extracellular enzymes in response to temperature: A cross-latitudinal study. *Global Change Biology*, *18*(4), 1468–1479. https://doi.org/10.1111/j.1365-2486.2011.02615.x

Guo, X., Gao, Q., Yuan, M., Wang, G., Zhou, X., Feng, J., et al. (2020). Gene-informed decomposition model predicts lower soil carbon loss due to persistent microbial adaptation to warming. *Nature Communications*, *11*(1), 4897. https://doi.org/10.1038/s41467-020-18706-z

Jian, S., Li, J., Wang, G., Kluber, L. A., Schadt, C. W., Liang, J., & Mayes, M. A. (2020). Multi-year incubation experiments boost confidence in model projections of long-term soil carbon dynamics. *Nature Communications*, *11*(1), 5864. https://doi.org/10.1038/s41467-020-19428-y

Kluber, L. A., Phillips, J. R., Singh, R., Sindhu, J., Wang, W., Schadt, C. W., & Mayes, M. A. (2020). Soil respiration and microbial biomass from soil incubations with 13C labeled additions [Dataset]. *Oak Ridge National Laboratory, TES SFA, U.S. Department of Energy*. https://doi.org/10.3334/CDIAC/ornlsfa.010

Li, J., Guo, C., Jian, S., Deng, Q., Yu, C.-L., Dzantor, K. E., & Hui, D. (2018). Nitrogen fertilization elevated spatial heterogeneity of soil microbial biomass carbon and nitrogen in switchgrass and gamagrass croplands. *Scientific Reports*, *8*(1), 1734. https://doi.org/10.1038/s41598-017-18486-5

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

Li, J., Wang, G., Mayes, M. A., Allison, S. D., Frey, S. D., Shi, Z., et al. (2019). Reduced carbon use efficiency and increased microbial turnover with soil warming. *Global Change Biology*, *25*(3), 900–910. https://doi.org/10.1111/gcb.14517

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

Luo, Y., Ahlstrom, A., Allison, S. D., Batjes, N. H., Brovkin, V., Carvalhais, N., et al. (2016). Toward more realistic projections of soil carbon dynamics by Earth system models. *Global Biogeochemical Cycles*, *30*(1), 40–56. https://doi.org/10.1002/2015GB005239

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

R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Schimel, J. (2023). Modeling ecosystem-scale carbon dynamics in soil: The microbial dimension. *Soil Biology and Biochemistry*, *178*, 108948. https://doi.org/10.1016/j.soilbio.2023.108948

Schimel, J., Weintraub, M. N., & Moorhead, D. (2022). Estimating microbial carbon use efficiency in soil: Isotope-based and enzyme-based methods measure fundamentally different aspects of microbial resource use. *Soil Biology and Biochemistry*, *169*, 108677. https://doi.org/10.1016/j.soilbio.2022.108677

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

Sinsabaugh, R. L., Manzoni, S., Moorhead, D. L., & Richter, A. (2013). Carbon use efficiency of microbial communities: Stoichiometry, methodology and modelling. *Ecology Letters*, *16*(7), 930–939. https://doi.org/10.1111/ele.12113

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. (2014). 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

Tang, J., & Riley, W. J. (2015). Weaker soil carbon–climate feedbacks resulting from microbial and abiotic interactions. *Nature Climate Change*, *5*(1), 56–60. https://doi.org/10.1038/nclimate2438

Tang, J., & Riley, W. J. (2019). Competitor and substrate sizes and diffusion together define enzymatic depolymerization and microbial substrate uptake rates. *Soil Biology and Biochemistry*, *139*, 107624. https://doi.org/10.1016/j.soilbio.2019.107624

Tao, X., Feng, J., Yang, Y., Wang, G., Tian, R., Fan, F., et al. (2020). Winter warming in Alaska accelerates lignin decomposition contributed by Proteobacteria. *Microbiome*, *8*(1), 1–12. https://doi.org/10.1186/s40168-020-00838-5

Tao, X., Yang, Z., Feng, J., Jian, S., Yang, Y., Bates, C. T., et al. (2024). Experimental warming accelerates positive soil priming in a temperate grassland ecosystem. *Nature Communications*, *15*(1), 1178. https://doi.org/10.1038/s41467-024-45277-0

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*(1–3), 19–33. https://doi.org/10.1007/s10533-011-9635-6

Wang, G. (2015). Multi-case version of the MEND model: MENDmult: 2015 release (version 1.1.0) [Software]. *GitHub*. https://github.com/wanggangsheng/MEND_mult.git

Wang, G., Gao, Q., Yang, Y., Hobbie, S. E., Reich, P. B., & Zhou, J. (2022). Soil enzymes as indicators of soil function: A step toward greater realism in microbial ecological modeling. *Global Change Biology*, *28*(5), 1935–1950. https://doi.org/10.1111/gcb.16036

Wang, G., Huang, W., Mayes, M. A., Liu, X., Zhang, D., Zhang, Q., et al. (2019). Soil moisture drives microbial controls on carbon decomposition in two subtropical forests. *Soil Biology and Biochemistry*, *130*, 185–194. https://doi.org/10.1016/j.soilbio.2018.12.017

Wang, G., Jagadamma, S., Mayes, M. A., Schadt, C. W., Megan Steinweg, J., Gu, L., & Post, W. M. (2015). Microbial dormancy improves development and experimental validation of ecosystem model. *The ISME Journal*, *9*(1), 226–237. https://doi.org/10.1038/ismej.2014.120

Wang, G., Jager, H. I., Baskaran, L. M., & Brandt, C. C. (2018). Hydrologic and water quality responses to biomass production in the Tennessee river basin. *Global Change Biology Bioenergy*, *10*(11), 877–893. https://doi.org/10.1111/gcbb.12537

Wang, G., Li, W., Wang, K., & Huang, W. (2021). Uncertainty quantification of the soil moisture response functions for microbial dormancy and resuscitation. *Soil Biology and Biochemistry*, *160*, 108337. https://doi.org/10.1016/j.soilbio.2021.108337

Wang, G., Mayes, M. A., Gu, L., & Schadt, C. W. (2014). Representation of dormant and active microbial dynamics for ecosystem modeling. *PLoS One*, *9*(2), e89252. https://doi.org/10.1371/journal.pone.0089252

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, K., Peng, C., Zhu, Q., Wang, M., Wang, G., Zhou, X., et al. (2019). Changes in soil organic carbon and microbial carbon storage projected during the 21st century using TRIPLEX-MICROBE. *Ecological Indicators*, *98*, 80–87. https://doi.org/10.1016/j.ecolind.2018.10.045

Wang, K., Peng, C., Zhu, Q., Zhou, X., Wang, M., Zhang, K., & Wang, G. (2017). Modeling global soil carbon and soil microbial carbon by integrating microbial processes into the ecosystem process model TRIPLEX-GHG. *Journal of Advances in Modeling Earth Systems*, *9*(6), 2368–2384. https://doi.org/10.1002/2017MS000920

Wieder, W. R., Allison, S. D., Davidson, E. A., Georgiou, K., Hararuk, O., He, Y., et al. (2015). Explicitly representing soil microbial processes in Earth system models. *Global Biogeochemical Cycles*, *29*(10), 1782–1800. https://doi.org/10.1002/2015GB005188

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

Wu, L., Zhang, Y., Guo, X., Ning, D., Zhou, X., Feng, J., et al. (2022). Reduction of microbial diversity in grassland soil is driven by long-term climate warming. *Nature Microbiology*, *7*(7), 1054–1062. https://doi.org/10.1038/s41564-022-01147-3

Ye, J. S., Bradford, M. A., Dacal, M., Maestre, F. T., & Garca-Palacios, P. (2019). Increasing microbial carbon use efficiency with warming predicts soil heterotrophic respiration globally. *Global Change Biology*, *25*(10), 3354–3364. https://doi.org/10.1111/gcb.14738

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

Zhang, X., Srinivasan, R., Zhao, K., & Liew, M. V. (2009). Evaluation of global optimization algorithms for parameter calibration of a computationally intensive hydrologic model. *Hydrological Processes*, *23*(3), 430–441. https://doi.org/10.1002/hyp.7152

Zhou, J., Xue, K., Xie, J., Deng, Y., Wu, L., Cheng, X., et al. (2012). Microbial mediation of carbon-cycle feedbacks to climate warming. *Nature Climate Change*, *2*(2), 106–110. https://doi.org/10.1038/nclimate1331

Zhu, Q., Riley, W. J., Tang, J., & Koven, C. D. (2016). Multiple soil nutrient competition between plants, microbes, and mineral surfaces: Model development, parameterization, and example applications in several tropical forests. *Biogeosciences*, *13*(1), 341–363. https://doi.org/10.5194/bg-13-341-2016