Microbial dynamics and soil physicochemical properties explain large-scale variations in soil organic carbon

Haicheng Zhang, Daniel S. Goll, Ying-Ping Wang, Philippe Ciais, William R. Wieder, Rose Z Abramoff, Yuanyuan Huang, Bertrand Guenet, Anne-katrin Prescher, Raphael A. Viscarra Rossel, et al.

DR. HAICHENG ZHANG (Orcid ID : 0000-0002-9313-5953)

DR. WILLIAM R WIEDER (Orcid ID : 0000-0001-7116-1985)

DR. BERTRAND GUENET (Orcid ID : 0000-0002-4311-8645)

PROF. RAPHAEL A. VISCARRA ROSSEL (Orcid ID : 0000-0003-1540-4748)

DR. GUOYI ZHOU (Orcid ID : 0000-0002-5667-7411)

Article type : Primary Research Articles

Microbial dynamics and soil physicochemical properties explain large scale variations in soil organic carbon

Running Title: Microbe & soil property explain SOC variation

Haicheng Zhang[1,2], Daniel S. Goll[1,3], Ying-Ping Wang[4], Philippe Ciais[1], William R. Wieder[5,6], Rose Abramoff[1], Yuanyuan Huang[1], Bertrand Guenet[1], Anne-Katrin Prescher[7], Raphael A. Viscarra Rossel[8], Pierre Barré[9], Claire Chenu[10], Guoyi Zhou[11], Xuli Tang[11]

[1]Le Laboratoire des Sciences du Climat et de l'Environnement, IPSL-LSCECEA/CNRS/UVSQ Saclay, 91191, Gif-sur-Yvette, France

[2]Department Geoscience, Environment & Society, Université Libre de Bruxelles, 1050 Bruxelles, Belgium

[3] Institute of Geography, University of Augsburg, Augsburg, Germany

[4]CSIRO Oceans and Atmosphere Private Bag 1, Aspendale, Vic 3195, Australia

[5]Climate and Global Dynamics Laboratory, National Center for Atmospheric Research, Boulder, CO,

USA

[6]*Institute of Arctic and Alpine Research, University of Colorado, Boulder, CO, USA*

[7]*Thünen Institute of Forest Ecosystems, Alfred-Möller-Straße 1, 16225 Eberswalde, Germany*

[8]*Soil & Landscape Science, School of Molecular & Life Sciences, Faculty of Science & Engineering, Curtin University, Perth, Western Australia, Australia.*

[9]*Laboratoire de Géologie de l'ENS, PSL Research University, UMR8538 du CNRS, 24 rue Lhomond, 75231, Paris, cedex 05, France*

[10]*UMR ECOSYS, INRA, AgroParisTech, Université Paris-Saclay, 78850 Thiverval-Grignon, France*

[11]*South China Botanical Garden, Chinese Academy of Sciences, Guangzhou 510650, China*

Correspondence: Haicheng Zhang (haicheng.zhang@lsce.ipsl.fr)

Type: primary research

Abstract

First-order organic matter decomposition models are used within most Earth System Models (ESMs) to project future global carbon cycling; these models have been criticized for not accurately representing mechanisms of soil organic carbon (SOC) stabilization and SOC response to climate change. New soil biogeochemical models have been developed, but their evaluation is limited to observations from laboratory incubations or few field experiments. Given the global scope of ESMs, a comprehensive evaluation of such models is essential using *in situ* observations of a wide range of SOC stocks over large spatial-scales before their introduction to ESMs. In this study, we collected a set of *in situ* observations of SOC, litterfall and soil properties from 206 sites covering different forest and soil types in Europe and China. These data were used to calibrate the model MIMICS (The MIcrobial-MIneral Carbon Stabilization model), which we compared to the widely used first-order model CENTURY. We show that, compared to CENTURY, MIMICS more accurately estimates forest SOC concentrations and the sensitivities of SOC to variation in soil temperature, clay content and litter input. The ratios of microbial biomass to total SOC predicted by MIMICS agree well with independent observations from

globally-distributed forest sites. By testing different hypotheses regarding (by using alternative process representations) of the physicochemical constraints on SOC deprotection and microbial turnover in MIMICS, the errors of simulated SOC concentrations across sites were further decreased. We show that MIMICS can resolve the dominant mechanisms of SOC decomposition and stabilization and that it can be a reliable tool for predictions of terrestrial SOC dynamics under future climate change. It also allows us to evaluate at large scale the rapidly evolving understanding of SOC formation and stabilization based on laboratory and limited filed observation.

KEYWORDS

Soil organic carbon, soil biogeochemical model, microbial physiology, soil physicochemical property, soil carbon stabilization, soil carbon classification, climate change

1 | INTRODUCTION

Soil organic carbon (SOC) is the largest terrestrial carbon (C) pool (Ciais et al., 2013), and contains more than three times as much C as either the atmosphere or terrestrial vegetation. Therefore, a small change (< 1 %) in the global SOC pool might drastically alter the land-atmosphere C balance (Heimann & Reichstein, 2008; Shi et al., 2018). SOC is also closely related to soil fertility, structure, water holding capacity and ecosystem biogeochemical cycles (Six et al., 2004; Campbell & Paustian, 2015). Dynamics of SOC have received increasing attention in many research areas ranging from small-scale projects for preserving or improving soil health, to large-scale climate change mitigation (e.g. the "4per1000" initiative) (Lal, 2016). Soil biogeochemical models are the main tools for estimating global land C stock and the interactions between SOC dynamic and changes in climate and land use.

The majority of global soil C models are developed based on first-order kinetics, in which the decomposition rate of organic matter is proportional to the pool size and turnover rate, modified by environmental factors (Parton et al., 1987; Manzoni & Porporato, 2009). These models are mathematically simple and stable, and have been proven effective for simulating soil organic matter dynamics (e.g. the decreasing trend

of remaining organic matter mass during decomposition experiments, Barré et al., 2010; Bonan et al., 2013). However, these models are unable to mechanically represent the transient SOC dynamics in response to increased fresh litter input (Fontaine et al., 2007; Guenet et al., 2010; Kuzyakov, 2010), likely because they lack explicit representation of microbial decomposition and SOC stabilization (Schmidt et al., 2011; Creamer et al., 2015). Earth System Models (ESMs) which use the first-order soil C models also show poor agreement with global spatial variation of SOC stock (Todd-Brown et al., 2013; Hararuk & Luo, 2014; Wu et al., 2018). Moreover, the conceptual SOC pools used in conventional models are largely not observable (Elliot et al., 1996; Abramoff et al., 2018; Robertson et al., 2019), making it challenging to validate conventional soil C models using field observations (Six et al., 2014; Viscarra Rossel et al., 2019).

New theories and soil biogeochemical models have been developed to explicitly represent microbial biomass and physiology (Allison, 2012; Cotrufo et al., 2013; Wieder et al., 2014b; Campbell et al., 2016; Abramoff et al., 2018, 2019; Huang et al., 2018; Robertson et al., 2019). These microbial models are valuable for testing specific responses of SOC at small spatial scales, such as the effect of short-term priming observed during litter manipulation experiments or the addition of labile organic matter to the incubated soil samples in the lab. However, they introduce parameters determined from short term experiments or under laboratory conditions. Thus, microbial models add uncertainty to large-scale simulations (Stockmann et al., 2013; Wang et al., 2014; Shi et al., 2018; Robertson et al., 2019), because most of these models are calibrated against observed litter or SOC decomposition rates obtained from limited laboratory or field experiments (Wieder et al., 2014b; Campbell et al., 2016; Georgiou et al., 2017). Robust datasets which can be used to comprehensively evaluate the simulated quasi-equilibrium SOC pool sizes are still scarce (Wieder et al., 2014a). Furthermore, it remains difficult to determine whether microbial explicit models outperform conventional first-order models on predicting large-scale SOC spatial gradients and temporal dynamics (Campbell & Paustian, 2015; Wieder et al., 2015, 2018). Microbial models have to be carefully calibrated and evaluated before

they are used to replace conventional first-order models in ESMs (Wieder et al., 2013; Wang et al., 2014).

Several studies have calibrated and validated microbial decomposition models (Wieder et al., 2013, 2015; Robertson et al., 2019) using globally gridded soil databases such as the Harmonized Word Soils Database (HWSD, FAO/IIASA/ISRIC/ISSCAS/JRC, 2012) and the Northern Circumpolar Soil Carbon Database (NCSDC, Tarnocai et al., 2009). However these global databases do not contain uncertainty estimates (Dai et al., 2018), and previous studies have identified significant differences between SOC estimates from these databases or between grid-scale estimates from these databases and point-scale *in situ* observations (Tifafi et al., 2018; Fig. S1 in supplementary material). In addition, there is still no reliable globally gridded database of plant litter input. Uncertainties in the boundary conditions (e.g. litter inputs simulated by ESMs and soil physical and chemical properties) used as model forcing data further hamper the use of these global databases for model evaluation. An alternative approach is to calibrate and evaluate the microbial-explicit SOC models using extensive in situ observations of SOC contents, soil properties, litterfall production and climate conditions. Moreover, to ensure that the tested microbial model can capture many key processes related to SOC decomposition and stabilization, rather than only simulate the total SOC contents, it is necessary to evaluate the simulated composition of different C pools to total SOC, the turnover time of each C pool, and the sensitivity of SOC content to litter input and soil properties.

In this study we compiled a large set of *in situ* observations of SOC concentrations for northern forests, as well as related soil property measurements (e.g. texture, bulk density and pH), annual litter input and climate from 206 forest sites distributed across different climate zones of Europe and China. Using this database, we calibrated and evaluated the first-order soil biogeochemical model CENTURY (Parton et al., 1987) and the microbial trait-based model MIMICS (MIcrobial-MIneral Carbon Stabilization, Wieder et al., 2015). To evaluate the simulated SOC composition, we acquired observations of the ratio of microbial biomass to total SOC,

and the SOC fractions that represent the different SOC pools in the total SOC stock from sites that are independent from the European and Chinese sites.

The aim of this study is to assess the strength and weakness of microbial implicit and microbial explicit models in simulating the stocks and composition of SOC with the intent of guiding future experiments and model developments. Specifically, we: 1) compared CENTURY and MIMICS with observed forest SOC concentrations at the continental scale, and explored the sources of model biases; 2) quantified the sensitivity of CENTURY- and MIMICS-simulated sensitivities of SOC concentration to changing soil conditions and litterfall inputs; 3) evaluated the MIMICS-simulated SOC compositions including ratios of microbial biomass to total SOC and the proportions of different SOC pools using observed values globally; 4) explored the main drivers of the variation in SOC composition. Finally we discussed the implications of our results for SOC modeling at global scales.

2 | MATERIALS AND METHODS

2.1 | Observation data on SOC concentration and soil properties

To calibrate and evaluate both soil C models under a wide range of climate conditions and forest types, we compiled observed SOC concentrations and the corresponding plant biomass, litterfall, soil properties (e.g. bulk density, soil texture, pH) and climate conditions (mean annual temperature) from 72 European forest sites and 134 Chinese forest sites (Fig. S2). The European sites are part of the International Co-operative Programme on Assessment and Monitoring of Air Pollution Effects on Forests (ICP Forests, http://icp-forests.net) operating under the UNECE Air Convention and featuring consistent methods and harmonized data across the whole network (Gleck et al., 2016; Ukonmaanaho et al.,2016). The Chinese forest sites belong to a reviewable and consistent nationwide inventory system established by the Chinese Ministry of Forestry (Tang et al., 2018). The forest stand ages at most sites are older than 40 years. *In situ* observations are mostly conducted during the period from 2000 to 2015, with durations ranging from one to more than 10 years. The observation sites cover four forest types (temperate needle-leaved evergreen forest

(TeNE), temperate broad-leaved evergreen forest (TeBE), temperate broad-leaved summer-green forest (TeBS), boreal needle-leaved evergreen forest (BoNE)) and more than 15 soil types (based on the FAO-90 soil classification in HWSD v1.2). Mean annual temperatures of the observation sites span a large range from -10 °C to higher than 20 °C (Fig. S3a). Values of mean annual total precipitation ranged from less than 300 mm yr^{-1} to more than 2000 mm yr^{-1} (Fig. S3b). Annual total litterfall production was between 100 g C m^{-2} yr^{-1} and 2000 g C m^{-2} yr^{-1} (Fig. S3c). Soil properties at the observation sites vary widely (Figs. S3d-i), with soil pH ranges from 4.5 to 8.5, and clay fraction ranges from 1% to 45%. Moreover, observation data at European ICP Forest sites provide measurements of SOC concentrations and soil properties at four different layers (0-10, 10-20, 20-40, 40-80 cm) of the top 80 cm soil, whereas data at Chinese sites provide the mean condition of the top 1 m soil.

At the European ICP Forest sites, leaf litterfall (including twig litterfall for some sites) was measured *in situ*, but not wood and root litterfall. We estimate the wood litterfall based on the ratios of wood litterfall to leaf litterfall, and the root litterfall based on the root turnover rates and the ratios of root biomass to leaf biomass (Table S1 in supplementary material). At Chinese sites, there are no *in situ* observations of litterfall. We calculated the leaf, wood and root litterfall from observed standing biomass (including leaf, wood and root) and the annual leaf and root turnover rates and the ratios of wood litterfall to leaf litterfall (Table S1). The leaf and root turnover rate, the ratios of wood and root litterfall to leaf litterfall and the ratios of root biomass to leaf biomass used in this study were obtained from a statistical analysis of extensive global observations (Zhang et al., 2014; Holland et al., 2015; Jia et al., 2016, Fig. S4).

C:N ratios of leaf litterfall at both European and Chinese sites were measured *in situ*. C:N ratios of wood and root litterfall, as well as the litterfall lignin:C ratios for each forest type were obtained from the global Fine-Root Ecology Database (FRED, Iversen et al., 2017) , the TRY database (Kattge et al., 2011) and the Long-Term Inter-site Decomposition Experiment Team (LIDET, Harmon et al., 2009).

The soil base saturation (BS, %), Cation Exchange Capacity (CEC, cmol kg^{-1})

and soil gravel content (% of volume) at each observation site were obtained from the Global Soil Dataset for Earth System Models (GSDE, Shangguan et al., 2014). Soil type was determined based on the map from HWSD v1.2. Annual mean soil water content (%) was extracted from the estimation of land surface model ORCHIDEE-trunk (r5504, Krinner et al., 2005). LAI and NDVI data were extracted from the GLASS (resolution: 0.05°, Liang et al., 2013) and GIMMS NDVI products (resolution: 8-km, Tucker et al., 2005), respectively. Evapotranspiration (ET) and the potential evapotranspiration (PET) were obtained from Jung et al. (2010) and the CRUNCEP v7 database (Viovy, 2018), respectively. More details of the datasets used in this study can be found in Table S1.

2.2 | Decomposition models

2.2.1 | CENTURY

We selected the CENTURY model (the version presented by Parton et al., 1987) to represent first-order soil biogeochemical models, because it has been widely incorporated into ESMs (e.g. Sitch et al., 2003; Krinner et al., 2005; Koven et al., 2013). In CENTURY, organic matter is separated into metabolic litter (high quality, LIT_m) and structural litter (low quality, LIT_s) and three SOC pools (active pool (SOC_{act}), slow pool (SOC_{slow}), passive pool (SOC_{pas})) with different turnover times (Fig. 1a). Fresh litter inputs are partitioned into metabolic and structural litter pools based on a linear function (f_{met}, dimensionless) of litter lignin to nitrogen (N) ratios (LN) (Parton et al., 1987):

$$f_{met} = \max(0.0, 0.85 - 0.013 \times LN) \quad (1)$$

There is no explicit representation of microbial biomass in CENTURY. The decomposition of litter and SOC is described by first order kinetics. At each daily time step, the decomposition of litter or SOC (mg C cm^{-3} day^{-1}) is calculated following:

$$\frac{\mathrm{d}C_s}{\mathrm{d}t} = I_c - k_{max} \times C_s \times f(tem) \times f(swc) \times f(clay) \quad (2)$$

where C_s (mg C cm^{-3}) is an individual litter or SOC pool, I_c (mg C cm^{-3} day^{-1}) is the C

input to the pool considered, k_{max} is the potential maximum turnover rate of C_s (day^{-1}) and is equal to the reciprocal of maximum turnover time. $f(tem)$, $f(swc)$ and $f(clay)$ are the soil temperature factor, moisture factor and clay factor modulating decomposition rate, respectively.

2.2.2 | MIMICS (default and modified versions)

The default version of MIMICS (*MIMICS-def*)

The MIcrobial-MIneral Carbon Stabilization model (MIMICS, Wieder et al., 2014b, 2015) explicitly considers the relationships among litter quality, functional tradeoffs in microbial physiology, and the physical and physicochemical protection of microbial byproducts in forming stable soil organic matter. Like CENTURY, MIMICS also has two types of litter pool: metabolic (LIT_m) and structural (LIT_s) litter (Fig. 1b), and the method used to partition fresh litter input into metabolic and structural pools (f_{met}, Fig. 1b) is the same as that used in CENTURY (Eq. 1). SOC in MIMICS is divided into three pools: the physically and physicochemically protected (SOC_p), the chemically recalcitrant (SOC_c) and available (SOC_a). Two microbial functional types are represented in MIMICS that roughly correspond to microorganisms with copiotrophic (r-strategy, MIC_r) and oligotrophic (k-strategy, MIC_k) growth strategies (Fig. 1b). The MIC_r is assumed to have higher growth and turnover rates and prefers to consume more labile litter (LIT_m), whereas the MIC_k has relatively lower growth and turnover rates and is more competitive when consuming low-quality litter (LIT_s) and chemically recalcitrant SOC (SOC_c).

C fluxes in MIMICS are simulated at an hourly (h) time step. Decomposition of litter and SOC pools (mg C cm^{-3} h^{-1}) is based on temperature-sensitive Michaelis–Menten kinetics (Schimel & Weintraub, 2003; Allison et al., 2010) through the equation:

$$\frac{dC_s}{dt} = I_c - \mathrm{MIC} \times \frac{V_{max} \times C_s}{K_m + C_s} \quad (3)$$

where C_s (mg C cm^{-3}) is a substrate pool (LIT or SOC) and MIC (mg C cm^{-3}) corresponds to the biomass of each microbial pool (MIC_r or MIC_k). I_c is the C input to the pool considered (mg C cm^{-3} h^{-1}). V_{max} and K_m are the microbial maximum

reaction velocity (mg C (mg MIC)$^{-1}$ h^{-1}) and half-saturation constant (mg C cm^{-3}), respectively. They are calculated as:

$$V_{max} = e^{V_{slope} \times T + V_{int}} \times av \times V_{mod} \quad (4)$$

$$K_m = e^{K_{slope} \times T + K_{int}} \times ak \times K_{mod} \quad (5)$$

where T is soil temperature (◦C), V_{mod} and K_{mod} represent the modifications of V_{max} and K_m based on assumptions regarding to microbial functional types, litter chemical quality and soil texture effects, av and ak are the tuning coefficient of V_{max} and K_m, respectively. V_{slope} and K_{slope} are two regression coefficients. V_{int} and K_{int} are the regression intercepts.

Decomposition rate of substrates and the microbial growth efficiency (MGE, Fig. 1b) determine the growth rate of microbes. The turnover of MIC_r and MIC_k (MIC_τ, mg C cm^{-3} h^{-1}) at each time step is calculated based on their specific turnover rate (k_{mic}, h^{-1}), annual total litterfall input (LIT_{tot}, g C m^{-2} yr^{-1}) and f_{met} by following:

$$\mathrm{MIC}_\tau = a_\tau \times k_{mic} \times e^{c \times f_{met}} \times \max\left(\min\left(\sqrt{LIT_{tot}}, 1.2\right), 0.8\right) \times \mathrm{MIC} \quad (6)$$

where a_τ (=1.0, dimensionless) is a tuning coefficient of k_{mic}. c is the regression coefficients, and its value is 0.3 for MIC_r and 0.1 for MIC_k. Turnover of microbial biomass provides C inputs to SOC pools (Fig. 1b). The fractions of microbial residues to different SOC pools are determined by soil clay content (f_{clay}) and the quality of litter inputs (lignin:N), and can be specifically calculated by following:

$$f_{rp} = \min\left(1.0,\ a_1 \times e^{1.3 \times f_{clay}}\right) \quad (7)$$

$$f_{kp} = \min\left(1.0,\ a_2 \times e^{0.8 \times f_{clay}}\right) \quad (8)$$

$$f_{rc} = \min\left(1.0 - f_{rp},\ a_4 \times e^{a_3 \times f_{met}}\right) \quad (9)$$

$$f_{kc} = \min\left(1.0 - f_{kp},\ a_5 \times e^{a_3 \times f_{met}}\right) \quad (10)$$

$$f_{ra} = 1.0 - f_{rp} - f_{rc} \quad (11)$$

$$f_{ka} = 1.0 - f_{kp} - f_{kc} \quad (12)$$

where f_{rp}, f_{kp}, f_{rc}, f_{kc}, f_{ra} and f_{ka} represent the fractions of MIC_r and MIC_k residues to SOC_P, SOC_c and SOC_a, respectively. LN is the lignin:N ratio. $a_{1\text{-}5}$ are coefficients and

their values in default MIMICS can be found in Table S1 in supplementary materials. In addition to microbial residues, a fraction of inputs ($f_{i,met}$ and $f_{i,stru}$) which bypasses litter and microbial biomass pools is transferred directly to corresponding SOC pools (Fig. 1b).

The transfer of SOC_p to SOC_a (D, mg C cm^{-3} h^{-1}), which is intended to represent the deprotection of SOC, i.e. desorption of physico-chemically protected SOC from mineral surfaces and/or the breakdown of aggregates de-protecting physically protected SOC, is calculated as a function of soil clay content (f_{clay}) by following:

$$D = 1.5 \times 10^{-5} \times k_d \times e^{-1.5 \times f_{clay}} \quad (13)$$

where k_d (=1.0, dimensionless) is a tuning coefficient of the deprotection rate. Some parameter values of the default MIMICS are provided in Table S1 in supplementary materials. Please see Wieder *et al*. (2014b, 2015) for more details of the structure, algorithms, parameters and underlying assumptions of MIMICS.

MIMICS with revised SOC deprotection rate (*MIMICS-D*)

In addition to the default version of MIMICS (MIMICS-def), we also developed and tested a new version of MIMICS (MIMICS-D) that considers the saturation of SOC protected by the mineral matrix (SOC_p). In the MIMICS-def, the deprotection rate of SOC_p in a specific soil was a fixed value determined by the abundance of the soil clay fraction (Eq. 13). However, field and laboratory research suggests that there might be an upper limit, or 'saturation level', in the amount of physicochemically and physically protected SOC that can be held in soil (Six et al., 2002; Stewart et al., 2007; Robertson et al., 2019). Deprotection rate of the SOC protected by the mineral matrix is closely related to this saturation degree (defined as the ratio of existing SOC_p to the soil maximum adsorption capacity; Kothawala et al., 2008; Wang et al., 2013). In this study, we did not calculate the maximum adsorption capacity directly, as it is determined by soil physical and chemical characteristics, and there is still no widely recognized method to calculate it (Lützow et al., 2006; Campbell & Paustian, 2015; Huang et al., 2018), The upper-limit of SOC_p was represented by assuming that the

deprotection rate increases exponentially with the pool size of SOC_p:

$$D = 1.5 \times 10^{-5} \times k_d \times e^{-1.5 \times f_{clay}} \times e^{k_{dp} \times SOC_p} \quad (14)$$

where k_{dp} is a coefficient for tuning the relationship between the deprotection rate (D) and the pool size of SOC_p.

MIMICS considering the impact of base saturation (BS) on deprotection rate (MIMICS-DB)

We tested several new modifications of MIMICS to see if the inclusion of soil chemical properties (BS and pH) could further decrease the uncertainties in simulated SOC concentrations. We modified the microbial maximum reaction velocity (V_{max}, Eq. 4), the C input rates to SOC_p (f_p and $f_{i,met}$ in Fig. 1b) and the deprotection rate of SOC_p with some simple linear or exponential functions of soil BS and pH, separately. In this study, we only present the results from the modification called MIMICS-DB, where the modified deprotection rate of SOC_p is calculated as:

$$D = 1.5 \times 10^{-5} \times k_d \times e^{-1.5 \times f_{clay}} \times e^{k_{dp} \times SOC_p} \times e^{k_{bs} \times BS} \quad (15)$$

where k_{bs} is a coefficient modifying the impacts of BS on the deprotection rate.

MIMICS considering density-dependent microbial turnover rate (MIMICS-DBT)

Following the method of Georgiou et al. (2017), we also incorporated a density-dependent microbial turnover rate into MIMICS. In this version (MIMICS-DBT), microbial turnover rate increases with growing microbial biomass density (MIC, mg C cm^{-3}) by modifying Eq. 6:

$$MIC_\tau = a_\tau \times k_{mic} \times e^{c \times f_{met}} \times \max\left(\min\left(\sqrt{LIT_{tot}}, 1.2\right), 0.8\right) \times (MIC)^{\beta} \quad (16)$$

where β is the density-dependence exponent.

2.3 | Model parameterization and validation against SOC concentrations

We assumed that all the forest sites included in this study are at steady state (i.e. no interannual variation of SOC, litterfall and stand biomass). CENTURY and the

four versions of MIMICS introduced above (Table 1) were then calibrated and evaluated against the 'equilibrium' SOC concentrations using observation data of soil texture, annual total litterfall and mean annual temperature. We also ignored the interannual and seasonal dynamics of climate and vegetation. Historical climate, litterfall input and soil properties were all assumed to be similar to the average condition during the observation period. Vertical discretization in SOC and soil properties is not considered in CENTURY and MIMICS. We focus only on the spatial variation of average SOC concentrations in the upper soil horizons (0-80 cm for European sites and 0-1 m for Chinese sites). The semi-analytic approach was used to calculate the steady state microbial and soil C pool sizes (Xia et al., 2012) based on annual total litterfall production (evenly distributed to each time step of simulation), annual mean soil temperature and moisture conditions and observed soil properties at each forest site.

Parameters of CENTURY and MIMICS were optimized against the observed SOC concentrations (Table 1). Although many parameters (e.g. carbon use efficiency and parameters related to the constraints of temperature and soil clay on C decomposition rate) of CENTURY and MIMICS can impact the simulated SOC concentrations, we only optimized the parameters which directly control the organic matter decomposition rates. Because these parameters generally contain large uncertainties and the simulated SOC stocks are generally more sensitive to these parameters than to other model parameters (Wieder et al., 2014b, 2015; Shi et al., 2018). Specifically, we added two scaling parameters k_{litt} and k_{soc} (dimensionless) in CENTURY to tune the turnover rates of litter and SOC pools, respectively.

$$k_{max\_litt\_opt} = k_{litt} \times k_{max\_litt} \quad (17)$$

$$k_{max\_soc\_opt} = k_{soc} \times k_{max\_soc} \quad (18)$$

where k_{max_litt} and $k_{max\_litt\_opt}$ are the default and optimized litter turnover rates, respectively. k_{max_soc} and $k_{max\_soc\_opt}$ are the default and optimized SOC turnover rates, respectively. The default litter and SOC turnover rates (see Table S2) were obtained from Parton et al. (1987). Optimization of only k_{litt} and k_{soc} may be not enough to minimize the uncertainties in the turnover rates of litter and SOC pools and the

simulated SOC concentrations. We therefore also tested the effectiveness of CENTURY on capturing observed SOC concentrations when five free parameters were introduced to tune the turnover rates of metabolic litter, structural litter, active SOC, slow SOC and passive SOC, respectively (Fig. S5).

For the MIMICS models, we optimized the scaling parameters (*av*, *ak* and k_d) of the microbial maximum reaction velocity (V_{max}, Eq.4), half-saturation constant (K_m, Eq. 5) and of the deprotection rate of SOC_p (Eqs. 13-15), as they are all closely related to the decomposition and the physical stabilization of organic matter (Wieder *et al*., 2014b, 2015). Parameters in the newly introduced equations (Eqs. 14-16) for modifying deprotection rates and microbial turnover rate were also optimized (Table 1).

Parameter optimization was performed using the shuffled complex evolution (SCE) algorithm developed by Duan et al. (1993, 1994), which has proven to be effective for global optimization by many previous studies (e.g. Muttil & Jayawardena, 2008; Franchini et al., 2009). Prior value and the range of each parameter used for the SCE algorithm are listed in Table S3. Root mean square error (RMSE, Eq. 19) between simulated (SOC_{sim_i}) and observed (SOC_{obs_i}) SOC concentrations (g C kg^{-1} soil) was used as the objective function, and parameters that minimized the RMSE were regarded as optimal.

$$RMSE = \sqrt{\left(\frac{\sum_{i=1}^{n}(SOC_{obs\_i} - SOC_{sim\_i})^2}{n}\right)} \quad (19)$$

where *n* is the number of observation sites. In addition to RMSE, the Akaike information criterion (AIC, Eq. 20), which considers both the goodness of fit and the number of free model parameters (n_{param}) were also used to evaluate the optimized models (Table 1).

$$AIC = n \times ln\left(\frac{\sum_{i=1}^{n}(SOC_{obs\_i} - SOC_{sim\_i})^2}{n}\right) + 2n_{param} \quad (20)$$

Our preliminary-analyses indicated that parameter optimizations of MIMICS based solely on observed SOC concentration might result in unrealistic estimates of SOC composition (e.g. the SOC_p pool approaching to zero at all sites) and of turnover

409 rates (e.g. the SOC_p turnover rates being significantly larger than SOC_a), although the
410 simulated concentrations of total SOC agreed well with the observations. To mitigate
411 this problem, some additional constraints on simulated SOC composition and turnover
412 rates were incorporated into our optimization scheme (see below). Parameter sets that
413 did not meet the imposed constraints on SOC composition and turnover rates were
414 excluded. Note that the simulated turnover rates of different SOC pools from
415 CENTURY are always consistent with the definition of SOC pools (i.e. the active
416 pool has the largest turnover rate, followed by the slow pool, and the passive pool has
417 the lowest turnover rate), and the simulated SOC composition (mainly determined by
418 the turnover rate of each pool, see section 3.2) did not show any 'abnormalities' (i.e.
419 no simulated SOC pool declined to very small values approaching zero), so we did
420 not incorporate additional constraints when optimizing the parameters of CENTURY.
421 Previous studies suggest that the organic C associated with soil minerals or stored
422 within soil aggregates, corresponding to the SOC_p pool of MIMICS, is the most stable
423 fraction of SOC with turnover times approaching hundreds to thousands of years.
424 Further, the recalcitrant SOC fractions composed by structurally complex compounds
425 corresponding to the SOC_c pool of MIMICS generally have longer turnover time than
426 the labile SOC fraction (Benbi *et al.*, 2014; Robertson *et al.*, 2019; Sokol *et al.*, 2019).
427 Therefore, we set a constraint that the simulated mean SOC_p turnover time for all of
428 the 206 observation sites must be longer than that of SOC_c, and that the mean SOC_c
429 turnover time must be longer than SOC_a.
430 Observations found that a large fraction (e.g. 10-50%) of SOC is in stable pool
431 (Lützow et al., 2007; Barré et al., 2010; Benbi et al., 2014; Viscarra Rossel et al,
432 2019). To avoid the optimized parameters giving a very low (approaching to zero)
433 estimate of the fraction of SOC_p, we also added as a constraint of model results with
434 optimized parameters that the simulated average proportion of SOC_p at the 206
435 observation sites (not for every individual site) must be larger than 5%, that average
436 proportion of SOC_c cannot exceed 70%, and that the total amount of SOC_p and SOC_c
437 should be higher than SOC_a.
438 Note that the parameters ($a_{1\text{-}5}$ in Eqs. 7-10) controlling the partition of microbial

residues to different SOC pools were modified before the parameters listed in Table 1 are optimized, because MIMICS did not give reasonable estimates of the SOC concentrations, compositions and the turnover rates simultaneously when only the parameters listed in Table 1 were calibrated. The modified values of $a_{1\text{-}5}$ are provided in Table S2.

To explore the sources of simulation errors (i.e. the difference between simulated and observed SOC concentrations), we first calculated the partial correlation coefficient between the errors of the simulated SOC concentration and different soil (e.g. texture, pH, BS and CEC), plant (NDVI and LAI) and climate (temperature, precipitation, ET) variables (see section 2.1 and Table S1 for the source of each variable). Then we fitted a linear mixed-effects (LME) model to quantify the combined contribution of the fixed-effects (soil, plant and climate variables listed above) and site-specific random-effects (e.g. soil type, forest type, stand age and micro-topography) on explaining the simulation errors. All the important variables that might potentially affect SOC dynamics, for example soil texture, temperature, pH, moisture, BS, CEC, bulk density, litterfall inputs, precipitation and ET, were included as fixed-effects in the LME. Observation site was used as a random-effect. We also fitted a multiple linear regression (MLR) with all of the fixed-effects of the LME as the predictor variables to quantify the relative contributions of fixed- and random-effects to the simulation errors. Then the relative contributions of fixed- and random-effects were quantified based on the coefficient of determination of the LME (R^2_{LME}) and MLR (R^2_{MLR}). The contributions of model choice (f_{model}), fixed-effects (f_{fixed}) and random-effects (f_{random}) to explaining the variation of SOC concentrations can be quantified by:

$$f_{model} = R^2_{model} \quad (21)$$

$$f_{fixed} = R^2_{MLR} \times \left(1 - R^2_{model}\right) \quad (22)$$

$$f_{model} = \left(R^2_{LME} - R^2_{MLR}\right) \times \left(1 - R^2_{model}\right) \quad (23)$$

where R^2_{model} is the determining coefficient of the regression equation between

simulated and observed SOC concentrations.

2.4 | Model evaluation against sensitivities of SOC concentrations to key model drivers

To assess whether each model simulated the variations of SOC concentrations for the right reasons, we first identified the key drivers of the spatial variations of SOC concentration, and then compared modeled sensitivities of SOC concentration to these drivers to the values derived from the observations. The potential key drivers we evaluated include soil temperature, moisture, clay content, litterfall input, the mean C:N ratio and the lignin:C ratio of litterfall. The sensitivities of organic matter decomposition rate to manipulated soil temperature, moisture and litter inputs have been widely investigated via laboratory and field experiments (Parton et al., 2007; Bonan et al., 2013; Sierra et al., 2015). However, no experiments have measured the sensitivity of equilibrium SOC stock to changing soil properties and litter inputs, as it would take decades to hundreds of years for the SOC pool to reach equilibrium after manipulating litter. Here we estimated the sensitivities by making use of observed spatial variation of SOC with different drivers, including soil temperature, water content, clay fraction, annual total litter input and the C:N ratio and lignin:C ratio of litter input. We assumed the soil-litter system is in steady-state, and the sensitivities of equilibrium SOC to different drivers were quantified by multiple linear regression. The regression coefficient of each driver was regarded as the observed sensitivity.

The sensitivities of simulated SOC concentration to soil and litter properties from optimized CENTURY and MIMICS were obtained using Monte Carlo simulations. We sampled 1000 sets of unique soil and litter input condition within the observed space of each variable using Latin Hypercube technique (Tang & Zhuang, 2009). All soil and litter variables were assumed to be uniformly distributed and the range of each variable was set based on the maximum and minimum observed values at the European and Chinese sites. For each combination of soil and litter input condition, the sensitivity (S_i) of SOC concentration to each variable (d_i) was calculated as

$$S_i = \frac{f(d_1,d_2,\cdots,d_i+\delta,\cdots d_n) - f(d_1,d_2,\cdots,d_i,\cdots d_n)}{\delta}$$

497 (24)

498 where δ is the step size of a change in variable d_i assumed to be one percent of the
499 difference between maximum and minimum d_i (i.e. $\delta = (d_{i\_max}\text{-}d_{i\_min})/100$).

500

501 **2.5 | Model evaluation against SOC composition**

502 We evaluated the simulated proportions of the different SOC pools using
503 observations from sites that are independent of the European and Chinese forest sites,
504 for which the model parameters were calibrated. The simulated ratios of microbial
505 biomass to total SOC were validated against 655 observations from forest sites around
506 the world (Xu et al., 2013). The simulated SOC composition from CENTURY and
507 MIMICS was compared to measurements of SOC composition from 505 sites under
508 native forests and grasslands in Australia (Viscarra Rossel & Hicks, 2015; Viscarra
509 Rossel et al. 2019). These data were partitioned into three fractions, the particulate
510 organic C (POC), humic organic C (HOC) and resistant organic C (ROC, which is the
511 mineral-associated organic carbon) based on the particle size and chemical
512 compositions of organic matter. We acknowledge the fact that the observed pools are
513 not modeled conceptual pools and we propose a correspondence between both in
514 Table S4. We compared the simulated SOC pools to the observed SOC fractions to
515 assess their correspondence in terms of their expected/assumed turnover rates.

516 **2.6 | Model evaluation against the key drivers of variations in SOC composition**

517 To determine whether the key drivers of variations in SOC composition in
518 MIMICS and CENTURY models are consistent with the observations, we calculated
519 the partial correlation coefficient between fraction of each SOC pool and different
520 model drivers using the simulated proportions of different SOC pools by optimized
521 MIMICS and CENTURY models at all of the 206 forest sites in Europe and China
522 (Fig. S2), and using the observed proportions of different SOC pools at the 505
523 Australia sites (Viscarra Rossel et al. 2019). The key drivers we considered in this
524 analysis include soil temperature, moisture, clay fraction, BS, annual litterfall input,
525 litter C:N and lignin:C ratios and the total SOC pool size). For each model driver, all
526 of the other drivers described above were used as the controlling factor for calculating

the partial correlation coefficient.

3 | RESULTS

3.1 | Evaluation of simulated SOC concentrations

Our evaluation indicates that MIMICS can better capture the observed spatial variation of SOC concentrations than CENTURY across European and Chinese forest sites. The default version MIMICS-def explains 48% observed SOC spatial variation, as compared to only 10% by CENTURY model (Fig. 2). MIMICS-D, MIMICS-DB and MIMICS-DBT explain 52%, 57% and 59% SOC spatial variation, respectively (Fig. 2). The RMSE and Akaike information criterion (AIC) indicate that all MIMICS versions estimate the spatial variation of SOC concentration more accurately than CENTURY, with MIMICS-DBT having the best performance overall (Fig. 2f). We also note that the CENTURY model with 5 free parameters for tuning turnover rates of litter and SOC pools (Fig. S5a) does not estimate SOC concentrations more accurately than the CENTURY with 2 free parameters (Table 1). CENTURY with 5 free parameters has a slightly smaller RMSE (16.89) but a higher AIC (1174.7) than the RMSE (16.97) and AIC (1170.5) respectively from CENTURY with 2 parameters (Fig. S5a).

There are systematic biases in the simulated SOC concentrations along the gradients of SOC pool size, soil properties, and climate and plant variables (Figs. 3 and S6). Both CENTURY and MIMICS overestimate the low SOC concentrations but underestimate the high concentrations (Figs. 2 and S6). The simulation biases of CENTURY are significantly correlated with soil (e.g. moisture, BS, pH, and bulk density), plant (e.g. litterfall, LAI) and climate (e.g. mean annual temperature and annual total precipitation) variables (Fig. 3), suggesting that CENTURY has structural biases in the processes depending upon those factors. Similar to CENTURY, the simulation bias of MIMICS is also significantly correlated with some soil and litterfall-related variables. By including the effect of BS on deprotection rate into MIMICS (MIMICS-DB), the significant relationships between simulation biases and soil, plant and climate variables are largely eliminated, but a significant negative

relationship between simulation biases and soil CEC appears. The significant relationship between simulation biases and annual litterfall input can be eliminated only when the density dependence of microbial turnover rate in MIMICS-DBT is represented. Moreover, the simulation biases of all models are positively related to soil bulk density (Fig. 3).

Soil properties, litter input rate and the plant and climate conditions together can only explain a small portion of the simulation biases in SOC concentrations, especially for MIMICS (Figs. S7, S8). The linear mixed-effects (LME) models which consider both fixed factors (i.e. the soil, litter and climate variables) and site-specific random factor (e.g. soil type, forest type, stand age and micro-topography) explain most of the variations in the simulation biases (Fig. S7). Further statistics indicated that the SOC variation explained by CENTURY, fixed factors and random factors are 10%, 27% and 54%, respectively (Fig. S8). But for MIMICS, the model itself explained the largest part (48-59%) of SOC variation, followed by the random factor (24-32%), with fixed factors explaining 5-9% of SOC variation (Fig. S8). Our further analysis on the potential contributors to random factors indicated that CENTURY estimations of SOC are consistently biased regardless of soil type, plant type and stand age (Fig. S9). But the estimations of SOC made by MIMICS are, with few exceptions, unbiased across sites with different soil types, plant types and stand ages. Overall, the constraints of soil, litter and climate factors on SOC stocks are significantly better represented in MIMICS than in CENTURY.

3.2 | Evaluation of simulated sensitivities of SOC concentration to key model drivers

Based on observations, SOC concentrations are sensitive to local soil temperature and soil clay content (Figs. 4a, c), but are not sensitive to local soil moisture or litter quantity and quality (Figs. 4b, d, e, f). On average, SOC concentration declines by 0.53 g C kg^{-1} soil with a 1 °C increase in soil temperature, and increases by 0.37 g C kg^{-1} soil with a 1 percent increase in soil clay fraction.

MIMICS models provide more accurate estimates of the observation-based partial sensitivity of SOC concentration to changes in soil temperature, compared to

CENTURY (Fig. 4a). With a 1 °C increase in soil temperature, the simulated SOC concentration declines by 0.4-0.55 g C kg^{-1} soil (median value) depending on the version of MIMICS. The sensitivity is comparable to the value calculated based on observation data, but significantly lower than the value simulated by CENTURY (-0.92±4.1 g C kg^{-1} soil $°C^{-1}$). Both CENTURY and MIMICS underestimate the observed sensitivity of SOC to soil clay fraction. Despite this, the sensitivities estimated by MIMICS (0.17 – 0.26 g C kg^{-1} soil $(clay\%)^{-1}$) are closer to the observed value than CENTURY (0.02 g C kg^{-1} soil $(clay\%)^{-1}$, Fig. 4c). In CENTURY or MIMICS, the sensitivities of SOC concentration to these variables generally show large variations. Overall, SOC simulated by CENTURY is more sensitive to the changes in soil condition and litter input than MIMICS.

3.3 | Evaluation of simulated SOC composition

The simulated ratios of microbial biomass (MIC) to total SOC stock (MIC/SOC) from the MIMICS models is broadly consistent with the observations collected from global forest sites (Xu *et al*., 2013), both in terms of mean (or median) value and the range of variation (Fig. 5). Overall, both observed and simulated MIC/SOC ranged from 0.005 to approximately 0.05, with a mean value of approximately 0.017 (0.015-0.019) and a median value of 0.013 (0.012-0.014).

MIMICS simulated fractions of SOC pools that are consistent with measurements of the Australian soil samples based on the particle size and chemical compositions of organic matter (Table S4), but CENTURY did not (Fig. 6). Observations at 505 Australian sites indicate that HOC (46-60%) accounts for the largest proportion of SOC, followed by the most stable pool ROC (25-33%). The labile pool POC makes up a small fraction (12-23%) of total SOC (Fig. 6a). MIMICS predicts a similar composition of SOC pools. The moderately stable pool (SOC_c) accounts for the largest proportion of total SOC, followed by the most stable pool protected by the mineral matrix (SOC_p), and the available pool (SOC_a, Fig. 6c). SOC composition simulated by CENTURY can be very different depending on the optimized turnover rates of the active, slow and passive SOC pools (Figs. 6b and S5b). Increasing turnover rate of a specific SOC pool generally results in a smaller proportion of this

pool compared to the total SOC (if the turnover rates of other SOC pools are assumed to be fixed).

3.4 | Key drivers of the variation in SOC composition

The key factors controlling the simulated SOC composition in CENTURY and MIMICS are different from the observations (Fig. 7). Based on observation data, soil moisture, clay fraction, BS and litter input show significant empirical correlations with SOC composition, whereas soil temperature shows no significant correlation. In both CENTURY and MIMICS, soil temperature strongly affects SOC composition. Higher temperature however decreases the 'stable' SOC fraction (SOC_{pas}) in CENTURY, but increases the stable fraction (SOC_p) in MIMICS. MIMICS can represent the impacts of litter input on SOC composition, but CENTURY does not. Similar to the observations, higher litter input rate increases the proportion of the stable SOC pools (ROC and SOC_p) but decreases the proportion of moderately stable pools (HOC and SOC_c). The simulated decreasing trend of labile SOC (SOC_a) with increasing litter input is contrary to the observation (POC). In MIMICS-DB and MIMICS-DBT, soil chemical properties represented by BS also show strong impact on SOC composition. Moreover, SOC composition also changes with the pool size of total SOC. It is necessary to note that the partial correlation coefficients might not be able to fully represent the relationships between SOC composition and soil and litter variables (Fig. 7), as SOC composition might not be linearly related to these variables (Fig. S10).

4 | DISCUSSION

Using *in situ* observations of SOC, litterfall and soil properties from 206 forest sites in Europe and China, we compared the performance of a first-order soil biogeochemical model (CENTURY) and four different versions of the microbial trait-based model (MIMICS) for simulating the large-scale spatial variation of SOC concentrations, the sensitivity of SOC concentration to key model drivers and the SOC composition. Our evaluation provides strong evidence that soil biogeochemical models with explicit microbial processes can be applied to simulate the large-scale SOC dynamics across different soil, vegetation and climate conditions. Below we

discuss in detail the implications of these results, uncertainties associated with the analysis, and an outlook for future data and model needs.

4.1 | Implications of simulation results

4.1.1 | Decomposition model should be calibrated and evaluated comprehensively

This study reveals the necessity to calibrate and evaluate MIMICS comprehensively. Preliminary parameter estimates for this study showed that although parameters optimized based solely on observed SOC concentrations can accurately estimate total SOC stocks; they may not be able to estimate SOC composition and turnover time. In order to avoid unreasonable estimates of SOC composition (e.g. SOC_p of MIMICS calibrated only against the SOC concentrations at European and China forest sits always approaches to zero) and C turnover times, we imposed additional constraints to restrict the ranges of proportions and turnover times of MIMICS SOC pools (see section 2.3). Our results highlight the need for comparing model results with total SOC and microbial biomass, SOC composition and turnover time, as well as the response of SOC to changed climate, litter input and soil properties with a wide range of observations. Moreover, the optimized parameter values of both CENTURY and MIMICS in this study (Table S3) are different from the default values calibrated against manipulated decomposition experiments (Parton et al., 1987; Wieder et al., 2015), suggesting that model parameters obtained based on local decomposition experiments might not work well at large spatial scales.

4.1.2 | Importance of explicitly representing microbial dynamics in decomposition model

Explicit representation of microbial biomass and substrate-limited growth rates is important for soil biogeochemical models to accurately capture the observed SOC concentration variations and the responses of SOC to climate changes (Wieder et al., 2014b; Campbell & Paustian, 2015). In our research, simulations of SOC concentration at forest sites using MIMICS were more accurate and parsimonious compared to using CENTURY (Fig. 2), and MIMICS better capture the observed sensitivities of SOC concentrations to temperature and soil clay than CENTURY. Conventional first-order models do not explicitly simulate microbial activity, but

instead strongly emphasizes the relationship between litter chemical recalcitrance and soil C stock (Jenkinson & Rayner, 1977; Parton et al., 1987; Wieder et al., 2014b). Recent analytical and experimental advances have demonstrated that molecular structure alone does not control SOC stability. Rather, microbial products of decomposition are the main precursors of stable SOC (Cotrufo et al., 2013; Kallenbach et al., 2016), suggesting that, in fact, environmental and biological controls predominate (Lützow et al., 2006; Schmidt et al., 2011; Lehmann & Kleber, 2015).

4.1.3 | Impacts of soil physiochemical properties on SOC decomposition and stabilization

Besides microbial dynamics, it is also necessary to accurately represent the effects of soil physiochemical properties on SOC dynamics in soil biogeochemical models, especially for the formation and release of SOC protected by the mineral matrix. It has been widely recognized that soil clay fractions can influence SOC stock and stabilization by promoting the sorption of organic C to mineral surfaces and entrapment into micropores (Schimel et al., 1994; Wagner et al., 2007). CENTURY uses the soil clay fraction to modify the decomposition rate of the active SOC pool and the C transfer from active to slow pool (Parton et al., 1987). As the active pool generally accounts for only a small fraction (c.a. 3.5%) of total SOC (Fig. 6b), this might explain why the sensitivity of SOC concentration to soil clay content in CENTURY is drastically underestimated compared to the observation-based sensitivity (Fig. 4c). In MIMICS, soil clay influences both the decomposition rate of available SOC pool and the deprotection rate of protected by the mineral matrix. MIMICS thus better represents current understanding of SOC stabilization processes and appears to more accurately estimate the sensitivity of SOC to soil clay fraction than CENTURY (Fig. 4c).

Numerous experimental studies also reported the significant impacts of soil chemical properties such as pH, exchangeable cations (e.g. Ca^{2+}) and extractable metals (e.g. iron- and aluminum-oxyhdroxides) on SOC dynamics (Six et al., 2004; Doetterl et al., 2015; Rasmussen et al., 2018; ViscarraRossel et al., 2019), and the

relative importance of these factors likely varies across scales and ecosystems (Jobbágy & Jackson, 2000; Schmidt et al., 2011; ViscarraRossel et al., 2019). Indeed, representing the diversity of mechanisms by which the soil physicochemical environment influences the persistence of soil organic matter in numerically tractable ways remains an outstanding challenge in models (Bailey et al. 2018). Our work suggests one opportunity to use base saturation (BS) as a proxy variable that can modify C deprotection rates from the SOC_p pool in MIMICS (MIMICS-DB). This modification significantly decreased the biases in simulated SOC concentrations (Fig. 2) and eliminated the systematic estimation biases along gradients of soil pH, clay content and annual precipitation at the observation sites (Fig. 3). Moreover, our analysis on the relative contributions of model choice, fixed effects and site-specific random effects to explaining the SOC variation (Fig. S8) reveals that the constraints of soil physical (e.g. temperature and clay content) and chemical (e.g. BS) properties on SOC dynamics has been better represented in MIMICS than in CENTURY, as the fixed effects including all potentially important soil variables can only explain a small part of the simulation errors of MIMICS, but a considerable part (~ 30%) of the simulation errors of CENTURY (Fig. S7, S8).

4.1.4| Impacts of litter inputs on SOC decomposition and stabilization

First order models like CENTURY assume a linear relationship with productivity and soil C stocks (Todd-Brown et al. 2013), and the same is true for default parameterizations of MIMICS. Our analysis shows that the simulated SOC concentrations from CENTURY and MIMICS models are systematically biased from observations along the gradients of local litterfall production, except for the MIMICS-DBT which considers the density-dependent turnover of microbes (Fig. 3). This suggests that at the community level, regulatory mechanisms like competition, space constraints and other controls that depend on the density of individuals (such as disease and production of toxins) may limit microbial population sizes (Hibbing et al., 2010; Kaiser et al., 2014; Kaiser et al., 2015) Indeed, a recent study from Georgiou et al. (2017) indicated that the density-dependent microbial processes can play an essential, but often overlooked role in regulating SOC dynamics. We recognize that

the parameterization of density dependent turnover implemented in MIMCS-DBT simplifies the complex community interactions that occur in soils, but they represent a tractable means for capturing the emergent dynamics in models that are intended for global-scale application and projections.

Litter input is not as important as soil physicochemical properties for predicting total SOC stock (Fig. 4d), but it nevertheless strongly affects SOC composition (Fig. 7), which determines the vulnerability of SOC (i.e. risk of C loss) to perturbations such as climate change and human disturbances. Litter quality can impact microbial C use efficiency and short-term SOC dynamics (Manzoni et al., 2017; Zhang et al., 2018), but evidence is inclusive on the significant role of litter quality in long-term SOC dynamics (Helfrich et al., 2008; Gentile et al., 2011). The effect of litter quality on SOC stabilization is mostly modulated by the extent of soil C saturation, and it may alter SOC stocks only when there is a saturation deficit (Castellano et al., 2015). Consistent with our results (Fig 7), previous studies also reported that litter quantity rather than quality is one of the main determinants of SOC stability (Carrington et al., 2012; Dungait et al., 2012). Experiments by Wang et al. (2016) suggested that the ratio between different SOC fractions is related to microbial biomass and community composition (which depends on the amounts of litter inputs), but not to litter chemical composition.

4.2 | Uncertainties in this study

Some uncertainties in our simulation results may be caused by biases of forcing and validation data. In this study, we assumed the forest and soil C at all observation sites are at equilibrium. However, even though most observation sites have a stand age older than 40 years and have not been strongly disturbed by fire or human activities (e.g. reforestation and deforestation can induce a 30% change in soil C stock, Don et al., 2011), the forest systems at some sites may not be at equilibrium, especially under the background of global climate change. Some uncertainties also arise due to lack of observations. Specifically, the wood and root litterfall at European sites have not been measured and Chinese observation data only provides

767 measurements of plant biomass but not litterfall, so we have adopted the leaf turnover
768 rates and ratios of wood litter and root litter to leaf litter from databases of plant traits
769 and litterfall production to calculate the total litterfall production at each observation
770 site (see section 2.1). Moreover, most of the litter C:N ratios and the lignin:C ratios
771 were obtained from previously compiled litterfall databases and publications and not
772 from site level observations, Thus, biases and uncertainties that exist in the litter input
773 data are poorly quantified.

774 Additional uncertainties are related to model structural assumptions and
775 parameterizations. Specifically, soil moisture has been widely regarded as one of the
776 primary physical factors that control microbial activity (Arnold et al., 2015; Manzoni
777 et al., 2016; Ghezzehei et al., 2019); however the soil moisture control over microbial
778 dynamics is not used in the current parameterization of MIMICS. Soil structure
779 (characterized by porosity or bulk density) determines soil O_2 availability and the
780 accessibility of C particles to microbes (Lützow et al., 2006; Davidson et al., 2012).
781 Soil nutrient availability (e.g. mineral nitrogen and phosphorus) strongly affects
782 microbial C use efficiency and growth rate (Manzoni *et al*., 2017). Again, soil
783 moisture, structure and nutrient availability have not been considered in this
784 implement MIMICS. Finally, neither of the models considered here implement
785 vertically resolved soil biogeochemistry, which are clearly important to capture soils
786 with strong vertical profiles or vertical perturbations such as in permafrost C (Kovenet
787 al. 2015; McGuire et al. 2018). The insufficient representation of interactions between
788 soil physicochemical properties, nutrient availability, microbial dynamics and SOC
789 stabilization therefore may induce additional uncertainties in our results. We
790 appreciate that these additional complexities in model form also generates greater data
791 demands to appropriately parameterize and evaluate models, but may be necessary to
792 build confidence in soil carbon projections (Bradford *et al*. 2016).

793 **4.3 | Outlooks and challenges**

794 A study by Wieder et al. (2014) demonstrated that MIMICS could capture the
795 observed temporal decreasing trends of litter and SOC stocks in field decomposition
796 experiments. Our evaluation further demonstrates that MIMICS can simulate SOC

stock and composition across ecosystems with different climate, and soil and forest types. MIMICS also represents the SOC decomposition and stabilization processes more realistically (e.g. explicitly represents microbial dynamics) than conventional first-order models. Therefore MIMICS can be used to replace the conventional decomposition models used in existing ESMs.

The parameters, structure and algorithms of MIMICS can still be improved. We encourage future studies to assess the global applicability of MIMICS or similar models based on more integrated *in situ* observations on plant biomass, litterfall (both aboveground and belowground), SOC stock and composition, soil physicochemical properties and local climate from more ecosystems, in particular observations from grasslands and tropical forests. We also encourage more studies to quantify the interactions between soil physicochemical properties, microbial dynamics and the stabilization of SOC. In this study, the MIMICS model considering the physicochemical constraints of soil properties on SOC deprotection rate and microbe turnover more accurately estimated SOC concentration than the default model (Fig. 2). But the empirical functions (Eqs. 13, 14) used to represent physicochemical constraints were built empirically based on analysis of the biases of simulated SOC concentration from the default version of MIMICS (Fig. 3). More experiments investigating influences of soil physicochemical properties on microbial activity and the C adsorption/desorption rate of mineral soil are needed to improve these empirical functions. Furthermore, many soil properties are significantly correlated (e.g. Fig. S12) and the changes in litter inputs and SOC contents can in return dramatically alter soil physical, chemical, and biological properties (Schmidt et al., 2011; Murphy et al., 2015). Thus, research focusing on the interactions between litter, SOC and different soil properties is also essential.

ACKNOWLEDGEMENTS

HZ, DSG, PC, YPW, RA, YH and BG acknowledge the IMBALANCE-P project of the European Research Council (ERC-2013-SyG- 610028). HZ acknowledges the 'Lateral-CNP' project (No. 34823748) of the Fonds de la Recherche Scientifique

827 –FNRS. YPW acknowledge the financial support by the National Environmental
828 Science Program Earth System and Climate Change. WRW is supported by the US
829 Department of Energy under award number BSS DE-SC0016364, US Department of
830 Agriculture NIFA 2015-67003-23485, and NASA Interdisciplinary Science Program
831 award number NNX17AK19G. We acknowledge the contribution of Shijie Han,
832 Sheng Du, Shenggong Li, Keping Ma, Junhua Yan, Youxin Ma and Genxu Wang
833 from the Chinese Academy of Sciences to collecting the Chinese observation data.
834 We also acknowledge the collection of data by partners of the official UNECE ICP
835 Forests Network (http://icp-forests.net/contributors). Part of the data was co-financed
836 by the European Commission (Data achieved at 23/04/2018).
837
838 **Data accessibility**: The European ICP forest data can always be requested from the
839 Programme Co-ordinating Centre (PCC, http://icp-forests.net/page/data-requests) of
840 ICP Forests in Eberswalde, Germany. The Chinese forest data can be obtained by
841 contacting the Prof. Tang X (xltang@scib.ac.cn) in South China Botanical Garden,
842 Chinese Academy of Sciences, Guangzhou, China. All of the other databases of soil,
843 climate, litterfall and vegetation are publicly accessible, and the specific references
844 and links to these databases are provided in section 2.1.
845
846 **REFERENCES**

847 Abramoff, R., Xu, X., Hartman, M., O'Brien, S., Feng, W., Davidson, E., Finzi, A., Moorhead, D., Schimel, J.,
848 Torn, M. & Mayes, M.A. (2018). The Millennial model: in search of measurable pools and
849 transformations for modeling soil carbon in the new century. *Biogeochemistry*, 137, 51-71.
850 Abramoff, R. Z., Torn, M. S., Georgiou, K., Tang, J. & Riley, W. J. (2019). Soil organic matter temperature
851 sensitivity cannot be directly inferred from spatial gradients. *Global Biogeochemical Cycles*.
852 doi:10.1029/2018gb006001
853 Akaike, H. (1974). A new look at the statistical model identification, IEEE T. on Automat. Cont., 19, 716–723.
854 Allison, S. D. (2012). A trait-based approach for modelling microbial litter decomposition. *Ecology Letters*, 15,
855 1058-1070.
856 Allison, S. D., Wallenstein, M. D. & Bradford, M. A. (2010). Soil-carbon response to warming dependent on

microbial physiology. *Nature Geoscience*, 3, 336-340.

Arnold, C., Ghezzehei, T. A. & Berhe, A. A. (2015). Decomposition of distinct organic matter pools is regulated by moisture status in structured wetland soils. *Soil Biology and Biochemistry*, 81, 28-37.

Bailey, V. L., Bond-Lamberty, B., De Angelis, K., Grandy, A. S., Hawkes, C. V., Heckman, K., . . . Wallenstein, M. D. (2018). Soil carbon cycling proxies: Understanding their critical role in predicting climate change feedbacks. *Global Chang Biology*, 24, 895-905.

Barré, P., Eglin, T., Christensen, B. T., Ciais, P., Houot, S., Kätterer, T., van Oort, F., Peylin, P., Poulton, P. R., Romanenkov, V. & Chenu, C. (2010). Quantifying and isolating stable soil organic carbon using long-term bare fallow experiments. *Biogeosciences*, 7, 3839-3850.

Benbi, D. K., Boparai, A. K. & Brar, K. (2014). Decomposition of particulate organic matter is more sensitive to temperature than the mineral associated organic matter. *Soil Biology and Biochemistry*, 70, 183-192.

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 Chang Biology*, 19, 957-974.

Bradford, M. A., Wieder, W. R., Bonan, G. B., Fierer, N., Raymond, P. A., & Crowther, T. W. (2016). Managing uncertainty in soil carbon feedbacks to climate change. *Nature Climate Change, 6*, 751-758.

Campbell, E. E. & Paustian, K. (2015). Current developments in soil organic matter modeling and the expansion of model applications: a review. *Environmental Research Letters*, 10, 123004.

Campbell, E. E., Parton, W. J., Soong, J. L., Paustian, K., Hobbs, N. T. & Cotrufo, M. F. (2016). Using litter chemistry controls on microbial processes to partition litter carbon fluxes with the Litter Decomposition and Leaching (LIDEL) model. *Soil Biology and Biochemistry*, 100, 160-174.

Carrington, E. M., Hernes, P. J., Dyda, R. Y., Plante, A. F. & Six, J. (2012). Biochemical changes across a carbon saturation gradient: Lignin, cutin, and suberin decomposition and stabilization in fractionated carbon pools. *Soil Biology and Biochemistry*, 47, 179-190.

Castellano, M. J., Mueller, K. E., Olk, D. C., Sawyer, J. E. & Six, J. (2015) Integrating plant litter quality, soil organic matter stabilization, and the carbon saturation concept. *Global Chang Biology*, 21, 3200-3209.

Ciais, P., Sabine, C., Bala, G., Bopp, L., Brovkin, V., Canadell, J., Chhabra, A., DeFries, R., Galloway, J., Heimann, M., Jones, C., Le Quéré, C., Myneni, R. B., Piao, S. L. & Thornton, P. (2013). Carbon and Other Biogeochemical Cycles. In: *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* eds.

T. F. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and M. P.M.). Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Cools, N. & De Vos, B. (2016). Part X: Sampling and Analysis of Soil. In: UNECE ICP Forests Programme Coordinating Centre (ed.): Manual on methods and criteria for harmonized sampling, assessment, monitoring and analysis of the effects of air pollution on forests. In, Thünen Institute of Forest Ecosystems, Eberswalde, Germany.

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? *Glob Chang Biol*, 19, 988-995.

Creamer, C. A., de Menezes, A. B., Krull, E. S., Sanderman, J., Newton-Walters, R. & Farrell, M. (2015). Microbial community structure mediates response of soil C decomposition to litter addition and warming. *Soil Biology and Biochemistry*, 80, 175-188.

Dai, Y., Shangguan, W., Wang, D., Wei, N., Xin, Q., Yuan, H., Zhang, S., Liu, S. & Yan, F. (2018) A review on the global soil datasets for earth system modeling. *SOIL Discussions*, 1-30.

Davidson, E. A., Samanta, S., Caramori, S. S. & Savage, K. (2012). The Dual Arrhenius and Michaelis-Menten kinetics model for decomposition of soil organic matter at hourly to seasonal time scales. *Global Change Biology*, 18, 371-384.

Doetterl, S., Stevens, A., Six, J., Merckx, R., Van Oost, K., Casanova Pinto, M., Casanova-Katny, A., Muñoz, C., Boudin, M., Venegas, E. & Boeckx, P. (2015). Soil carbon storage controlled by interactions between geochemistry and climate. *Nature Geoscience*, 8, 780–783.

Don, A., Schumacher, J. & Freibauer, A. (2011). Impact of tropical land-use change on soil organic carbon stocks – a meta-analysis. *Global Change Biology*, 17, 1658-1670.

Duan, Q., Gupta, V. & Sorooshian, S. (1993). Shuffled complex evolution approach for effective and efficient global minimization. *Journal of Optimization Theory and Its Applications*, 76, 501-521.

Duan, Q., Sorooshian, S. & Gupta, V. K. (1994). Optimal use of the SCE-UA global optimization method for calibrating watershed models. *Journal of Hydrology*, 158, 265-284.

Dungait, J. A. J., Hopkins, D. W., Gregory, A. S. & Whitmore, A.P. (2012). Soil organic matter turnover is governed by accessibility not recalcitrance. *Global Change Biology*, 18, 1781-1796.

Elliott, E. T., Paustian, K. & Frey, S. D. (1996). Modeling the Measurable or Measuring the Modelable: A Hierarchical Approach to Isolating Meaningful Soil Organic Matter Fractionations. In: Powlson D. S.,

917 Smith P., Smith J. U. (eds) *Evaluation of Soil Organic Matter Models. NATO ASI Series* (*Series I:*
918 *Global Environmental Change*), vol 38. Springer, Berlin, Heidelberg.
919 FAO/IIASA/ISRIC/ISSCAS/JRC (2012). Harmonized World Soil Database (version 1.2). In, FAO, Rome, Italy
920 and IIASA, Laxenburg, Austria.
921 Fontaine, S., Barot, S., Barré, P., Bdioui, N., Mary, B. & Rumpel, C. (2007). Stability of organic carbon in deep
922 soil layers controlled by fresh carbon supply. *Nature*, 450, 277-280.
923 Frankenberg, C., Fisher, J. B., Worden, J., Badgley, G., Saatchi, S.S., Lee, J. E., Toon, G. C., Butz, A., Jung, M. &
924 Kuze, A. (2011). New global observations of the terrestrial carbon cycle from GOSAT: Patterns of plant
925 fluorescence with gross primary productivity. *Geophysical Research Letters*, 38, 351-365.
926 Gentile, R., Vanlauwe, B. & Six, J. (2011). Litter quality impacts short- but not long-term soil carbon dynamics in
927 soil aggregate fractions. *Ecological Applications*, 21, 695-703.
928 Georgiou, K., Abramoff, R. Z., Harte, J., Riley, W. J. & Torn, M. S. (2017). Microbial community-level regulation
929 explains soil carbon responses to long-term litter manipulations. *Nature Communications*, 8, 1223.
930 Ghezzehei, T. A., Sulman, B., Arnold, C. L., Bogie, N. A. & Berhe, A. A. (2019). On the role of soil water
931 retention characteristic on aerobic microbial respiration. *Biogeosciences*, 16, 1187-1209.
932 Hararuk, O. & Luo, Y. (2014). Improvement of global litter turnover rate predictions using a Bayesian MCMC
933 approach. *Ecosphere*, 5, art163.
934 Harmon, M. E., Silver, W. L., Fasth, B., Chen, H. U. A., Burke, I. C., Parton, W. J., Hart, S. C. & Currie, W. S.
935 (2009). Long-term patterns of mass loss during the decomposition of leaf and fine root litter: an intersite
936 comparison. *Global Change Biology*, 15, 1320-1338.
937 Hartono, A., Funakawa, S. & Kosaki, T. (2005). Phosphorus Sorption-Desorption Characteristics of Selected Acid
938 Upland Soils in Indonesia. *Soil Science and Plant Nutrition*, 51, 787-799.
939 Heimann, M. & Reichstein, M. (2008). Terrestrial ecosystem carbon dynamics and climate feedbacks. *Nature*, 451,
940 289-292.
941 Helfrich, M., Ludwig, B., Potthoff, M. & Flessa, H. (2008). Effect of litter quality and soil fungi on
942 macroaggregate dynamics and associated partitioning of litter carbon and nitrogen. *Soil Biology and*
943 *Biochemistry*, 40, 1823-1835.
944 Hibbing, M. E., Fuqua, C., Parsek, M. R. & Peterson, S. B. (2010) Bacterial competition: surviving and thriving in
945 the microbial jungle. *Nature Reviews Microbiology*, 8, 15-25.
946 Holland, E. A., Post, W. M., Matthews, E., Sulzman, J., Staufer, R. & Krankina, O. (2015). A Global Database of

947 Litterfall Mass and Litter Pool Carbon and Nutrients. Data set. Available on-line [http://daac.ornl.gov]
948 from Oak Ridge National Laboratory Distributed Active Archive Center. In, Oak Ridge, Tennessee,
949 USA.
950 Huang, Y., Guenet, B., Ciais, P., Janssens, I.A., Soong, J. L., Wang, Y., Goll, D., Blagodatskaya, E. & Huang, Y.
951 (2018). ORCHIMIC (v1.0), A microbe-driven model for soil organic matter decomposition designed for
952 large-scale applications. *Geoscientific Model Development*, 11, 2111-2138.
953 Iversen, C. M., McCormack, M. L., Powell, A. S., Blackwood, C. B., Freschet, G. T., Kattge, J., Roumet, C.,
954 Stover, D. B., Soudzilovskaia, N. A., Valverde-Barrantes, O. J., van Bodegom, P. M. & Violle, C. (2017).
955 A global Fine-Root Ecology Database to address below-ground challenges in plant ecology. *New*
956 *Phytologist*, 215, 15-26.
957 Jenkinson, D. & Rayner, J. (1977). The turnover of soil organic matter in some of the Rothamsted classical
958 experiments. *Soil Science*, 123, 298-305.
959 Jia, B., Zhou, G. & Xu, Z. (2016). Forest litterfall and its composition: a new data set of observational data from
960 China. *Ecology*, 97, 1365.
961 Jobbágy, E. G. & Jackson, R. B. (2000). The Vertical Distribution of Soil Organic Carbon and Its Relation to
962 Climate and Vegetation. *Ecological Applications*, 10, 423-436.
963 Jung, M., Reichstein, M., Ciais, P., Seneviratne, S.I., Sheffield, J., Goulden, M. L., Bonan, G., Cescatti, A., Chen,
964 J., de Jeu, R., Dolman, A. J., Eugster, W., Gerten, D., Gianelle, D., Gobron, N., Heinke, J., Kimball, J.,
965 Law, B. E., Montagnani, L., Mu, Q., Mueller, B., Oleson, K., Papale, D., Richardson, A. D., Roupsard,
966 O., Running, S., Tomelleri, E., Viovy, N., Weber, U., Williams, C., Wood, E., Zaehle, S. & Zhang, K.
967 (2010). Recent decline in the global land evapotranspiration trend due to limited moisture supply. *Nature*,
968 467, 951-954.
969 Kaiser, C., Franklin, O., Dieckmann, U. & Richter, A. (2014). Microbial community dynamics alleviate
970 stoichiometric constraints during litter decay. *Ecology Letters*, 17, 680-690.
971 Kaiser, C., Franklin, O., Richter, A. & Dieckmann, U. (2015). Social dynamics within decomposer communities
972 lead to nitrogen retention and organic matter build-up in soils. *Nat Commun*, 6, 8960.
973 Kallenbach, C., Frey, S. & Grandy, S. (2016). Direct evidence for microbial-derived soil organic matter formation
974 and its ecophysiological controls. *Nature Communications*, 7, 13630.
975 Kattge, J., *et al*. (2011). TRY - a global database of plant traits. *Global Change Biology*, 17, 2905-2935.
976 Kothawala, D. N., Moore, T. R. & Hendershot, W. H. (2008). Adsorption of dissolved organic carbon to mineral

977 soils: A comparison of four isotherm approaches. *Geoderma*, 148, 43-50.
978 Koven, C. D., Lawrence, D. M., & Riley, W. J. (2015). Permafrost carbon−climate feedback is sensitive to deep
979 soil carbon decomposability but not deep soil nitrogen dynamics. *Proceedings of the National Academy*
980 *of Sciences*, 112, 3752-3757.
981 Koven, C. D., Riley, W. J., Subin, Z. M., Tang, J. Y., Torn, M. S. Collins, W. D., Bonan, G. B., Lawrence, D. M.
982 & Swenson, S. C. (2013). The effect of vertically resolved soil biogeochemistry and alternate soil C and
983 N models on C dynamics of CLM4. *Biogeosciences*, 10, 7109-7131.
984 Krinner, G., Viovy, N., de Noblet-Ducoudré, N., Ogée, J., Polcher, J., Friedlingstein, P., Ciais, P., Sitch, S. &
985 Prentice, I.C. (2005). A dynamic global vegetation model for studies of the coupled
986 atmosphere-biosphere system. *Global Biogeochemical Cycles*, 19
987 Kuzyakov, Y. (2010). Priming effects: Interactions between living and dead organic matter. *Soil Biology and*
988 *Biochemistry*, 42, 1363-1371.
989 Lützow, M. v., Kögel-Knabner, I., Ekschmitt, K., Flessa, H., Guggenberger, G., Matzner, E. & Marschner, B.
990 (2007). SOM fractionation methods: Relevance to functional pools and to stabilization mechanisms. *Soil*
991 *Biology and Biochemistry*, 39, 2183-2207.
992 Lützow, M. v., Kogel-Knabner, I., Ekschmitt, K., Matzner, E., Guggenberger, G., Marschner, B. & Flessa, H.
993 (2006). Stabilization of organic matter in temperate soils: mechanisms and their relevance under
994 different soil conditions - a review. *European Journal of Soil Science*, 57, 426-445.
995 Lal, R. (2016). Beyond COP 21: Potential and challenges of the "4 per Thousand" initiative. *Journal of Soil and*
996 *Water Conservation*, 71, 20A-25A.
997 Lehmann, J. & Kleber, M. (2015). The contentious nature of soil organic matter. *Nature*, 528, 60-68.
998 Liang, S., Zhao, X., Liu, S., Yuan, W., Cheng, X., Xiao, Z., Zhang, X., Liu, Q., Cheng, J., Tang, H., Qu, Y., Bo, Y.,
999 Qu, Y., Ren, H., Yu, K. & Townshend, J. (2013). A long-term Global LAnd Surface Satellite (GLASS)
1000 data-set for environmental studies. *International Journal of Digital Earth*, 6, 5-33.
1001 Manzoni, S. & Porporato, A. (2009). Soil carbon and nitrogen mineralization: Theory and models across scales.
1002 *Soil Biology and Biochemistry*, 41, 1355-1379.
1003 Manzoni, S., Moyano, F., Kätterer, T. & Schimel, J. (2016). Modeling coupled enzymatic and solute transport
1004 controls on decomposition in drying soils. *Soil Biology and Biochemistry*, 95, 275-287.
1005 Manzoni, S., Taylor, P., Richter, A., Porporato, A. & Agren, G. I. (2012). Environmental and stoichiometric
1006 controls on microbial carbon-use efficiency in soils. *New Phytologist*, 196, 79-91.

1007 Manzoni, S., Capek, P., Mooshammer, M., Lindahl, B. D., Richter, A. & Santruckova, H. (2017). Optimal
1008 metabolic regulation along resource stoichiometry gradients. *Ecol Lett*, 20, 1182-1191.
1009 McGuire, A. D., Lawrence, D. M., Koven, C., Clein, J. S., Burke, E., Chen, G. & Zhuang, Q. (2018). Dependence
1010 of the evolution of carbon dynamics in the northern permafrost region on the trajectory of climate
1011 change. *Proceedings of the National Academy of Sciences*. 115, 3882-3887.
1012 Muttil, N. & Jayawardena, A. W. (2008). Shuffled Complex Evolution model calibrating algorithm: enhancing its
1013 robustness and efficiency. *Hydrological Processes*, 22, 4628-4638.
1014 Parton, W., Silver, W. L., Burke, I. C., Grassens, L., Harmon, M. E., Currie, W. S., King, J. Y., Adair, E. C.,
1015 Brandt, L.A., Hart, S. C. & Fasth, B. (2007). Global-scale similarities in nitrogen release patterns during
1016 long-term decomposition. *Science*, 315, 361-364.
1017 Parton, W. J., Schimel, D. S., Cole, C. V. & Ojima, D. S. (1987). Analysis of Factors Controlling Soil Organic
1018 Matter Levels in Great Plains Grasslands1. *Soil Science Society of America Journal*, 51, 1173-1179.
1019 Rasmussen, C., Heckman, K., Wieder, W. R., Keiluweit, M., Lawrence, C. R., Berhe, A. A., Blankinship, J. C.,
1020 Crow, S. E., Druhan, J. L., Hicks Pries, C. E., Marin-Spiotta, E., Plante, A.F., Schädel, C., Schimel, J. P.,
1021 Sierra, C. A., Thompson, A. & Wagai, R. (2018). Beyond clay: towards an improved set of variables for
1022 predicting soil organic matter content. *Biogeochemistry*, 137, 297-306.
1023 Robertson, A. D., Paustian, K., Ogle, S., Wallenstein, M. D., Lugato, E. & Cotrufo, M. F. (2019). Unifying soil
1024 organic matter formation and persistence frameworks: the MEMS model. *Biogeosciences*, 16,
1025 1225-1248.
1026 Schimel, D. S., Braswell, B. H., Holland, E. A., McKeown, R., Ojima, D. S., Painter, T. H., Parton, W. J. &
1027 Townsend, A. R. (1994). Climatic, edaphic, and biotic controls over storage and turnover of carbon in
1028 soils. *Global Biogeochemical Cycles*, 8, 279-293.
1029 Schimel, J. P. & Weintraub, M. N. (2003). The implications of exoenzyme activity on microbial carbon and
1030 nitrogen limitation in soil: a theoretical model. *Soil Biology and Biochemistry*, 35, 549-563.
1031 Schmidt, M. W., Torn, M. S., Abiven, S., Dittmar, T., Guggenberger, G., Janssens, I. A., Kleber, M.,
1032 Kogel-Knabner, I., Lehmann, J., Manning, D. A., Nannipieri, P., Rasse, D. P., Weiner, S. & Trumbore,
1033 S.E. (2011) Persistence of soil organic matter as an ecosystem property. *Nature*, 478, 49-56.
1034 Shangguan, W., Dai, Y., Duan, Q., Liu, B. & Yuan, H. (2014). A global soil data set for earth system modeling.
1035 *Journal of Advances in Modeling Earth Systems*, 6, 249-263.
1036 Shi, Z., Crowell, S., Luo, Y. & Moore, B., III. (2018). Model structures amplify uncertainty in predicted soil

carbon responses to climate change. *Nature Communications*, 9, 2171.

Sierra, C. A., Trumbore, S. E., Davidson, E. A., Vicca, S. & Janssens, I. (2015). Sensitivity of decomposition rates of soil organic matter with respect to simultaneous changes in temperature and moisture. *Journal of Advances in Modeling Earth Systems*, 7, 335-356.

Six, J., Bossuyt, H., Degryze, S. & Denef, K. (2004). A history of research on the link between (micro)aggregates, soil biota, and soil organic matter dynamics. *Soil and Tillage Research*, 79, 7-31.

Six, J., Feller, C., Denef, K., Ogle, S.M., de Moraes, J.C. & Albrecht, A. (2002). Soil organic matter, biota and aggregation in temperate and tropical soils - Effects of no-tillage. *Agronomie*, 22, 755-775.

Six, J. & Paustian, K. (2014) Aggregate-associated soil organic matter as an ecosystem property and a measurement tool. *Soil Biology & Biochemistry*, 68, A4-A9.

Sitch, S., Smith, B., Prentice, I. C., Arneth, A., Bondeau, A., Cramer, W., Kaplan, J. O., Levis, S., Lucht, W., Sykes, M. T., Thonicke, K., & Venevsky, S. (2003) Evaluation of ecosystem dynamics, plant geography and terrestrial carbon cycling in the LPJ dynamic global vegetation model, *Global Change Biology*, 9, 161–185.

Sokol, N. W., Sanderman, J. & Bradford, M. A. (2019). Pathways of mineral-associated soil organic matter formation: Integrating the role of plant carbon source, chemistry, and point of entry. *Global Change Biology*, 25, 12-24.

Stewart, C. E., Paustian, K., Conant, R. T., Plante, A. F. & Six, J. (2007). Soil carbon saturation: concept, evidence and evaluation. *Biogeochemistry*, 86, 19-31.

Stewart, C. E., Plante, A. F., Paustian, K., Conant, R. T. & Six, J. (2008). Soil Carbon Saturation: Linking Concept and Measurable Carbon Pools. *Soil Science Society of America Journal*, 72, 379.

Stockmann, U., Adams, M. A., Crawford, J. W., Field, D. J., Henakaarchchi, N., Jenkins, M., Minasny, B., McBratney, A. B., Courcelles, V. d. R. d., Singh, K., Wheeler, I., Abbott, L., Angers, D. A., Baldock, J., Bird, M., Brookes, P. C., Chenu, C., Jastrow, J. D., Lal, R., Lehmann, J., O'Donnell, A. G., Parton, W. J., Whitehead, D. & Zimmermann, M. (2013). The knowns, known unknowns and unknowns of sequestration of soil organic carbon. *Agriculture, Ecosystems & Environment*, 164, 80-99.

Tang, J. & Zhuang, Q. (2013). A global sensitivity analysis and Bayesian inference framework for improving the parameter estimation and prediction of a process-based Terrestrial Ecosystem Mode. *Journal of Geophysical Research:Atmospheres*, D15303. doi:10.1029/2009JD011724

Tang, X., Zhao, X., Bai, Y., Tang, Z., Wang, W., Zhao, Y., Wan, H., Xie, Z., Shi, X., Wu, B., Wang, G., Yan, J.,

Ma, K., Du, S., Li, S., Han, S., Ma, Y., Hu, H., He, N., Yang, Y., Han, W., He, H., Yu, G., Fang, J. & Zhou, G. (2018). Carbon pools in China's terrestrial ecosystems: New estimates based on an intensive field survey. *Proc Natl Acad Sci U S A*, 115, 4021-4026.

Tarnocai, C., Canadell, J. G., Schuur, E. A. G., Kuhry, P., Mazhitova, G. & Zimov, S. (2009). Soil organic carbon pools in the northern circumpolar permafrost region. *Global Biogeochemical Cycles*, 23, GB2023, doi:10.1029/2008GB003327

Thornton, P. E. & Rosenbloom, N. A. (2005). Ecosystem model spin-up: Estimating steady state conditions in a coupled terrestrial carbon and nitrogen cycle model. *Ecological Modelling*, 189, 25-48.

Tifafi, M., Guenet, B. & Hatté C. (2018). Large Differences in Global and Regional Total Soil Carbon Stock Estimates Based on SoilGrids, HWSD, and NCSCD: Intercomparison and Evaluation Based on Field Data From USA, England, Wales, and France. *Global Biogeochemical Cycles*, 32, 42-56.

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

Tucker, C. J., Pinzon, J. E., Brown, M. E., Slayback, D. A., Pak, E. W., Mahoney, R., Vermote, E. F. & El Saleous, N. (2005). An extended AVHRR 8-km NDVI dataset compatible with MODIS and SPOT vegetation NDVI data. *International Journal of Remote Sensing*, 26, 4485-4498.

Ukonmaanaho, L., Pitman, R., Bastrup-Birk, A., Breda, N. & Rautio, P. (2016). Part XIII: Sampling and Analysis of Litterfall. In: UNECE ICP Forests Programme Co-ordinating Centre (ed.): Manual on methods and criteria for harmonized sampling, assessment, monitoring and analysis of the effects of air pollution on forests. In, Thünen Institute for Forests Ecosystems, Eberswalde, Germany.

Viovy, N. (2018). CRUNCEP Version 7 - Atmospheric Forcing Data for the Community Land Model. Research Data Archive at the National Center for Atmospheric Research,. In, Computational and Information Systems Laboratory. http://rda.ucar.edu/datasets/ds314.3/.

Viscarra Rossel, R. A. & Hicks, W. S. (2015). Soil organic carbon and its fractions estimated by visible-near infrared transfer functions. *European Journal of Soil Science*, 66, 438-450.

Viscarra Rossel, R. A., Lee, J., Berhrens, T., Luo, Z., Baldock, J. & Richards, A. (2019). Continental-scale soil carbon composition and vulnerability modulated by regional environmental controls. *Nature Geoscience*, 12, 547-552.

Wagner, S., Cattle, S. R. & Scholten, T. (2007). Soil-aggregate formation as influenced by clay content and

organic-matter amendment. *Journal of Plant Nutrition and Soil Science*, 170, 173-180.

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, 255-272.

Wang, Q., He, T. & Liu, J. (2016). Litter input decreased the response of soil organic matter decomposition to warming in two subtropical forest soils. *Scitific Reports*, 6, 33814.

Wang, Y. P., Chen, B.C., Wieder, W. R., Leite, M., Medlyn, B. E., Rasmussen, M., Smith, M. J., Agusto, F. B., Hoffman, F. & Luo, Y. Q. (2014). Oscillatory behavior of two nonlinear microbial models of soil carbon decomposition. *Biogeosciences*, 11, 1817-1831.

Wieder, W. R., Bonan, G. B. & Allison, S.D. (2013). Global soil carbon projections are improved by modelling microbial processes. *Nature Climate Change*, 3, 909-912.

Wieder, W. R., Boehnert, J. & Bonan, G. B. (2014a). Evaluating soil biogeochemistry parameterizations in Earth system models with observations. *Global Biogeochemical Cycles*, 28, 211-222.

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

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, 1789-1808.

Wieder, W. R., Hartman, M. D., Sulman, B. N., Wang, Y-P, Koven, C. D. & Bonan, G. B. (2018). Carbon cycle confidence and uncertainty: Exploring variation among soil biogeochemical models. *Global Change Biology*., 24, 1563–1579.

Wu, D., Piao, S., Liu, Y., Ciais, P. & Yao, Y. (2018). Evaluation of CMIP5 Earth System Models for the Spatial Patterns of Biomass and Soil Carbon Turnover Times and Their Linkage with Climate. *Journal of Climate*, 31, 5947-5960.

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

Xu, X., Thornton, P. E. & Post, W. M. (2013). A global analysis of soil microbial biomass carbon, nitrogen and phosphorus in terrestrial ecosystems. *Global Ecology and Biogeography*, 22, 737-749.

Zhang, H., Yuan, W., Dong, W. & Liu, S. (2014). Seasonal patterns of litterfall in forest ecosystem worldwide.

Ecological Complexity, 20, 240-247.

Zhang, H., Goll, D. S., Manzoni, S., Ciais, P., Guenet, B. & Huang, Y. (2018). Modeling the effects of litter stoichiometry and soil mineral N availability on soil organic matter formation using CENTURY-CUE (v1.0). *Geoscientific Model Development*, 11, 4779-4796.

Table

Table 1 Tested models in this study and parameters subject to optimization of each model. k_{litt} and k_{soc} tune the turnover rate of litter and SOC pools in CENTURY, respectively. a_v and a_k are parameters tune microbial maximum reaction velocity (Eq. 4) and half-saturation constant (Eq. 5). k_d, k_{dp} and k_{bs} tune the deprotection rate of SOC_p (Eqs. 13-15). β tunes the density-dependent microbial turnover rate (Eq. 16).

Model	Optimized parameters
CENTURY	k_{litt}, k_{soc}
MIMICS-def	a_v, a_k, k_d
MIMICS-D	a_v, a_k, k_d, k_{dp}
MIMICS-DB	a_v, a_k, k_d, k_{dp}, k_{bs}
MIMICS-DBT	a_v, a_k, k_d, k_{dp}, k_{bs}, β

Figure

Figure 1 Soil C pools and fluxes represented in CENTURY (a) and MIMICS (b). In both models, litter inputs (Lit_{inp}) are partitioned into metabolic and structural litter pools (LIT_m and LIT_s) based on litter quality (f_{met}). The soil organic carbon (SOC) in CENTURY are divided into active (SOC_{act}), slow (SOC_{slow}) and passive (SOC_{pas}) pools. CUE is the carbon use efficiency of decomposed litter or SOC. In MIMICS, decomposition of litter and available SOM pools (SOC_a) are governed by temperature sensitive Michaelis–Menten kinetics (V_{max} and K_m). Microbial growth efficiency (MGE) determines the partitioning of C fluxes entering microbial biomass pools vs.

heterotrophic respiration. Turnover of the microbial biomass (τ) depends on microbial functional type (MIC_r and MIC_k), and is partitioned into available, physically and physicochemically protected, and chemically recalcitrant SOC pools (SOC_a, SOC_p, and SOC_c, respectively). $f_{i,met}$ and $f_{i,stru}$ denote the fraction of decomposed metabolic litter to SOC_p and the fraction of decomposed structural litter to SOC_c, respectively. f_p and f_c denote the fraction of τ partitioned to SOC_p and the fraction of τ partitioned to SOC_c, respectively.

Figure 2 Comparison of CENTURY (a) and MIMICS (b-e) for simulating large-scale variation of SOC concentrations across the 206 forest sites in Europe and China. RMSE is the root mean square error, and AIC is the Akaike information criterion. MIMICS versions include the default model (MIMICS-def), revised SOC deprotection rate (MIMICS-D), using base saturation to modify deprotection rates (MIMICS-DB); and density-dependent microbial turnover rate (MIMICS-DBT; see section 2.2.2).

Figure 3 Partial correlation coefficients between the biases of simulated SOC concentrations and the climate condition, amount and quality of litter input, and soil physical and chemical properties. MAT: mean annual temperature (°C), MAP: mean annual total precipitation (mm), MAP-PET: the difference between annual total precipitation and potential evapotranspiration (mm), ET: evapotranspiration (mm), LAI_{max}: mean of the annual maximum leaf area index at the observation site during the period from 1982 to 2000, LAI_{trend}: change trend of the LAI_{max} during the period from 1982 to 2000 (yr^{-1}), $NDVI_{max}$: mean of the annual maximum normalized difference vegetation index at the observation site during the period from 1982 to 2000, LAI_{trend}: change trend of the $NDVI_{max}$ during the period from 1982 to 2000 (yr^{-1}), $litter_{ab}$: aboveground litter-C stock (g C m^{-2}), SWC: soil water content, BD: bulk density (g cm^{-3}), BS: base saturation (0-1, dimensionless), CEC: Cation of exchange capacity (cmol kg^{-1}). Partial correlation coefficients between -0.14 and 0.14 were not significant ($p > 0.05$).

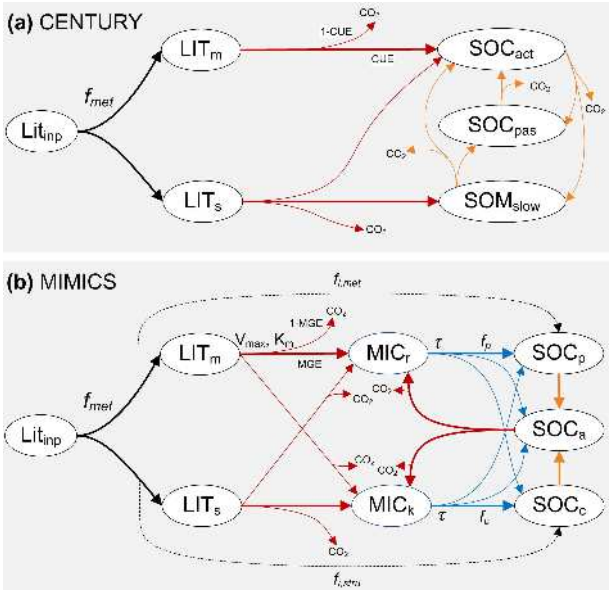
Figure 4 Sensitivity of simulated SOC concentration to mean annual temperature (S_{MAT}, a), soil water content (S_{SWC}, b), soil clay fraction (S_{clay}, c), annual litterfall input ($S_{litterfall}$, d), the C:N ratio of litterfall ($S_{C:N}$, e) and the lignin:C ratio of litterfall ($S_{lignin:C}$, f). The blue and red dashed lines denote insignificant and significant (p<0.05) sensitivity calculated based on observation data, respectively. The solid line in each box denotes the median value. Box boundaries show the 25th and 75th percentiles, whiskers denote the 10th and 90th percentiles, and the black dots denote the 5th and 95th percentiles.

Figure 5 Comparison between the simulated ratio of microbial C (MIC) to total SOC from different versions of MIMICS and the observed values at globally-distributed forest sites. The dashed and solid lines in each box are the mean and median value, respectively. Box boundaries show the 25th and 75th percentiles, whiskers denote the 10th and 90th percentiles, the dots below and above each box denote the 5th and 95th percentiles, respectively. The 655 samples of observed MIC/SOC at globally-distributed forest sites are collected by Xu et al., 2013.

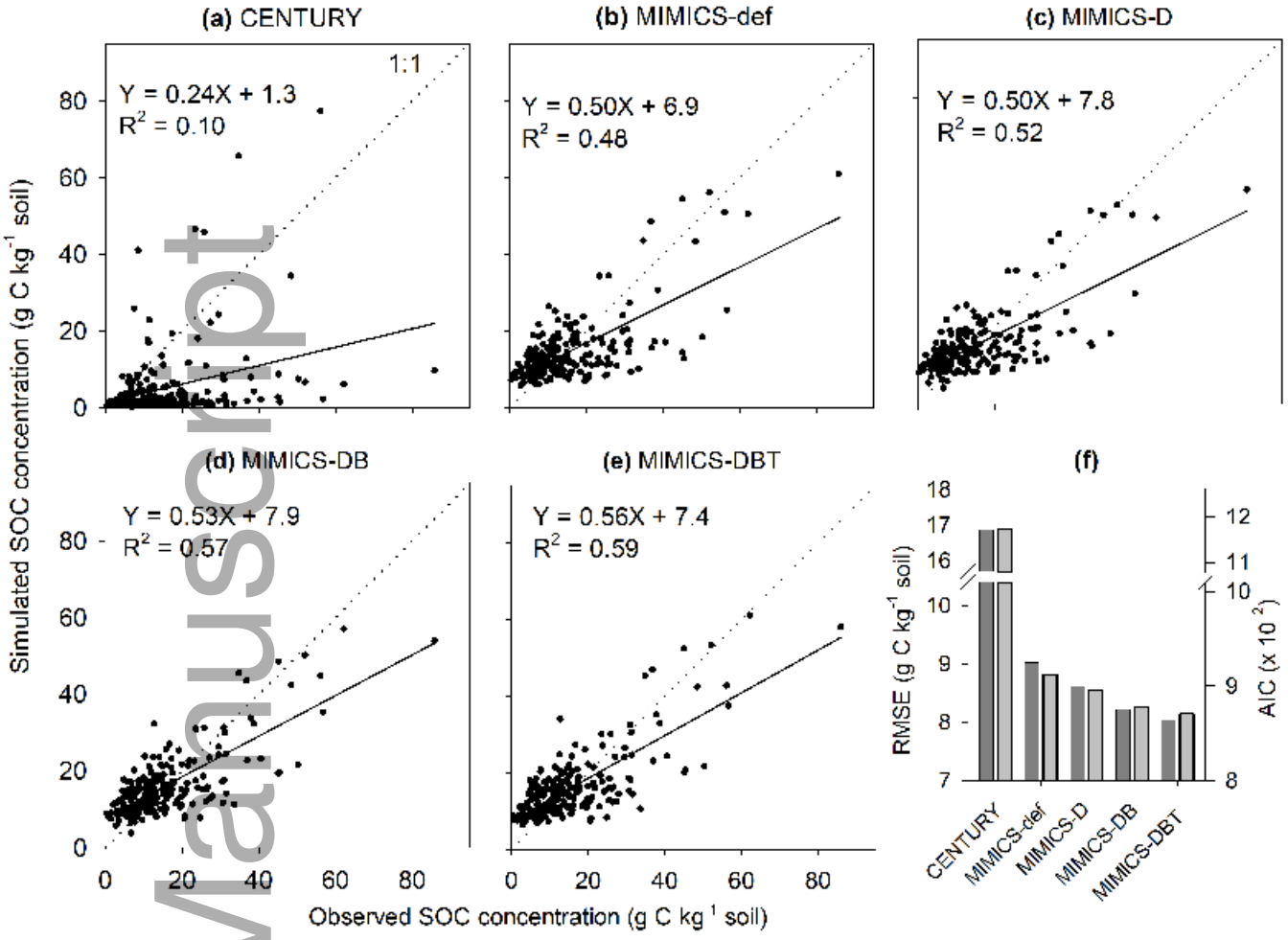
Figure 6 Comparison between the simulated SOC compositions from optimized MIMICS (a) and CENTURY (b) model and the observed SOC compositions at 505 sites in Australia (c). The observation data in Australia are obtained from Viscarra-Rossel et al. (2019). Viscarra-Rossel et al. partitioned total SOC into three fractions with different particle-sizes: the particulate organic carbon (POC), the humic organic carbon (HOC) and the resistant organic carbon (ROC, which is the mineral-associated organic carbon). The line in each box denotes median value. Box boundaries show the 25th and 75th percentiles, whiskers denote the 10th and 90th percentiles, and the dots below and above each box denote the 5th and 95th percentiles.

Figure 7 Partial correlation coefficients between fraction of each SOC pool and model drivers, including mean annual temperature (MAT, °C), soil water content

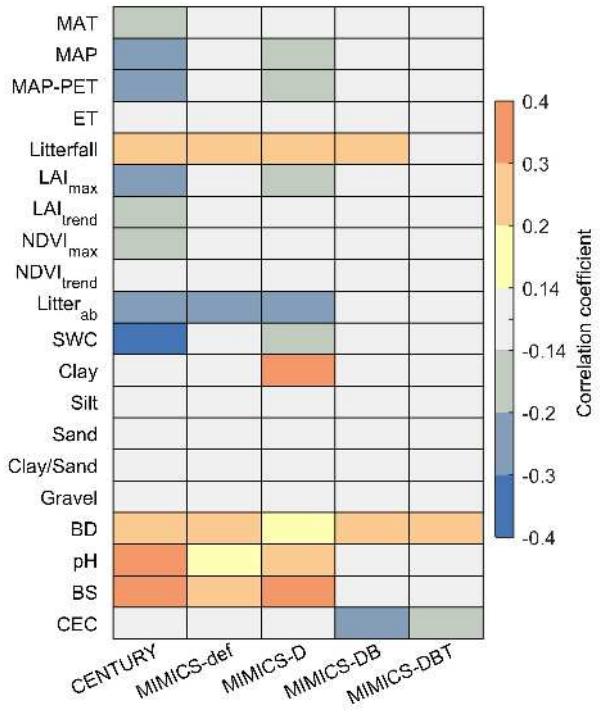
(SWC, dimensionless), soil clay content (clay, dimensionless), annual total litterfall production (Litterfall, g C m^{-2} yr^{-1}), litter C:N ratio (C:N), litter lignin:C ratio (Lignin:C), base saturation (BS, 0-1, dimensionless) and total SOC concentration (SOC), Figure (a) Obs show the results based on observation data from Australia. Figure (b)-(f) showed the results based on optimized CENTURY and MIMICS models. Partial correlation coefficients between -0.14 and 0.14 were not significant ($p > 0.05$).



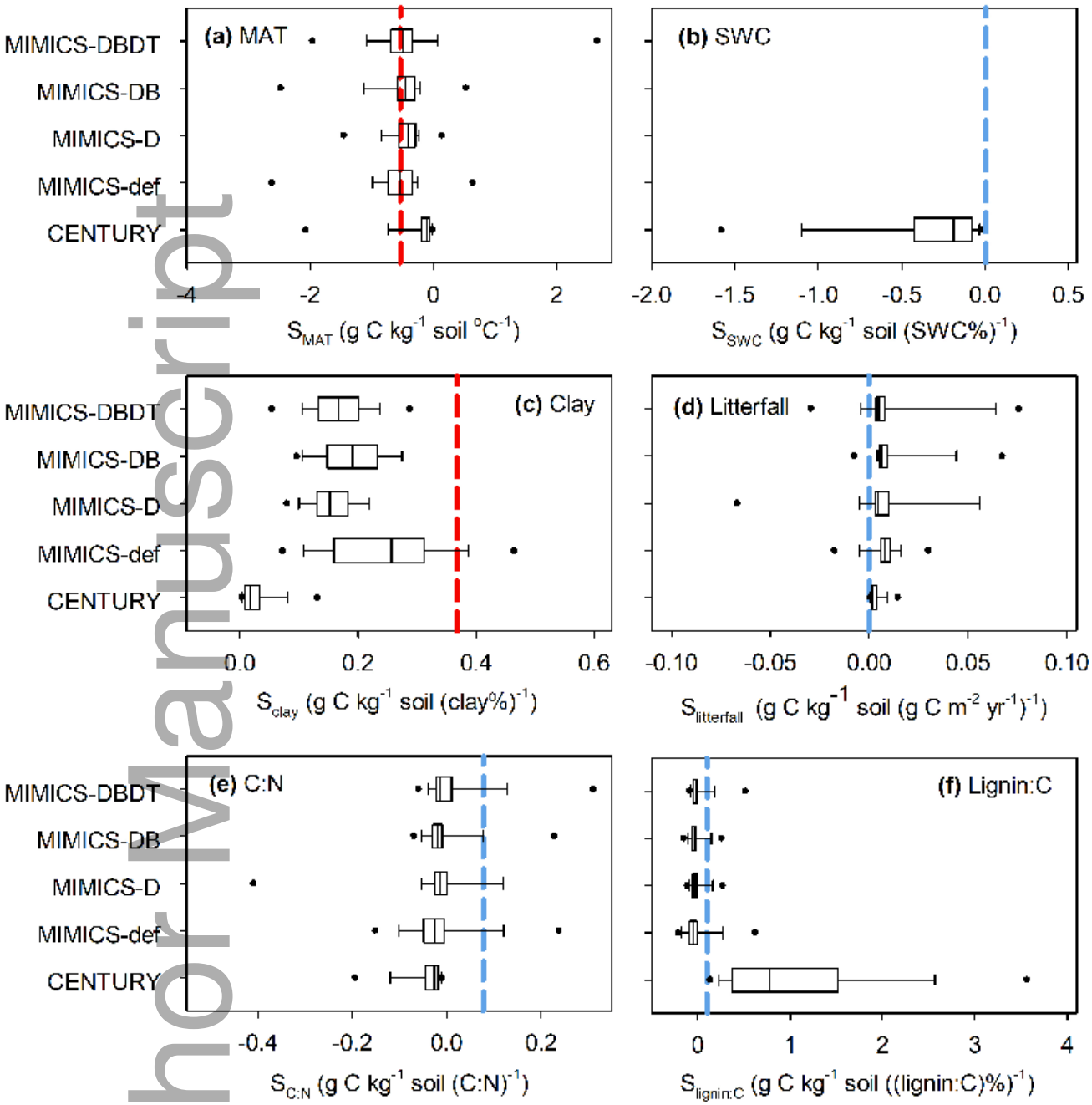
gcb_14994_f1.jpg



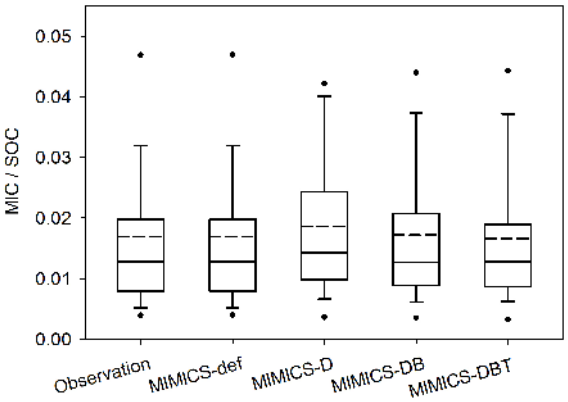
gcb_14994_f2.jpg



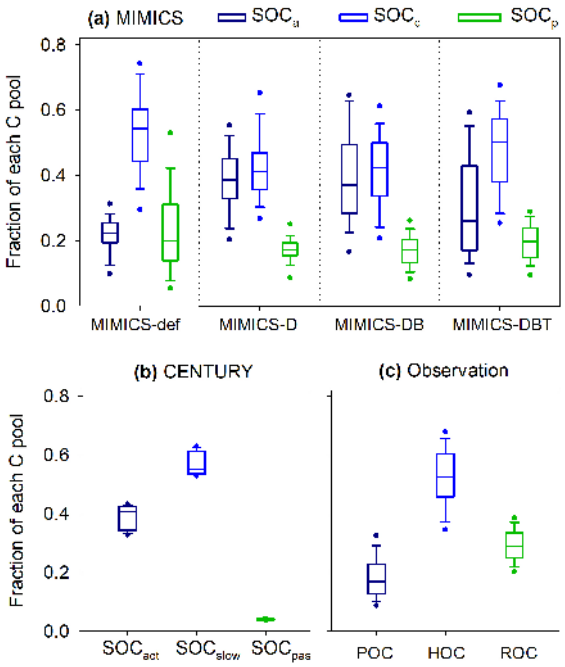
gcb_14994_f3.jpg



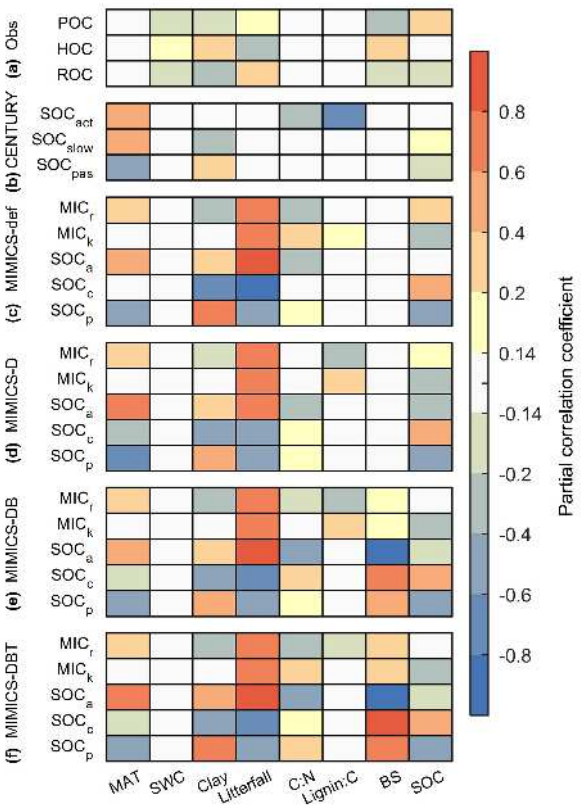
gcb_14994_f4.jpg



gcb_14994_f5.jpg



gcb_14994_f6.jpg



gcb_14994_f7.jpg