Global Change Biology

Global Change Biology (2015), doi: 10.1111/gcb.12827

Microbial models with data-driven parameters predict stronger soil carbon responses to climate change

OLEKSANDRA HARARUK[1,2], MATTHEW J. SMITH[2] and YIQI LUO[1,3]

[1]*Department of Microbiology and Plant Biology, University of Oklahoma, Norman, OK, USA,* [2]*Computational Science Laboratory, Microsoft Research, Cambridge, UK,* [3]*Center for Earth System Science, Tsinghua University, Beijing, China*

Abstract

Long-term carbon (C) cycle feedbacks to climate depend on the future dynamics of soil organic carbon (SOC). Current models show low predictive accuracy at simulating contemporary SOC pools, which can be improved through parameter estimation. However, major uncertainty remains in global soil responses to climate change, particularly uncertainty in how the activity of soil microbial communities will respond. To date, the role of microbes in SOC dynamics has been implicitly described by decay rate constants in most conventional global carbon cycle models. Explicitly including microbial biomass dynamics into C cycle model formulations has shown potential to improve model predictive performance when assessed against global SOC databases. This study aimed to data-constrained parameters of two soil microbial models, evaluate the improvements in performance of those calibrated models in predicting contemporary carbon stocks, and compare the SOC responses to climate change and their uncertainties between microbial and conventional models. Microbial models with calibrated parameters explained 51% of variability in the observed total SOC, whereas a calibrated conventional model explained 41%. The microbial models, when forced with climate and soil carbon input predictions from the 5th Coupled Model Intercomparison Project (CMIP5), produced stronger soil C responses to 95 years of climate change than any of the 11 CMIP5 models. The calibrated microbial models predicted between 8% (2-pool model) and 11% (4-pool model) soil C losses compared with CMIP5 model projections which ranged from a 7% loss to a 22.6% gain. Lastly, we observed unrealistic oscillatory SOC dynamics in the 2-pool microbial model. The 4-pool model also produced oscillations, but they were less prominent and could be avoided, depending on the parameter values.

Keywords: carbon cycle, carbon-climate feedback, data assimilation, model calibration, soil biogeochemistry, soil organic matter

Received 24 March 2014 and accepted 18 October 2014

Introduction

Soils contain the largest fraction of global terrestrial carbon (C), storing more C than the vegetation and atmosphere combined (Falkowski *et al.*, 2000; House *et al.*, 2002). Changing climate could accelerate soil organic carbon (SOC) decomposition (Fang *et al.*, 2005), increasing atmospheric CO_2 concentrations, which could cause further climate warming (Falkowski *et al.*, 2000). Such interdependencies between SOC and climate highlight the importance of accurate prediction of global SOC distributions and their feedbacks to climate change.

Predictions of the contemporary global SOC content from 11 Earth system models (ESMs) participating in the 5th Coupled Model Intercomparison Project (CMIP5; Taylor *et al.*, 2011) vary sixfold, ranging from 510 to 3040 Pg C (Todd-Brown *et al.*, 2013b). Of 11 models, only six produced SOC pools that were within the range of the Harmonized World Soil Database estimates (HWSD, 890–1660 Pg C) and none explained more than 16% of the spatial variability in the HWSD soil C (Todd-Brown *et al.*, 2013b). The differences in model representations of contemporary soil carbon result in a wide range of soil responses to climate change. For example, under the high radiative forcing scenario (RCP 8.5), model projections of changes in global SOC by 2100 ranged between a 72 Pg C loss and a 253 Pg C gain (Todd-Brown *et al.*, 2013a). This emphasizes the need for improving the models used for soil carbon cycle simulation, or at least understanding whether and how improvements in predictive performance may be achieved.

All CMIP5 models simulate soil carbon decomposition as a first-order decay process (Todd-Brown *et al.*, 2013b). Such formulations (which we call 'conventional') represent the decomposing activity of microbes as decay constants, modified by environmental functions, and assume that the amount of decomposed SOC is linearly dependent on the SOC stocks. These conventional models cannot account for some microbial processes observed in experimental studies, such as the priming effect (Kuzyakov *et al.*, 2000; Fontaine *et al.*, 2004, 2007), microbial acclimation to increasing temperatures (Luo *et al.*, 2001; Chen & Tian, 2005; Peng *et al.*, 2009), and CO_2-induced changes in the microbial com-

Correspondence: Oleksandra Hararuk, tel. +1 405 589 1067, fax: +1 405 325 7619; e-mail: ohararuk@ou.edu

munity composition (Carney *et al.*, 2007). Schimel & Weintraub (2003) argue that SOC decomposition should not be represented by decay constants because the decomposition rate is regulated by extracellular enzyme concentrations. Instead, they propose that SOC dynamics should be modeled using functions based on Michaelis–Menten kinetics, which account for the concentrations of enzymes.

In recent years several, enzyme-driven decomposition models (which we call 'microbial models') were developed (Schimel & Weintraub, 2003; Allison *et al.*, 2010; German *et al.*, 2012; Wang *et al.*, 2012). These simulate the acclimation of soil respiration to elevated temperatures (Allison *et al.*, 2010) as well as the priming effect (Schimel & Weintraub, 2003). Moreover, replacing a conventional model with a microbial model in the Community Land Model improved its predictive accuracy for the global SOC distribution (Wieder *et al.*, 2013). However, microbial models may also produce responses not observed in nature: A recent analysis illustrated that microbial models produce unrealistic oscillatory responses to small perturbations (Wang *et al.*, 2013). Such unrealistic dynamical properties emphasize the importance of investigating whether and why more realistic biological assumptions lead to more realistic biological predictions at the spatial and temporal scales they are intended to be applied.

Improvements to model predictions can also be achieved by calibration of the model parameters. With the increase in the number of globally observed datasets, more studies have focused on assimilating data into carbon cycle models to estimate their parameters. For instance, Ise & Moorcroft (2006) and Zhou *et al.* (2009) assimilated global SOC (Global Soil Data Task Group, 2000) data into a C cycle model to constrain SOC temperature sensitivities; Hararuk *et al.* (2014) researched how assimilating global SOC data changed model parameters and SOC feedbacks to changing climate; and Smith *et al.* (2013) assimilated multiple datasets into a global terrestrial carbon cycle model to explore structural as well as parameter uncertainties. All of these studies parameterized conventional carbon cycle models, and to date, limited research has been carried out on calibrating global microbial decomposition models and studying the climate change feedbacks they predict (though see Wieder *et al.*, 2013).

This study was to (1) calibrate two microbial models to the global estimated distributions of total soil organic carbon and microbial biomass carbon; (2) compare the performance of the calibrated microbial and conventional models at predicting the contemporary SOC distribution; (3) test for the presence of unrealistic oscillations in the microbial model dynamics; and (4) compare SOC responses to climate change between microbial and conventional models and evaluate the uncertainties of those responses.

Materials and methods

Models

We performed Bayesian parameter estimation on two microbial models and compared the parameter estimates and model predictions to those of a calibrated conventional model (Hararuk *et al.*, 2014). For the conventional model, we used Community Land Model coupled with Carnegie-Ames-Stanford Approach biogeochemistry submodel (CLM-CASA'; Parton *et al.*, 1993; Oleson *et al.*, 2008, 2004). This model has the soil carbon cycle compartment modeled as a 3-pool system and has C transfers among the pools regulated by temperature, soil moisture, soil clay content, and SOC pool sizes (Fig. 1a). For the 2-pool microbial model formulation (Fig. 1b), we used the model described by German *et al.* (2012), but with altered calculations of half-saturation constants and temperature limitation of C uptake so as to make them comparable to the ones in the 4-pool microbial model (Allison *et al.*, 2010). For the 4-pool microbial model, we used the model introduced by Allison *et al.* (2010).

The two-pool microbial model of German *et al.* (2012) is described as

$$\frac{\mathrm{d}MIC}{\mathrm{d}t} = CUE \times V_{max} \times MIC\frac{SOC}{Km + SOC} - r_d \times MIC \quad (1)$$

$$\frac{\mathrm{d}SOC}{\mathrm{d}t} = Input_{soil} + r_d \times MIC - V_{max} \times MIC\frac{SOC}{Km + SOC} \quad (2)$$

with

$$CUE = CUE_{slope} \times T_S - CUE_0 \quad (3)$$

$$V_{max} = V_{max_0} \times \exp\left(-\frac{E_a}{R \times (T_S + 273)}\right) \times \exp\left(-par_{clay} \times clay\right) \quad (4)$$

$$Km = Km_{slope} \times T_S + Km_0 \times \exp\left(par_{lig} \times lignin\right) \quad (5)$$

where MIC and SOC are the microbial biomass and soil organic carbon pools (Fig. 1b), respectively; $Input_{soil}$ is the carbon transferred to the soil from the litter pool; V_{max} is the temperature adjusted rate of SOC decomposition; Km is the half-saturation constant for substrate-limited soil organic carbon decomposition rate; r_d is the microbial death rate; CUE is the microbial carbon use efficiency; T_S is soil temperature; R is the gas constant (8.31 J K^{-1} mol^{-1}); CUE_0 and CUE_{slope} are the baseline microbial carbon use efficiency and its dependency on temperature, respectively; V_{max_0} is the maximum rate of microbial carbon uptake; E_a is the activation energy of SOC decomposition; and Km_0 and Km_{slope} are the baseline half-saturation constant and its dependency on temperature, respectively.

Microbial respiration, when normalized by microbial biomass $\left((1 - CUE) \times V_{max} \times \frac{SOC}{Km+SOC}\right)$, has been reported to be nonlinearly dependent on soil clay content (Müller & Höper, 2004). We therefore modified the original model of German *et al.* (2012) to include an exponential function of clay limita-

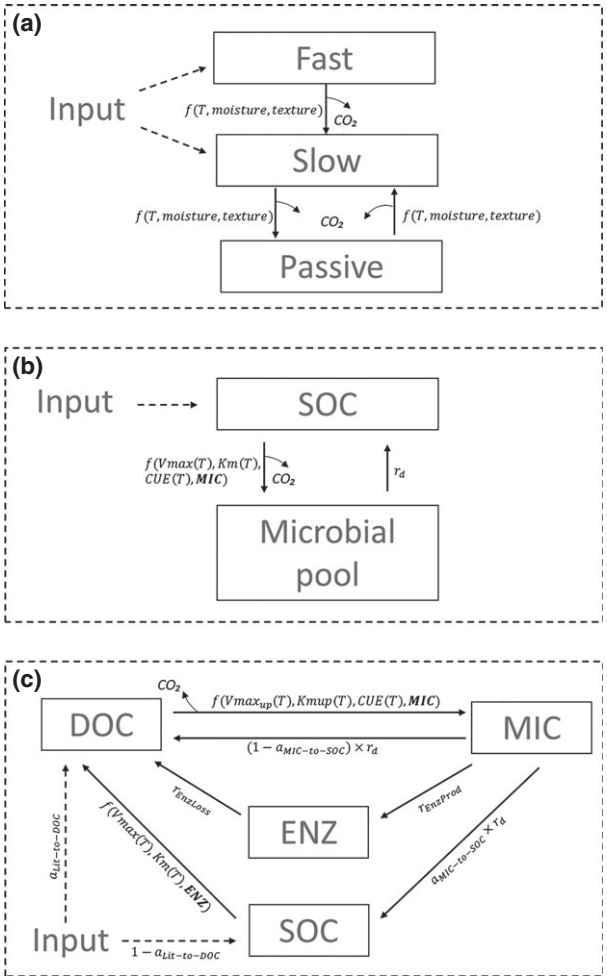

Fig. 1 Schematic representation of the conventional (a), 2-pool (b), and 4-pool (c) microbial models.

tion of decomposition and estimated a parameter par_{clay} to test whether it was different from 0 in the microbial model formulation. We also mimic the substrate quality limitation by adjusting the baseline half-saturation constant by a lignin-dependent correction factor, with its magnitude regulated by par_{lig}, because previous studies have reported substrate quality limitations on decomposition (Taylor *et al*., 1989; Vance & Chapin, 2001; Cusack *et al*., 2009).

The 4-pool microbial model from Allison *et al*. (2010) is

$$\frac{dMIC}{dt} = V_{maxup} \times MIC \frac{DOC}{Kmup + DOC} \times CUE - r_d \times MIC - r_{EnzProd} \times MIC \quad (6)$$

$$\frac{dDOC}{dt} = a_{lit\text{-}to\text{-}DOC} \times Input_{soil} + r_d \times MIC \times (1 - a_{MIC\text{-}to\text{-}SOC}) + V_{max} \times ENZ \frac{SOC}{Km + SOC} + r_{EnzLoss} \times ENZ - V_{maxup} \times MIC \frac{DOC}{Kmup + DOC} \quad (7)$$

$$\frac{dSOC}{dt} = a_{lit\text{-}to\text{-}DOC} \times Input_{soil} + r_d \times MIC \times a_{MIC\text{-}to\text{-}SOC} - V_{max} \times ENZ \frac{SOC}{Km + SOC} \quad (8)$$

$$\frac{dENZ}{dt} = r_{EnzProd} \times MIC - r_{EnzLoss} \times ENZ \quad (9)$$

where ENZ and DOC are enzyme and dissolved organic carbon pools, respectively; V_{maxup} is the temperature adjusted rate of DOC uptake by microbes; $Kmup$ is a half-saturation constant limiting microbial uptake of DOC; $r_{EnzProd}$ is a rate of enzyme production; $Input_{soil}$ is C transferred from litter to soil; $a_{lit\text{-}to\text{-}DOC}$ is the fraction of $Input_{soil}$ that is transferred to DOC; $a_{MIC\text{-}to\text{-}SOC}$ is the fraction of dead microbes transferred to soil; and $r_{EnzLoss}$ is the rate of enzyme loss. The functions in the 2- and 4-pool models were dependent on temperature as follows:

$$CUE = CUE_{slope} \times T_S - CUE_0 \quad (10)$$

$$V_{maxup} = V_{maxup_0} \times \exp\left(-\frac{E_{aup}}{R \times (T_s + 273)}\right) \quad (11)$$

$$Kmup = Kmup_{slope} \times T_S + Kmup_0 \quad (12)$$

$$V_{max} = V_{max_0} \times \exp\left(-\frac{E_a}{R \times (T_S + 273)}\right) \times \exp\left(-par_{clay} \times clay\right) \quad (13)$$

$$Km = Km_{slope} \times T_S + Km_0 \times \exp(par_{lig} \times lignin) \quad (14)$$

where V_{maxup_0} is the maximum rate of microbial DOC uptake; E_{aup} is the activation energy of DOC uptake; $Kmup_0$ and $Kmup_{slope}$ are baseline half-saturation constants for substrate limitation of DOC uptake and its dependency on temperature, respectively.

Data

We used two soil carbon datasets for Bayesian parameter estimation: A global map of total soil carbon content for the top 1 m of soil generated by International Geosphere-Biosphere Programme – Data and Information System (IGBP-DIS; Global Soil Data Task Group, 2000); and a global map of soil microbial biomass distribution for the top 1 m (Xu *et al*., 2013). The IGBP-DIS dataset has been widely used for production of new datasets (House *et al*., 2002), the assessment of terrestrial C uptake (Freibauer *et al*., 2004), for model evaluation (Kucharik *et al*., 2000; Delire *et al*., 2003), and for model improvement (Ise & Moorcroft, 2006; Zhou *et al*., 2009; Smith *et al*., 2013). The global microbial dataset has been used previously for parameterization of a biogeochemical model (Waring *et al*., 2014). Prior to using the datasets in the data assimilation routine, we randomly separated all the grid cells into two groups as in Smith *et al*. (2013) and Hararuk *et al*. (2014), used one group for model parameter estimation, and the other for evaluation to guard against overfitting.

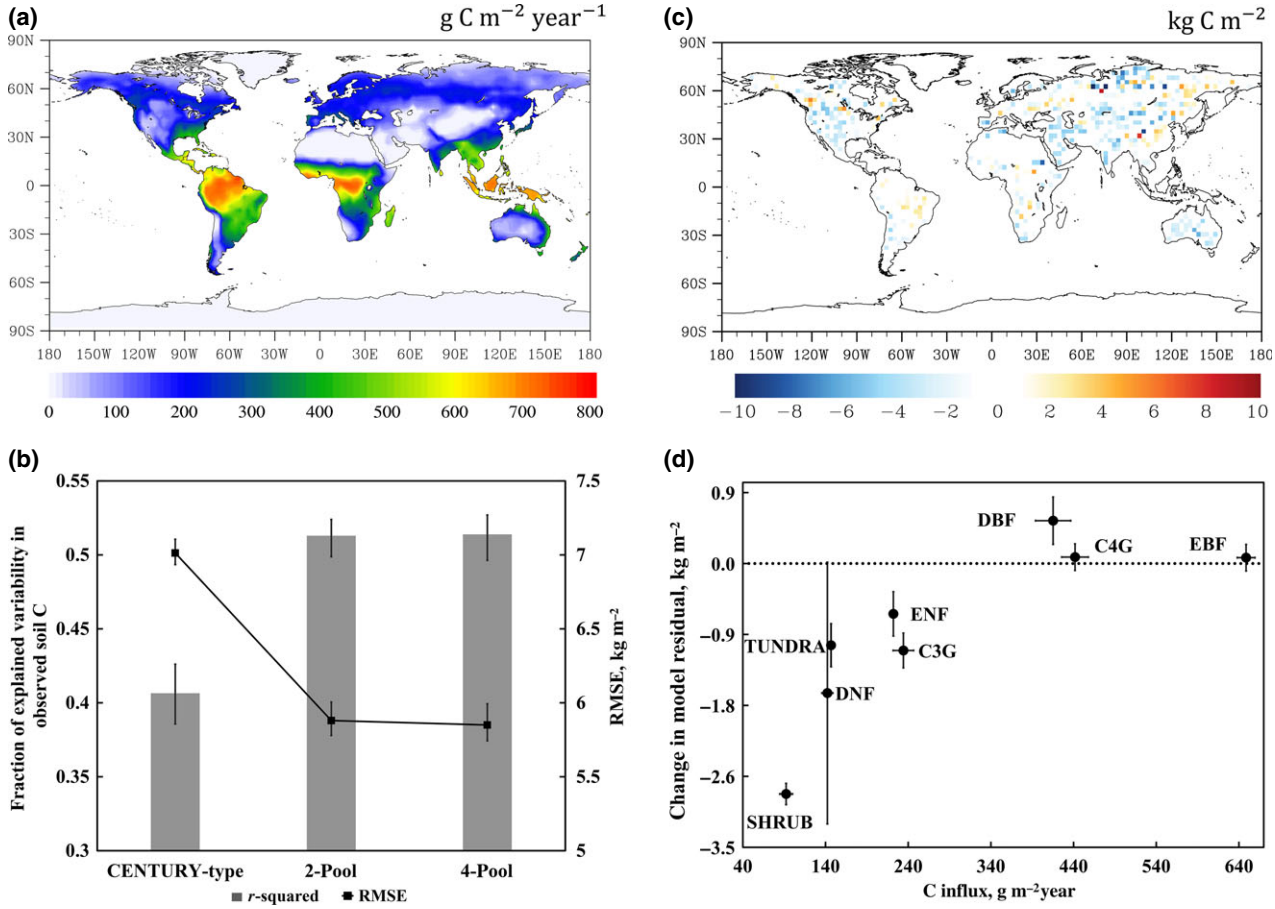

Fig. 2 Annual soil C influx used to drive the soil submodels (a); performance of calibrated microbial and conventional models (b); spatial distribution of changes in the residuals' magnitudes after switching from conventional to microbial model formulation (c), pixels used for calibration are not shown; plant-functional-type-level changes in the magnitudes of model residuals (d), error bars are standard errors. C3G – C3 grasslands; C4G – C4 grasslands; DBF – deciduous broadleaf forest; DNF – deciduous needleleaf forest; EBF – evergreen broadleaf forest; ENF – evergreen needleleaf forest; SHRUB – shrublands.

The global litter lignin content was provided as part of the CLM-CASA' package: The plant-functional-type-level estimates of lignin were applied to the MODIS-derived distributions of plant functional types used in CLM (Lawrence & Chase, 2007; Oleson *et al*., 2007). The map of soil clay content was originally developed by the International Geosphere-Biosphere Programme (Global Soil Data Task Group, 2000) and was also provided as a part of the CLM-CASA' package. We also used 30-year averages of soil temperatures and soil C input produced by CLM-CASA' and calibrated by Hararuk *et al.* (2014) (Fig. 2a) as external model inputs. Soil temperatures were calculated using air temperatures from global reanalysis data (Qian *et al*., 2006), and soil C input was strongly correlated with MODIS NPP data (r^2 = 0.75; Fig. S1). Soil C input was 45% of NPP values, which, given the equilibrium assumption, implied that 55% of the incoming C was returned back to the atmosphere via respiration from the litter pools.

Parameter estimation

We calibrated parameters in the two microbial models using Bayesian probabilistic inversion. According to Mosegaard & Sambridge (2002), Bayesian inversion can be summarized as

$$p(c|Z) = v_c \times p(Z|c) \times p(c) \tag{15}$$

where $p(c\,|\,Z)$ is posterior probability density function of model parameters c; $p(Z\,|\,c)$ is a likelihood function of parameters c; $p(c)$ is prior probability density function of parameters c; and v_c is a normalization constant. We assumed that the prediction errors were normally distributed and uncorrelated, and calculated the likelihood function, $p(Z\,|\,c)$, as

$$p(Z|c) = v_L \times \exp\left\{-\sum_{j=1}^{2}\sum_{i=1}^{k}\frac{(Z_{i,j} - X_{i,j})^2}{2\sigma_{i,j}^2}\right\} \tag{16}$$

where $Z_{i,j}$ is total soil C reported by IGBP-DIS (j = 1) or soil microbial biomass reported in Xu *et al*. (2013) (j = 2) in the ith gridcell, $X_{i,j}$ is total soil C or microbial biomass C simulated by the models at a corresponding gridcell; $\sigma_{i,j}^2$ is the variance of a jth measurement at ith gridcell (the standard deviation of each measurement at the ith grid cell is assumed to be 30% of the C pool value at that gridcell as in Hararuk *et al*., 2014); k is the total number of gridcells; and v_L is a constant.

We assigned minimum and maximum values to the parameters and used an adaptive Metropolis (AM) algorithm (Haario *et al*., 2001), a Markov Chain Monte Carlo method, to sample from the posterior parameter distributions. When assigning prior parameter ranges, we were guided by the literature, hypothesis testing, and sampling efficiency. For instance, activation energy ranges were assigned within the ranges reported in Fang & Moncrieff (2001) and Fang *et al*.(2006); CUE sensitivities were assigned between 0 and maximum reported in Devêvre & Horwáth (2000). We tested whether SOC decomposition was sensitive to soil clay content (Eqns 4 and 13) by assigning an initial sensitivity to 0 and a broad prior range; the same hypothesis testing was performed for litter lignin content. Broad prior ranges decreased parameter acceptance rate making the data assimilation algorithm inefficient; therefore, when literature data were not available, we assigned the range that yielded reasonable acceptance rates and was ±20–30% of the value reported in Allison *et al*. (2010).

We generated a parameter chain by running the AM algorithm in two steps: a proposing step and a moving step. In the proposing step, a new parameter set c^{new} was generated from a previously accepted parameter set c^{k-1} through a proposal distribution ($c^{new} \mid c^{k-1}$). In the moving step, a probability of acceptance $P(c^{k-1} \mid c^{new})$ was calculated as in (Marshall *et al*., 2004):

$$P(c^{k-1}|c^{new}) = \min\left\{1, \frac{p(Z|c^{new})p(c^{new})}{p(Z|c^{k-1})p(c^{k-1})}\right\} \quad (17)$$

The value of $P(c^{k-1} \mid c^{new})$ was then compared with a random number U from 0 to 1. The parameter set c^{new} was then accepted if $P(c^{k-1} \mid c^{new}) \geq U$, otherwise c^k was set to c^{k-1}.

The AM algorithm required an initial parameter covariance matrix, which we generated from test runs of 50 000 simulations for 2-pool and 4-pool microbial models, assuming a uniform proposal distribution as in Xu *et al*. (2006):

$$c^{new} = c^{k-1} + r \times \frac{c^{max} - c^{min}}{D} \quad (18)$$

where c^{max} and c^{min} are upper and lower parameter limits, and r is a random number between −0.5 and 0.5, and D = 5. We constructed a covariance matrix C_0 on the basis of the test run and modified the proposal step to be

$$c^{new} = N\left(c^{k-1}, C_k\right) \quad (19)$$

$$C_k = \begin{cases} C_0 & k \leq k_0 \\ S_d \mathrm{Cov}(c_0, \ldots, c_{k-1}) & k > k_0 \end{cases} \quad (20)$$

where $k_0 = 2000$; $S_d = 2.38\sqrt{8}$ for the 2-pool model and $S_d = 2.38\sqrt{15}$ for the 4-pool model (Gelman *et al*., 1996).

We made five parallel runs (each run containing 500 000 iterations) starting at dispersed initial points in the parameter space. All parameter and likelihood chains converged to the same frequency distributions. During a simulation, we equated Eqns (1–2) or (6–9) (depending on the model) to zero and solved the model for the two (or four) carbon pool sizes at each gridcell (the semi-analytical spin-up approach; Xia *et al*., 2012). We assumed that the spatial relationships of total soil C and microbial biomass C with environmental factors would represent the temporal relationships and that year-to-year changes in soil C pools were close to zero – an approach previously used by Ise & Moorcroft (2006) and Smith *et al*. (2013). We discarded the first half of the simulations (the burn-in phase) and confirmed the convergence the second half using Gelman–Rubin diagnostics (Gelman & Rubin, 1992).

Nonlinearities typically present in biological models, as well as a lack of data constraints, often mean that posterior model parameter estimates may not be independent of each other: Our posterior parameter probability distributions may be correlated if the effects of changes of any one parameter on model predictions is at least partly dependent on the values for other model parameters. High correlations indicate high confidence about the relationship between parameters but without necessarily indicating high confidence (narrow parameter ranges) about the values of the individual parameters that are correlated. In such cases, it is advisable to use the joint probability distributions when using the model to make predictions rather than individual parameter estimates. Such correlations can also be used to identify modeled processes that need additional data to estimate parameter values. To detect such processes and the degree to which our parameter estimates are independent of each other, we calculated parameter correlations from posterior parameter distributions (sample size = 10 000).

Forward model runs and stability analyses

To evaluate the consequences of our parameter uncertainties for uncertainties in soil C feedbacks under climate change, we ran the calibrated microbial models forward, driving them with a climate change scenario (increasing CO_2 and temperatures), and sampling from the joint posterior parameter distributions obtained in the previous analysis (sampling from the joint parameter distributions avoids potential problems with correlated parameters when making predictions). We used the Community Earth System Model (CESM) output for the Representative Concentration Pathway 8.5 (RCP8.5) experiment (specifically, the simulated temperature and soil C influx) to drive soil C pools. The CESM model output was provided as part of the Coupled Model Intercomparison Project Phase 5 (CMIP5) and was accessed from http://pcmdi9.llnl.gov. Under RCP 8.5, CESM simulated a 3.5 K increase in mean global temperature, and an atmospheric CO_2 increase to 1150 ppm, by the year 2100 (Keppel-Aleks *et al*., 2013). We first used the 2006–2010 data to generate initial pools using the semi-analytical model spin-up approach (Xia *et al*., 2012) and then ran the microbial models forward in time to the year 2100, generating soil C feedbacks to the changing climate. We

then compared the model predictions to the feedbacks predicted by a calibrated conventional model (Hararuk *et al.*, 2014).

Once, we ran the forward simulations, it became evident that we needed to test the microbial models for their intrinsic propensity to generate oscillations in carbon pool sizes, given that soil carbon pool sizes at local to global scales occasionally exhibited multiyear oscillations. We investigated the intrinsic propensity of the carbon pools to oscillate following the techniques described in Svirezhev (2002) and Wang *et al.* (2013): (1) Using maximum-likelihood parameter values, we calculated the Jacobian matrix for both microbial models at each gridcell, indicating the sensitivity of the system state to small perturbations (these small perturbations can come from any source); (2) we then calculated the eigenvalues of the Jacobians when the soil pool sizes were at equilibrium, where complex conjugate eigenvalues with negative real parts indicated an oscillatory return to equilibrium while only real negative eigenvalues indicated a monotonic return to equilibrium (eigenvalues with positive real parts indicated an unstable equilibrium but these did not occur in our study); (3) we estimated the period of oscillations from the imaginary component of an eigenvalue as $p = \frac{2\pi}{i}$, where p was the oscillation period, and i was the imaginary component of an eigenvalue; (4) we calculated the approximate times required to damp the oscillations as $t \approx (-r)^{-1}$, where r was the real part of an eigenvalue (negative if the models were convergent). Note that the calculated estimates of oscillation period and decay rate may not accurately reflect the values produced in forward simulations; the estimates are only valid for the dynamics following small perturbations to the stable equilibrium, whereas our forward simulations represent continuous and directional changes to the stable equilibrium. However, we include the eigenvalue-derived oscillation periods and decay rates here because they correspond well to those observed in simulations (as also noted by Wang *et al.*, 2013), and they also provide intuitive meaning to the contrasting eigenvalues calculated for different parameter sets and regions of the world.

Results

Performance of the microbial models after calibration

After calibration, the microbial models explained a higher fraction of the variation in the SOC data than the conventional model (Fig. 2b). Additionally, the microbial models had lower spatial RMSE than the calibrated conventional model. As indicated by RMSE's and r^2, the two microbial models produced similar SOC distributions; therefore, we make further total SOC fit comparisons using a conventional model and one of the microbial models, giving the illustrations for both microbial models in the Supplementary Information.

Microbial models performed notably better than the calibrated conventional model in terms of soil C prediction in the low-temperature regions and in the regions with small soil C inputs, as indicated by reduction in the magnitudes of model residuals (Figs 2c,d and S2). These differences in SOC predictions were caused by the differences in the SOC residence times because the soil C pools are determined by the soil C inputs and the residence times (Luo *et al.*, 2003) and the former were identical for conventional and microbial model formulations. In the conventional models, the spatial patterns of SOC residence times are determined mainly by temperature (Todd-Brown *et al.*, 2013b), whereas in microbial models residence times are controlled by both temperature and SOC inputs (mediated by microbial biomass change, Fig. 1b and c). Fresh C input stimulates microbial biomass growth, which increases the rate of old SOC decomposition (the priming effect, Kuzyakov *et al.*, 2000; Fontaine *et al.*, 2004, 2007); therefore, the microbial models simulated lower SOC residence times than the conventional model in the areas with high SOC input (Fig. S3). In the regions with low SOC input (e.g. shrublands and tundra), SOC residence times in the microbial models were higher than predicted by the conventional model. This was due to the nonlinearity of substrate limitation in the microbial models (Eqns 1–2 and 6–8) (the conventional model assumed a linear effect of substrate limitation on the microbial activity; Parton *et al.*, 1993), as well as the dependency of residence times on microbial biomass.

Both the 2-pool and the 4-pool microbial models explained ~30% of the spatial variability in observed microbial biomass C after calibration (Fig. 3). The biome with the largest residuals was deciduous needleleaf forest: Microbial biomass was underpredicted there on average by 550 g C m^{-2}. The largest overpredictions were in evergreen needleleaf forest and evergreen broadleaf forest: on average 50 g C m^{-2}. The patterns of model residuals in Fig. 3b and c suggested that the model formulations did not capture the decomposition processes in deciduous needleleaf forests.

Posterior parameter distributions and correlations

Many parameters were constrained by assimilating total SOC and microbial biomass C data into the two microbial models (Fig. 4). The parameters most constrained were the microbial death rate; the temperature sensitivity of microbial carbon use efficiency; the baseline microbial carbon use efficiency; activation energy; the lignin effect on the half-saturation constant of SOC decomposition; and the clay effect on the rate of the SOC decomposition. The unconstrained parameters were associated with the substrate limitation of C pool decomposition, the temperature sensitivity of microbial DOC uptake, and the dynamics of the enzyme pool. This indicates the need for assimilation of C flux data to constrain these parameters because, according to the

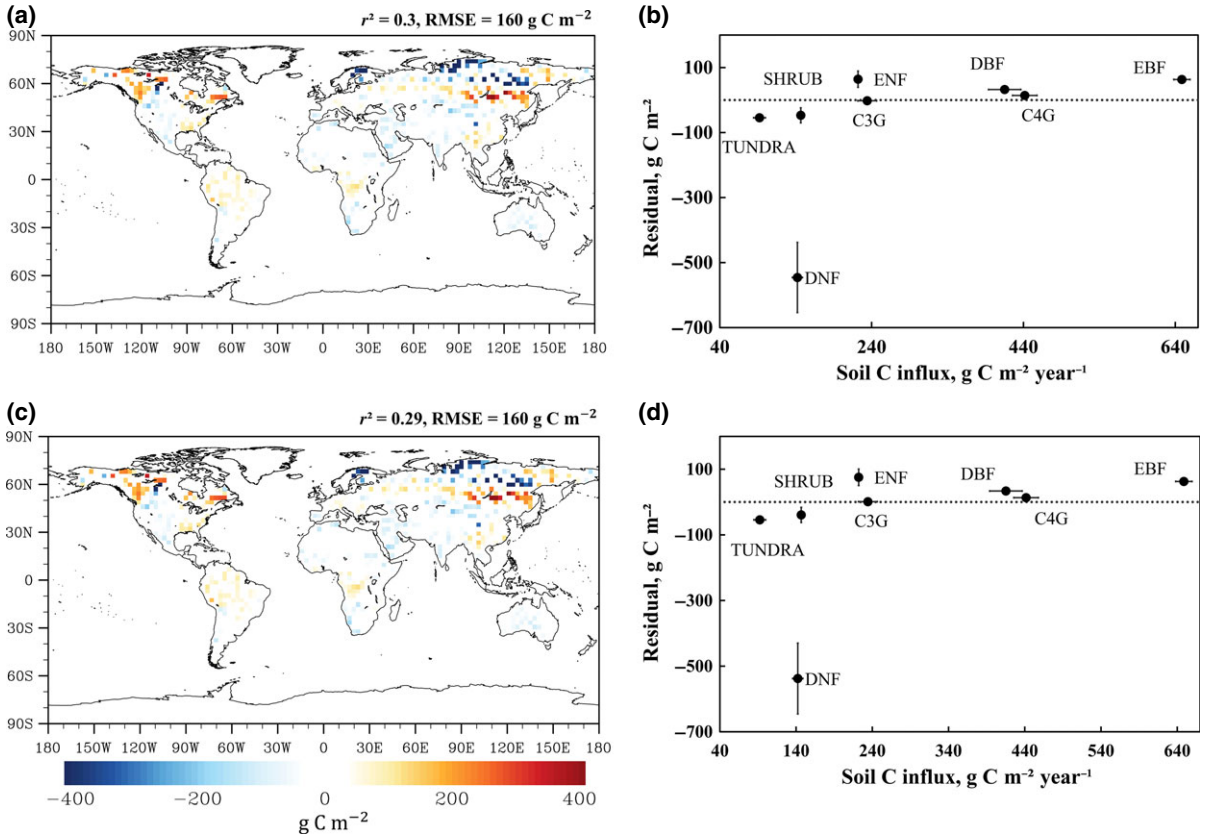

Fig. 3 Spatial distribution of model residuals for microbial biomass simulated by calibrated 2-pool (a, b) and 4-pool (c, d) microbial models. Pixels used for model calibration are not shown. C4G – C4 grasslands; DBF – deciduous broadleaf forest; DNF – deciduous needleleaf forest; EBF – evergreen broadleaf forest; ENF – evergreen needleleaf forest; SHRUB – shrublands.

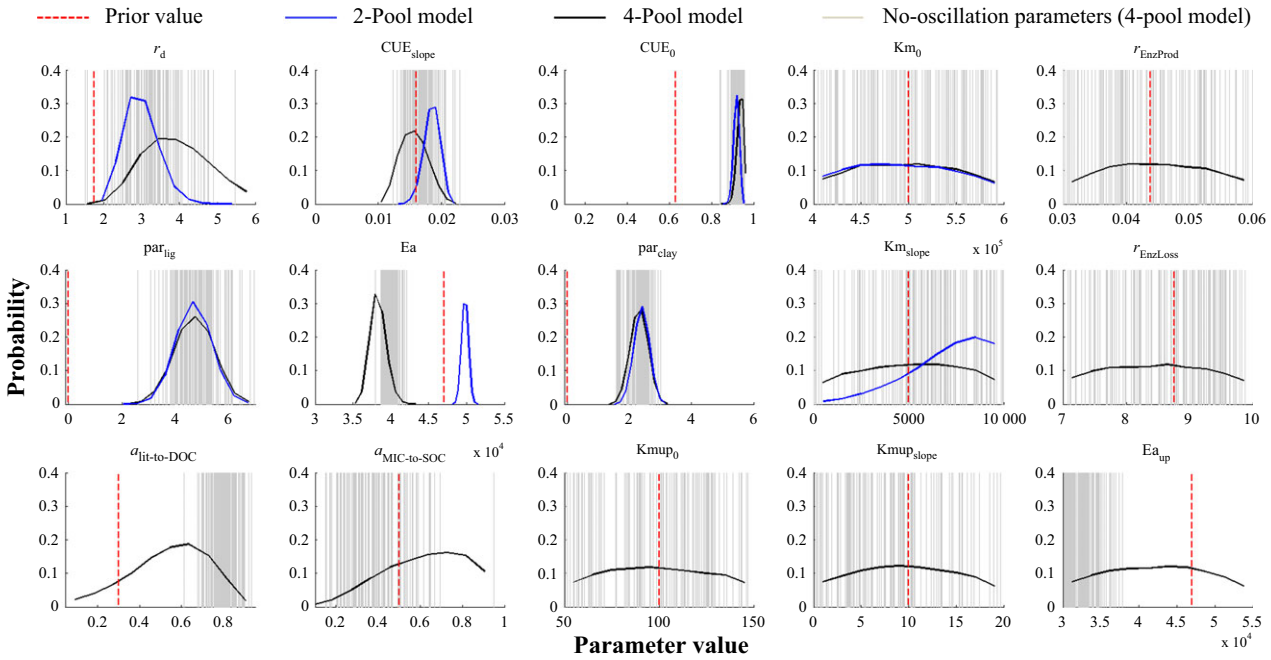

Fig. 4 Posterior probability density functions of the microbial models' parameters. See Models in Materials and methods or Table S1 for parameter definitions.

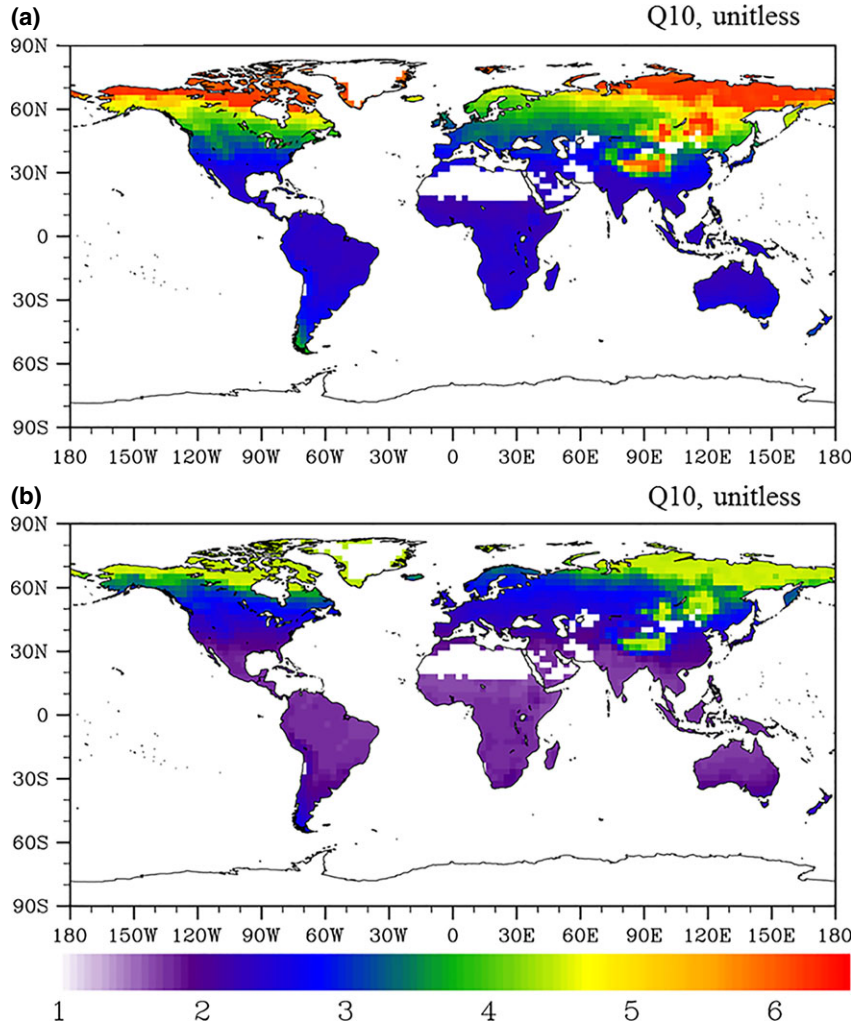

Fig. 5 Spatial distribution of the temperature sensitivities of SOC turnover rates in the 2-pool microbial model (a) and 4-pool microbial model (b). Constrained Q10 for the conventional model was constant and equal to 1.86.

model Eqns (1–2) and (6–8), flux data would contribute additional information on the temperature sensitivity and substrate limitation of microbial activity.

The lignin and clay effects on SOC decomposition, originally not included in the model, were significantly larger than zero and converged to the same values in both microbial models (Fig. 4). No observations were available to evaluate the lignin regulation for the half-saturation constant; however, our estimates for par_{clay} were close to the range calculated from observations. Observed par_{clay} ranged from 1.94 (Wang *et al.*, 2003) to 3.02 (Müller & Höper, 2004), whereas the 95% confidence intervals (CI) for the 2-pool model were 1.99–2.8 and 1.86–2.78 for the 4-pool model (Table S1). The temperature sensitivity of microbial carbon use efficiency was also estimated to be larger than zero and was 0.016–0.02°C^{-1} for the 2-pool model and 0.012–0.019°C^{-1} for the 4-pool model. For the 4-pool model, these were close to the observed range of 0.01–0.014°C^{-1} (Devêvre & Horwáth, 2000; Steinweg *et al.*, 2008; Frey *et al.*, 2013).

We calculated temperature sensitivities (Q10′s) of heterotrophic respiration for both microbial models (Fig. 5a and b) because the models had differing structures and parameters associated with soil responses to temperature. We calculated Q10 as the ratio of respiration after one day of a 10°C temperature increase from mean annual temperature to equilibrium respiration at mean annual temperature. The lower temperature dependency of microbial use efficiency and lower activation energy of SOC decomposition in the 4-pool model than in the 2-pool model led to lower temperature sensitivities in the 4-pool model. Both microbial models simulated spatially variable Q10′s, unlike the conventional model which assumed Q10 was constant across space, with higher values in the low-temperature

Table 1 Parameter correlations in the 2-pool microbial model

Parameters	r_d	CUE_{slope}	CUE_0	Km_0	Km_{slope}	par_{lig}	E_a	par_{clay}
r_d	1.00							
CUE_{slope}	**−0.75**	1.00						
CUE_0	**0.93**	**−0.61**	1.00					
Km_0	−0.07	0.08	−0.06	1.00				
Km_{slope}	0.33	−0.43	0.30	−0.04	1.00			
par_{lig}	−0.10	0.11	−0.07	−0.04	0.11	1.00		
E_a	**−0.63**	0.46	**−0.60**	−0.48	−0.27	−0.36	1.00	
par_{clay}	0.18	−0.23	0.16	0.01	−0.06	−0.19	−0.36	1.00

significance of bold values indicates strongly correlated parameters.

regions and lower Q10′s in the high-temperature regions – a widely reported pattern indicative of temperature acclimation of microbial activity (Luo *et al*., 2001; Chen & Tian, 2005; Peng *et al*., 2009). This acclimation was caused by substrate limitation and the sensitivity of CUE to temperature (Fig. S4). High CUE facilitated an increase in the microbial pool, and as respiration was proportional to microbial biomass, the large microbial pool produced a strong feedback to elevated temperature. In addition to temperature, microbial biomass was regulated by the substrate: The temperature sensitivity decreased if there was an insufficient amount of available substrate for the maximum potential biomass increase for the given temperature (Fig. S4), a phenomenon also observed in the field (Hartley *et al*., 2007).

Of 28 parameter pairs in the 2-pool model, only five were strongly correlated (Table 1): microbial death rate was negatively correlated with the activation energy of SOC decomposition, positively correlated with baseline carbon use efficiency, and negatively correlated with the temperature dependency of carbon use efficiency. The correlated parameters regulated C influx and outflux in the microbial pool, emphasizing that we did not have influx or outflux data to separate those parameters. Baseline CUE and the CUE dependency on temperature were negatively correlated, which was expected due to the CUE equation formulation (Eqns 3 and 10). Activation energy of SOC decomposition was negatively correlated with baseline CUE, and as the two parameters had the opposing effects on regulating microbial C uptake, we attributed this correlation to lack of the observed data to separate the two processes.

Of 105 parameter pairs in the 4-pool model, only four pairs were correlated (Table 2). As in the 2-pool model, microbial death rate was positively correlated with the baseline CUE and negatively with the degree of temperature dependency of CUE; two latter parameters were also negatively correlated with each other. The activation energy of SOC decomposition was negatively correlated with the fraction of dead microbes transferred to the SOC pool, which was probably also caused by the absence of C flux data.

SOC responses to climate change

We ran the models forward in time under the RCP8.5 climate change scenario to illustrate the differences in soil C responses between conventional and microbial models. The initial soil C pool sizes differed between the conventional and microbial models (Fig. 6a–c): the calibrated microbial models predicted a higher global soil C content than the calibrated conventional model (by 180–200 Pg C), and the predictions of all soil C cycle model formulations fell within the 95% CI of the observed SOC content (890–1660 Pg C; Todd-Brown *et al*., 2013b). We illustrate soil C sensitivities to climate change as cumulative relative soil C changes because the initial pool sizes were different among the models. In the microbial models, soil C was more sensitive to climate change than in the conventional model: The 2-pool and 4-pool microbial models simulated between 8% and 11% cumulative SOC losses, and the conventional model simulated a 2.5% SOC loss after 95 years of climate change.

Under the RCP8.5 scenario, after 95 years of climate change, the mean annual global temperatures increased by 3.5°C (Keppel-Aleks *et al*., 2013), and the annual global soil C input increased by 20% (Fig. S5). As mentioned earlier, the SOC residence times in the microbial models were regulated not only by temperatures, but also by SOC input. The temperature increases, along with the increase in SOC inputs, stimulated microbial biomass growth and therefore increased SOC decomposition. In the conventional model, fresh C input resulted in soil C accumulation and temperature increases caused SOC loss. Increases in both temperature and SOC input led to a higher relative SOC loss in the microbial models than in the conventional model, despite the lower Q10′s of the former.

Among the three models, the 2-pool microbial model had the highest uncertainty in SOC responses under

Table 2 Parameter correlations in the 4-pool microbial model

Parameters	r_d	$r_{EnzProd}$	$r_{EnzLoss}$	$a_{lit\text{-}to\text{-}DOC}$	$a_{MIC\text{-}to\text{-}SOC}$	CUE_{slope}	CUE_0	$Kmup_{slope}$	$Kmup_0$	Km_{slope}	Km_0	par_{lig}	Ea_{up}	E_a	par_{clay}
r_d	1.00														
$r_{EnzProd}$	0.00	1.00													
$r_{EnzLoss}$	−0.01	−0.01	1.00												
$a_{lit\text{-}to\text{-}DOC}$	−0.42	−0.08	0.04	1.00											
$a_{MIC\text{-}to\text{-}SOC}$	−0.09	−0.01	0.02	−0.40	1.00										
CUE_{slope}	**−0.89**	−0.01	0.01	0.47	0.09	1.00									
CUE_0	**0.93**	0.02	−0.01	−0.39	−0.09	**−0.81**	1.00								
$Kmup_{slope}$	0.00	−0.01	−0.02	0.02	−0.01	0.00	0.00	1.00							
$Kmup_0$	−0.02	−0.03	0.01	−0.01	0.01	0.02	−0.02	0.03	1.00						
Km_{slope}	0.01	0.02	0.02	0.00	0.00	0.01	0.00	0.01	−0.03	1.00					
Km_0	0.01	0.02	0.00	0.34	0.05	−0.02	0.00	0.02	0.00	−0.02	1.00				
par_{lig}	0.02	−0.02	−0.01	−0.07	0.01	−0.02	0.02	0.01	−0.01	−0.03	0.15	1.00			
Ea_{up}	−0.01	0.02	0.00	0.00	0.00	0.00	−0.01	0.00	0.01	−0.02	0.00	0.01	1.00		
E_a	−0.44	0.39	−0.22	0.49	**−0.65**	0.39	−0.41	0.00	0.00	−0.21	−0.05	−0.23	0.00	1.00	
par_{clay}	0.04	0.02	−0.01	0.15	0.04	−0.03	0.03	0.01	0.00	0.02	−0.04	−0.17	0.03	−0.14	1.00

significance of bold values indicates strongly correlated parameters.

the RCP8.5 scenario (Fig. 6e), ranging between 1.5% and 12.5% loss after 95 years of climate change. This larger uncertainty does not appear to have resulted from larger parameter uncertainties in the 2-pool model than in the 4-pool model (Fig. 4), rather it appears to be a consequence of the oscillatory dynamics predicted by the 2-pool model. The 2-pool model exhibited oscillations in the dynamics of SOC over the 95-year period (Fig. S6), consistent with the linear analysis of the model which implied that the average global period of oscillations was 57 years (Fig. 7a) with the range of around 38 years and taking about 140 years required to damp the oscillations.

The uncertainty in soil C feedbacks in the 4-pool model was not as high as in the 2-pool model because the globally averaged periods of oscillations in the former were larger and convergence times were shorter: 1.9 and 116.5 years, with the times required to damp the oscillations 17.5 h and 4.4 years, respectively (Fig. 7b,e,c and f). Moreover, in posterior parameter space for the 4-pool model, there were parameter combinations that did not produce oscillations (Fig. 4, gray lines; the linearized model about the steady state had four negative real eigenvalues). The distribution of nonoscillatory behaviors throughout parameter space as well as preliminary analyses of model behaviors using the posterior parameter estimates (omitted for brevity) strongly implies the existence of thresholds in activation energies, temperature dependence of microbial carbon use efficiency, and partitioning of the C flux from litter to DOC beyond which the model begins to exhibit oscillatory dynamics. High ratios of oscillation period to their convergence times and existence of sample runs with no oscillations led to lower uncertainty in SOC feedbacks in the 4-pool model than in the 2-pool model.

Discussion

Progress in simulating SOC dynamics

Most carbon cycle models have been simulating SOC decomposition as a first-order decay process with the variability in the models' parameterization causing substantial uncertainties in SOC predictions (Todd-Brown *et al.*, 2013b). In recent years, studies have been focusing on reducing this uncertainty by applying data-model fusion techniques to constrain the model parameters (Ise & Moorcroft, 2006; Zhou *et al.*, 2009; Smith *et al.*, 2013; Hararuk *et al.*, 2014). Still, data-constrained conventional models leave plenty of unexplained variation in the data (Smith *et al.*, 2013; Hararuk *et al.*, 2014).

Microbial models have shown potential to simulate global SOC stocks better than conventional models (Wieder *et al.*, 2013), and our results illustrated that

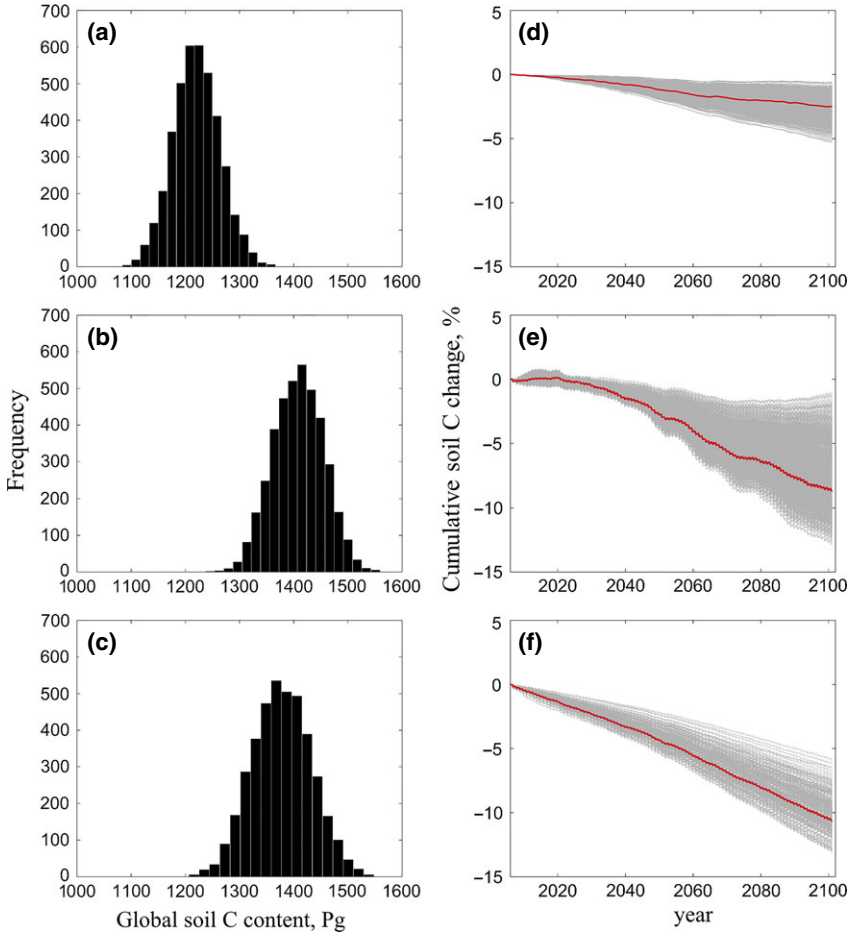

Fig. 6 Global frequency distributions of total SOC pools produced by a calibrated CENTURY-type model (a, from Hararuk *et al.*, 2014), 2-pool and 4-pool microbial models (b and c, respectively), and the cumulative SOC changes under RCP8.5 scenario (d–f, red lines are maximum-likelihood cumulative changes, and gray lines are sample runs, representing uncertainty).

SOC representation by microbial models was better than that by conventional models once both model formulations were calibrated against data. However, better representation of SOC distribution may not indicate that microbial models are better than conventional models at predicting SOC dynamics. For example, microbial models have been criticized for producing unrealistic SOC dynamics, such as oscillations in SOC pools (Wang *et al.*, 2013). While there are reports of oscillations in microbial biomass and respiration observed in the incubation studies (Lloyd *et al.*, 1982; Cenciani *et al.*, 2008), the phenomenon needs to be investigated further prior to being considered true for natural ecosystems. In our study, oscillations in the SOC pools were present in both the 2-pool and 4-pool microbial models (Fig. 7). However, they were less prominent in the 4-pool model due to short oscillation periods and rapid convergence times (Fig. S6). Moreover, many parameter combinations in the 4-pool model did not generate oscillations in the SOC pools at all (Fig. 4). This indicates that the 4-pool microbial model suffers less from unrealistic mathematical artifacts than the 2-pool model. Future studies should investigate whether the 4-pool microbial model could be used to construct a better performing alternative to conventional models for simulating soil carbon dynamics, providing improved predictions of carbon stocks and fluxes while avoiding mathematical artifacts (such as through only using regions of parameter space that do not lead to intrinsically generated oscillations).

Uncertainties in future SOC dynamics

Future climate change can be either mitigated or worsened by terrestrial feedbacks. For the past two decades, soils have been slowing the rate of atmospheric CO_2 increase by sequestering over 0.45 Pg C per year (Pan *et al.*, 2011); however, it remains uncertain whether soils will continue to sequester C under changing temperatures and CO_2 concentrations. CMIP5 models are

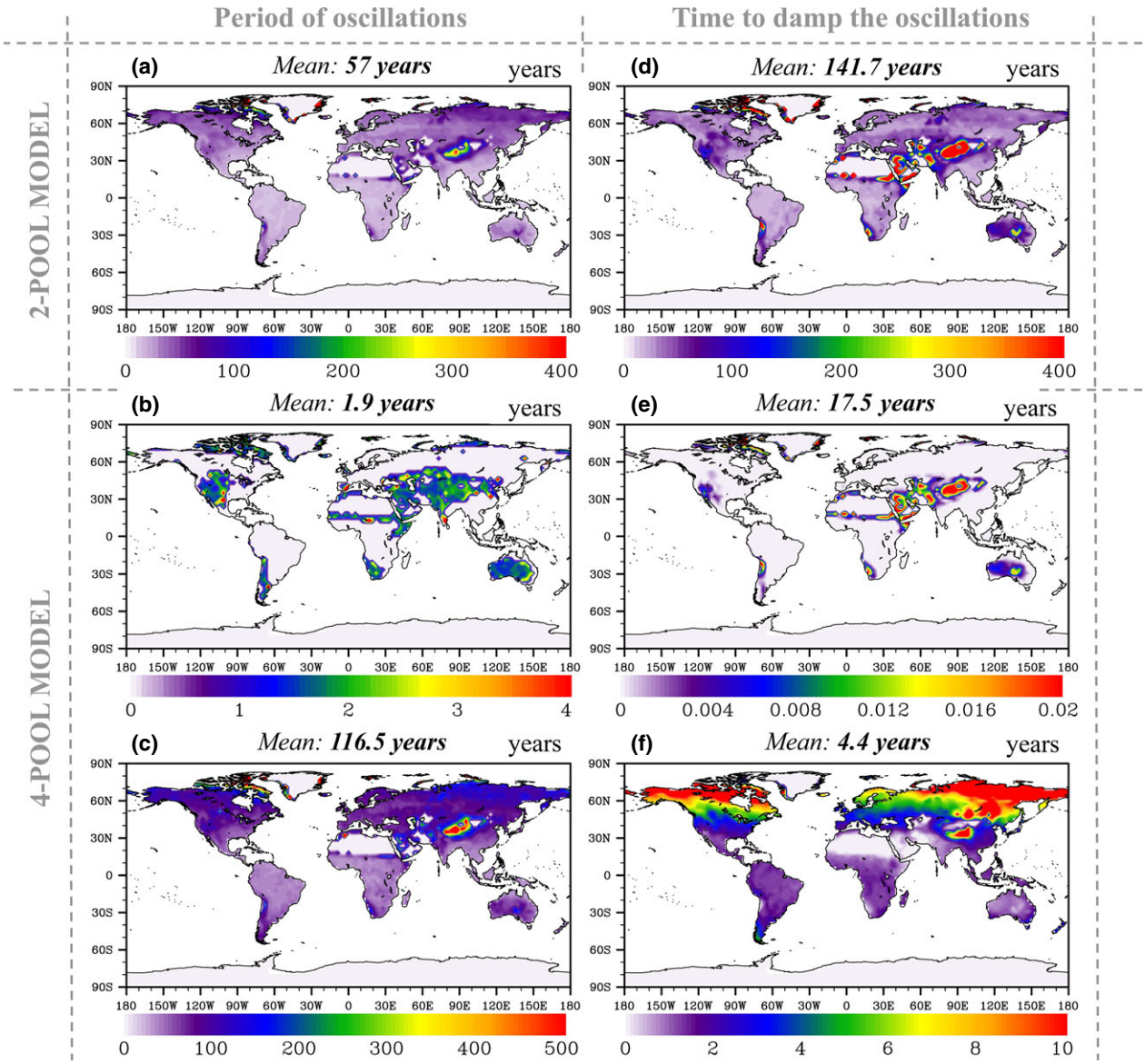

Fig. 7 Spatial distribution of oscillations in the 2-pool (a) and 4-pool (b, c) models, and time required to damp the oscillations (d–f). The 4-pool model had higher ratios of the oscillation period to convergence time than the two-pool model, which indicated that oscillations will be damped early in their evolution. This figure was produced using only those parameter sets that led to oscillations in the 4-pool model. Note that the oscillation properties were estimated by eigenvalue analysis not by direct simulation (see Materials and methods).

not unanimous in simulating soil responses to the RCP8.5 scenario, ranging from a 7% loss of soil C after 95 years of climate change to a 22.6% gain in SOC (Todd-Brown *et al.*, 2013a). A data-constrained conventional model predicted a 1–6% loss in soil C after 95 years (Hararuk *et al.*, 2014), which was within the CMIP5 range. The calibrated microbial models, however, simulated much stronger negative response of soil C to RCP8.5 scenario than conventional models: a 2–13% loss (2-pool model) and 6–13% (4-pool model) loss in SOC. This study also narrowed the range of potential SOC responses to RCP8.5 scenario presented in Wieder *et al.* (2013) and suggested that outcomes with zero carbon loss and carbon gain, projected by most CMIP5 models, were not likely.

One of the key differences between conventional and microbial model formulations is the effect of SOC input on SOC residence times: In the conventional model, the SOC input rate does not affect residence time, whereas in the microbial models, the SOC input is negatively related to residence time (Fig. S3). This is because increases in the SOC input rate result in increases in the microbial pool size, which in turn decreases the total SOC residence time (priming effect) as the turnover rate

of SOC is proportional to the size of microbial pool (Eqns 2, 8 and 9). According to the microbial model structure, the magnitude of the priming effect and its uncertainty depends on a multitude of parameters regulating microbial pool size and activity. Of those parameters, most attention has been paid to CUE and its temperature sensitivity, which has been shown to generate large uncertainties in SOC feedbacks (Allison *et al.*, 2010; Wieder *et al.*, 2013). Our data-constrained CUE sensitivities had more narrow ranges than those explored in the literature and therefore were unlikely to be the largest source of uncertainty.

The microbial turnover rate had wide posterior ranges in both models (2.33–3.78 per years for the 2-pool model and 2.59–5.36 per years for the 4-pool model), and being the only parameter that determined depletion of microbial pool, was likely the cause of most of the uncertainties in the predictions of soil carbon, particularly for the 4-pool model. Another poorly constrained parameter that could have caused substantial uncertainties in the 4-pool model predictions was $a_{lit\text{-}to\text{-}DOC}$, controlling the proportional allocation of C influx among soil carbon pools. The destination of fresh carbon input has previously been reported to cause large uncertainties in the SOC feedbacks (Allison *et al.*, 2010).

Future improvements

Although the microbial models (particularly, the 4-pool model) demonstrate the best model performance in simulating the global SOC distribution to date, there is much room for improvement. The data indicate that 48% of variability in global total SOC and 70% of variability in global microbial biomass distributions remains unexplained, implying that substantial improvements in predictive ability can be achieved. The extent to which this potential for improvement can be realized will be influenced by advances in our understanding of the SOC processes and properties as well as improvements in our understanding of how much predictive ability is realistically achievable (Luo *et al.*, 2014). The nature of environmental limitation on microbial dynamics also remains uncertain: CUE has been reported to decrease with increasing temperature; however, this relationship may not be linear as the rate of CUE decrease diminishes under high temperatures (Devêvre & Horwáth, 2000; Wetterstedt & Ågren, 2011; Frey *et al.*, 2013). CUE probably also depends on organic matter quality (Frey *et al.*, 2013; Sinsabaugh *et al.*, 2013). Lastly, a recent report suggests that changes in microbial activity with temperature are not caused by changing CUE, but are a result of temperature sensitivity of microbial turnover rates (Hagerty *et al.*, 2014). Together, these uncertainties highlight that more studies are needed to investigate controls over CUE, microbial death rates, and the partitioning of soil C input between DOC and SOC pools.

The oscillatory dynamics observed in microbial models by Wang *et al.* (2013) and in this study (Figs 7 and S6) are rarely observed in nature and need to be further investigated. If these turn out to be artifacts of the model assumptions then they can be avoided by calibrating the microbial models with an additional constraint so that parameters generating oscillations are discarded. This would be impossible for the 2-pool model as it inevitably produces oscillations, but could be plausible for the 3-pool model of Wang *et al.* (2013); which only exhibited oscillations in litter and microbial biomass, but not in soil organic carbon. A perhaps more robust, but potentially challenging, approach would be to modify the microbial formulations to scale from the microscale processes to the macro-scale dynamics in a way that properly accounts for the complexities of soil microbial community dynamics. A similar argument was made by Franks (2009) when highlighting the occurrence of the same problems when modeling marine microbial dynamics in a similarly simplistic way.

In this study, we relied on the assumption that soil carbon is in equilibrium. This assumption may introduce inaccuracies in parameter estimation: Zhou *et al.* (2013) illustrate that estimating parameters under the equilibrium assumption results in the underestimation of soil carbon residence times in comparison with parameter estimates made without the equilibrium assumption. This implies that parameter estimates in this study may underestimate the potential soil carbon sink and overestimate soil carbon losses under a climate change scenario. Improvements in the parameter estimates can be achieved by assimilating dynamic observations into the model, such as soil respiration, as well as changes in pool sizes.

The correlations between the parameter distributions inferred for our microbial models were largely associated with parameters controlling the rates of C influx and outflux from the soil. The presence of parameter correlations emphasized the potential benefits of incorporating data on the dynamics of soil organic carbon when performing parameter calibration. Obtaining accurate estimates of these parameters could be essential for accurate estimation of SOC changes under climate change. The relatively low number of parameter correlations in both models implied that the level of complexity in our models was appropriate given that we only used contemporary SOC and microbial C stocks as data constraints. Adding more parameters and processes while keeping the number of datasets fixed could lead to the models becoming overparame-

terized, in which the more parameter-rich models make worse predictions against holdout or independent data. Indeed, an information theory metric such as AIC would imply that the 2-pool microbial model is preferable to the 4-pool model on the basis of parsimony because it achieves a very similar predictive accuracy with fewer parameters. However, given that model performance was assessed using holdout data, we conclude that the additional parameters in the 4-pool model are not making its predictive accuracy worse. Moreover, the 4-pool model clearly does not include the unrealistic dynamics of the 2-pool model, and so we conclude that it is preferable to the 2-pool model in terms of the realism of its predictions.

Another assumption we made was that standard deviations of total and microbial soil carbon were 30% of their value at each gridcell. It is possible that the use of observational error estimates would alter our parameter estimates, with consequences for our predictions, but unfortunately such data are not available. In addition, the datasets we used probably did not contribute much information to model parameters regulating short time scale processes because soil pools are largely the result of the long-term processes. Time series of soil heterotrophic respiration could be used to estimate these unconstrained parameters.

As we show in this study, data assimilation facilitates fairly accurate (50% variation explained) SOC stocks simulation by microbial models, and the following structural analysis identifies the 4-pool model as preferable due to the absence of prominent oscillations in SOC dynamics. Microbial models simulated spatially variable temperature sensitivities of SOC decomposition with the variability pattern similar to the observed one, whereas the conventional model assumed Q10 was constant. Data assimilation also narrowed the global ranges for the temperature sensitivities of CUE, established a data-informed range for soil clay content limitation on C uptake rate, as well as the range for substrate quality limitation of C uptake by microbes. The parameter values obtained in this study can be used in future modeling efforts, as well as the initial values for further parameter optimization.

Acknowledgements

This study was financially supported by US Department of Energy, Terrestrial Ecosystem Sciences grant DE SC0008270 and US National Science Foundation (NSF) grant DBI 0850290, EPS 0919466, DEB 0743778, DEB 0840964, EF 1137293. We acknowledge the World Climate Research Programme's Working Group on Coupled Modelling, which is responsible for CMIP, and we thank the climate modeling groups Community Earth System Model Contributors for producing and making available their model output. For CMIP, the .S. Department of Energy's Program for Climate Model Diagnosis and Intercomparison provides coordinating support and led development of software infrastructure in partnership with the Global Organization for Earth System Science Portals. Lastly, we thank two anonymous reviewers for their helpful advice, which led to improvements in this paper.

References

Allison SD, Wallenstein MD, Bradford MA (2010) Soil-carbon response to warming dependent on microbial physiology. *Nature Geoscience*, **3**, 336–340.

Carney KM, Hungate BA, Drake BG, Megonigal JP (2007) Altered soil microbial community at elevated CO_2 leads to loss of soil carbon. *Proceedings of the National Academy of Sciences USA*, **104**, 4990–4995.

Cenciani K, Freitas SDS, SaM C, Airoldi C (2008) Microbial enzymatic activity and thermal effect in a tropical soil treated with organic materials. *Scientia Agricola*, **65**, 674–680.

Chen H, Tian H-Q (2005) Does a general temperature-dependent Q10 model of soil respiration exist at biome and global scale? *Journal of Integrative Plant Biology*, **47**, 1288–1302.

Cusack DF, Chou WW, Yang WH, Harmon ME, Silver WL, The Lidet T (2009) Controls on long-term root and leaf litter decomposition in neotropical forests. *Global Change Biology*, **15**, 1339–1355.

Delire C, Foley JA, Thompson S (2003) Evaluating the carbon cycle of a coupled atmosphere-biosphere model. *Global Biogeochemical Cycles*, **17**, 1012.

Devêvre OC, Horwáth WR (2000) Decomposition of rice straw and microbial carbon use efficiency under different soil temperatures and moistures. *Soil Biology and Biochemistry*, **32**, 1773–1785.

Falkowski P, Scholes RJ, Boyle E *et al.* (2000) The global carbon cycle: a test of our knowledge of earth as a system. *Science*, **290**, 291–296.

Fang C, Moncrieff JB (2001) The dependence of soil CO_2 efflux on temperature. *Soil Biology and Biochemistry*, **33**, 155–165.

Fang C, Smith P, Moncrieff JB, Smith JU (2005) Similar response of labile and resistant soil organic matter pools to changes in temperature. *Nature*, **433**, 57–59.

Fang C, Smith P, Smith JU (2006) Is resistant soil organic matter more sensitive to temperature than the labile organic matter? *Biogeosciences*, **3**, 65–68.

Fontaine S, Bardoux G, Abbadie L, Mariotti A (2004) Carbon input to soil may decrease soil carbon content. *Ecology Letters*, **7**, 314–320.

Fontaine S, Barot S, Barre P, Bdioui N, Mary B, Rumpel C (2007) Stability of organic carbon in deep soil layers controlled by fresh carbon supply. *Nature*, **450**, 277–280.

Franks PJS (2009) Planktonic ecosystem models: perplexing parameterizations and a failure to fail. *Journal of Plankton Research*, **31**, 1299–1306.

Freibauer A, Rounsevell MDA, Smith P, Verhagen J (2004) Carbon sequestration in the agricultural soils of Europe. *Geoderma*, **122**, 1–23.

Frey SD, Lee J, Melillo JM, Six J (2013) The temperature response of soil microbial efficiency and its feedback to climate. *Nature Climate Change*, **3**, 395–398.

Gelman A, Rubin DB (1992) Inference from iterative simulation using multiple sequences. *Statistical Science*, **7**, 457–511.

Gelman A, Roberts G, Gilks W (1996) Efficient metropolis jumping hules. *Bayesian Statistics*, **5**, 599–608.

German DP, Marcelo KRB, Stone MM, Allison SD (2012) The Michaelis-Menten kinetics of soil extracellular enzymes in response to temperature: a cross-latitudinal study. *Global Change Biology*, **18**, 1468–1479.

Global Soil Data Task Group (2000) *Global Gridded Surfaces of Selected Soil Characteristics (IGBP-DIS).* [Global Gridded Surfaces of Selected Soil Characteristics (International Geosphere-Biosphere Programme - Data and Information System)]. Data set, Oak Ridge National Laboratory Distributed Active Archive Center, Oak Ridge, TN.

Haario H, Saksman E, Tamminen J (2001) An adaptive Metropolis algorithm. *Bernoulli*, **7**, 223–242.

Hagerty SB, Van Groenigen KJ, Allison SD *et al.* (2014) Accelerated microbial turnover but constant growth efficiency with warming in soil. *Nature Climate Change*, **4**, 903–906.

Hararuk O, Xia J, Luo Y (2014) Evaluation and improvement of a global land model against soil carbon data using a Bayesian MCMC method. *Journal of Geophysical Research: Biogeosciences*, **119**, 403–417.

Hartley IP, Heinemeyer A, Ineson P (2007) Effects of three years of soil warming and shading on the rate of soil respiration: substrate availability and not thermal acclimation mediates observed response. *Global Change Biology*, **13**, 1761–1770.

House JI, Colin Prentice I, Le Quéré C (2002) Maximum impacts of future reforestation or deforestation on atmospheric CO_2. *Global Change Biology*, **8**, 1047–1052.

Ise T, Moorcroft PR (2006) The global-scale temperature and moisture dependencies of soil organic carbon decomposition: an analysis using a mechanistic decomposition model. *Biogeochemistry*, **80**, 217–231.

Keppel-Aleks G, Randerson JT, Lindsay K *et al.* (2013) Atmospheric carbon dioxide variability in the community earth system model: evaluation and transient dynamics during the twentieth and twenty-first centuries. *Journal of Climate*, **26**, 4447–4475.

Kucharik CJ, Foley JA, Delire C *et al.* (2000) Testing the performance of a dynamic global ecosystem model: Water balance, carbon balance, and vegetation structure. *Global Biogeochemical Cycles*, **14**, 795–825.

Kuzyakov Y, Friedel JK, Stahr K (2000) Review of mechanisms and quantification of priming effects. *Soil Biology and Biochemistry*, **32**, 1485–1498.

Lawrence PJ, Chase TN (2007) Representing a new MODIS consistent land surface in the Community Land Model (CLM 3.0). *Journal of Geophysical Research: Biogeosciences*, **112**, G01023.

Lloyd D, Edwards SW, Fry JC (1982) Temperature-compensated oscillations in respiration and cellular protein content in synchronous cultures of Acanthamoeba castellanii. *Proceedings of the National Academy of Sciences*, **79**, 3785–3788.

Luo Y, Wan S, Hui D, Wallace LL (2001) Acclimatization of soil respiration to warming in a tall grass prairie. *Nature*, **413**, 622–625.

Luo Y, White LW, Canadell JG *et al.* (2003) Sustainability of terrestrial carbon sequestration: A case study in Duke Forest with inversion approach. *Global Biogeochemical Cycles*, **17**, 1021.

Luo Y, Keenan TF, Smith M (2014) Predictability of the terrestrial carbon cycle. *Global Change Biology*, doi:10.1111/gcb.12766.

Marshall L, Nott D, Sharma A (2004) A comparative study of Markov chain Monte Carlo methods for conceptual rainfall-runoff modeling. *Water Resources Research*, **40**, W02501.

Mosegaard K, Sambridge M (2002) Monte Carlo analysis of inverse problems. *Inverse Problems*, **18**, R29–R54.

Müller T, Höper H (2004) Soil organic matter turnover as a function of the soil clay content: consequences for model applications. *Soil Biology and Biochemistry*, **36**, 877–888.

Oleson KW, Dai Y, Bonan G *et al.* (2004) Technical description of the community land model (CLM). NCAR Tech. Note NCAR/TN-461+ STR, 173.

Oleson K, Niu G, Yang Z *et al.* (2007) *CLM3. 5 documentation*. National Center for Atmospheric Research, Boulder, CO, 34 pp.

Oleson K, Niu G, Yang Z *et al.* (2008) Improvements to the Community Land Model and their impact on the hydrological cycle. *Journal of Geophysical Research*, **113**, G01021.

Pan Y, Birdsey RA, Fang J *et al.* (2011) A large and persistent carbon sink in the world's forests. *Science*, **333**, 988–993.

Parton WJ, Scurlock JMO, Ojima DS *et al.* (1993) Observations and modeling of biomass and soil organic matter dynamics for the grassland biome worldwide. *Global Biogeochemical Cycles*, **7**, 785–809.

Peng S, Piao S, Wang T, Sun J, Shen Z (2009) Temperature sensitivity of soil respiration in different ecosystems in China. *Soil Biology and Biochemistry*, **41**, 1008–1014.

Qian T, Dai A, Trenberth KE, Oleson KW (2006) Simulation of global land surface conditions from 1948 to 2004. Part I: forcing data and evaluations. *Journal of Hydrometeorology*, **7**, 953–975.

Schimel JP, Weintraub MN (2003) The implications of exoenzyme activity on microbial carbon and nitrogen limitation in soil: a theoretical model. *Soil Biology and Biochemistry*, **35**, 549–563.

Sinsabaugh RL, Manzoni S, Moorhead DL, Richter A (2013) Carbon use efficiency of microbial communities: stoichiometry, methodology and modelling. *Ecology Letters*, **16**, 930–939.

Smith MJ, Purves DW, Vanderwel MC, Lyutsarev V, Emmott S (2013) The climate dependence of the terrestrial carbon cycle, including parameter and structural uncertainties. *Biogeosciences*, **10**, 583–606.

Steinweg JM, Plante AF, Conant RT, Paul EA, Tanaka DL (2008) Patterns of substrate utilization during long-term incubations at different temperatures. *Soil Biology and Biochemistry*, **40**, 2722–2728.

Svirezhev YM (2002) Simple spatially distributed model of the global carbon cycle and its dynamic properties. *Ecological Modelling*, **155**, 53–69.

Taylor BR, Parkinson D, Parsons WFJ (1989) Nitrogen and lignin content as predictors of litter decay rates: a microcosm test. *Ecology*, **70**, 97–104.

Taylor KE, Stouffer RJ, Meehl GA (2011) An overview of CMIP5 and the experiment design. *Bulletin of the American Meteorological Society*, **93**, 485–498.

Todd-Brown KEO, Randerson JT, Hopkins F *et al.* (2013a) Changes in soil organic carbon storage predicted by Earth system models during the 21st century. *Biogeosciences Discussions*, **10**, 18969–19004.

Todd-Brown KEO, Randerson JT, Post WM, Hoffman FM, Tarnocai C, EaG S, Allison SD (2013b) Causes of variation in soil carbon simulations from CMIP5 Earth system models and comparison with observations. *Biogeosciences*, **10**, 1717–1736.

Vance ED, Chapin FS (2001) Substrate limitations to microbial activity in taiga forest floors. *Soil Biology and Biochemistry*, **33**, 173–188.

Wang WJ, Dalal RC, Moody PW, Smith CJ (2003) Relationships of soil respiration to microbial biomass, substrate availability and clay content. *Soil Biology and Biochemistry*, **35**, 273–284.

Wang G, Post WM, Mayes MA (2012) Development of microbial-enzyme-mediated decomposition model parameters through steady-state and dynamic analyses. *Ecological Applications*, **23**, 255–272.

Wang YP, Chen BC, Wieder WR *et al.* (2013) Oscillatory behavior of two nonlinear microbial models of soil carbon decomposition. *Biogeosciences Discussions*, **10**, 19661–19700.

Waring B, Weintraub S, Sinsabaugh R (2014) Ecoenzymatic stoichiometry of microbial nutrient acquisition in tropical soils. *Biogeochemistry*, **117**, 101–113.

Wetterstedt JÅM, Ågren GI (2011) Quality or decomposer efficiency – which is most important in the temperature response of litter decomposition? A modelling study using the GLUE methodology. *Biogeosciences*, **8**, 477–487.

Wieder WR, Bonan GB, Allison SD (2013) Global soil carbon projections are improved by modelling microbial processes. *Nature Climate Change*, **3**, 909–912.

Xia JY, Luo YQ, Wang YP, Weng ES, 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 Discussions*, **5**, 1259–1271.

Xu T, White L, Hui D, Luo Y (2006) Probabilistic inversion of a terrestrial ecosystem model: Analysis of uncertainty in parameter estimation and model prediction. *Global Biogeochemical Cycles*, **20**, GB2007.

Xu X, Thornton PE, Post WM (2013) A global analysis of soil microbial biomass carbon, nitrogen and phosphorus in terrestrial ecosystems. *Global Ecology and Biogeography*, **22**, 737–749.

Zhao M, Running SW (2010) Drought-induced reduction in global terrestrial net primary production from 2000 through 2009. *Science*, **329**, 940–943.

Zhou T, Shi P, Hui D, Luo Y (2009) Global pattern of temperature sensitivity of soil heterotrophic respiration (Q10) and its implications for carbon-climate feedback. *Journal of Geophysical Research: Biogeosciences*, **114**, G02016, doi:10.1029/2008JG000850.

Zhou T, Shi P, Jia G, Luo Y (2013) Nonsteady state carbon sequestration in forest ecosystems of China estimated by data assimilation. *Journal of Geophysical Research: Biogeosciences*, **118**, 1369–1384.

Supporting Information

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

Table S1. Model parameter description, prior ranges, posterior ranges, and their convergence indices.
Figure S1. Comparison of CLM-CASA soil C input with a 7-year average of MODIS data (Zhao & Running, 2010).
Figure S2. Changes in the magnitudes of the residuals between the calibrated microbial 2-pool (a, c) and 4-pool (b, d) models and conventional model.
Figure S3. Distribution of soil C residence times in the environmental space simulated by a calibrated conventional model (a) and microbial models (b).
Figure S4. Change of temperature sensitivity of soil C residence time with respect to microbial carbon use efficiency and substrate limitation.
Figure S5. Relative change in SOC input in response to the RCP8.5 scenario simulated by CESM.
Figure S6. Forward model runs under RCP8.5 climate change scenario for the calibrated 2-pool and 4-pool models in boreal (a, b), temperate (c, d) and tropical (e, f) climate zones.